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:	Wednesday, January 16, 2013

Work Order No.: NRC-3036 Pages 1-325

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

UNITED STATES OF AMERICA

NUCLEAR REGULATORY COMMISSION

+ + + + +

ADVISORY COMMITTEE ON REACTOR SAFEGUARDS

(ACRS)

+ + + + +

RELIABILITY AND PRA SUBCOMMITTEE

+ + + + +

WEDNESDAY

JANUARY 16, 2013

+ + + + +

ROCKVILLE, MARYLAND

+ + + + +

The Subcommittee met at the Nuclear Regulatory Commission, Two White Flint North, Room T2B1, 11545 Rockville Pike, at 8:30 a.m., John W. Stetkar, Chairman, presiding.

COMMITTEE MEMBERS:

JOHN W. STETKAR, Chairman

J. SAM ARMIJO, Member

DENNIS C. BLEY, Member

MICHAEL L. CORRADINI, Member

DANA A. POWERS, Member

HAROLD B. RAY, Member

JOY REMPE, Member

STEPHEN P. SCHULTZ, Member

WILLIAM J. SHACK, Member

DESIGNATED FEDERAL OFFICIAL:

JOHN LAI

NRC Staff

Sean Peters, RES

Jing Xing, RES

Erasmia Lois, RES

James Chang, RES

Others

Gareth Parry, ERIN Engineeering

Harry Liao, Sandia Lab

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

C-O-N-T-E-N-T-S (CONTINUED)

P R O C E E D I N G S

11:08 a.m.

CHAIRMAN STETKAR: [presiding] John Lai is the Designated Federal Official for this meeting.

The Subcommittee will hear the latest developments of HRA methods and its applications in response to Commission's SRMM-062010. The staff will also update the Subcommittee on the Halden Reactor Project related to HRA and the NRC HRA Data Collection Project. We will hear presentations from the NRC staff and its contractors and EPRI.

There will be a phone bridge line. To preclude interruption of the meeting, the phone will be placed on listen-in mode during the presentations and Committee discussions.

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

The entire meeting will be open to public attendance. 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 a part of this 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 throughout the meeting room when addressing the Subcommittee. The participants should first identify themselves and speak with sufficient clarity and volume so that they may be readily heard.

Our recorder is apparently caught in traffic. We are being recorded here. And so, all of the preliminary information will be on the record. We may have to take a break when the recorder shows up, so that they can get set up, just to sort of alert you to that situation.

I believe one or two of our members may need to make some statements.

MEMBER BLEY: Yes, Mr. Stetkar, I have a personal conflict of interest with some of the work that was done here. So, in the work that I was involved in, I will have to not participate in the Committee discussions.

CHAIRMAN STETKAR: Thank you.

MEMBER REMPE: And I appear to perhaps

have some organizational conflict of interest issues, but it is my understanding a waiver is being prepared to address that concern. But, until then, probably I need to recuse myself from some aspects of the discussion.

CHAIRMAN STETKAR: Thank you.

And with that, we will now proceed with the meeting and Sean Peters from the staff.

MR. PETERS: I am Sean Peters. I am the Branch Chief for the Human Factors and Liabilities Branch in the Office of Research.

And I would like to thank the ACRS and our team of engineers here from Sandia, Idaho, ERIN Engineering, the Electric Power Research Institute, Paul Scherrer Institute, and the University of Maryland.

What you will see here is a culmination of years of research in the Office of Research promoting and advancing the state-of-the-art and the state-of-practice in HRA throughout the agency. These products that we are going to present today were originally Commission-directed activities to promote HRA, to promote the state-of-the-art in HRA. Over time, we have gotten some of these products into user need-driven bases.

But what we are doing is we are coming in here as part of the Commission direction to interact and seek comments and questions from the ACRS. We hope what will happen through this Subcommittee interaction is give us the opportunity to make more useful and more scientifically-sound HRA products.

And so, with that -- I don't want to take up too much time -- I would like to let Jing go ahead and give the presentation. This is Jing Xing, one of our senior researchers in the Human Factors Reliability Branch.

CHAIRMAN STETKAR: Sean, something I was thinking about -- and, Jing, maybe we can discuss this at the end of the meeting -- as you mentioned, this project has been underway for quite some time. The original SRM actually was issued to ACRS to work with the staff. So, we are integrally-involved with this project.

We have not had a full Committee briefing on this topic in a very long time, meaning I don't remember when the last one was. We may want to think, if we are getting to a point where things are reaching a fairly high level of maturity, we may want to think about scheduling that. And you may want to think about whether you want a formal letter from the ACRS.

I will just float that. You may want to discuss it among yourselves today.

MR. PETERS: I think since we have worked with the ACRS, I think that a formal letter would be a route we would like to go in the future.

CHAIRMAN STETKAR: Okay.

MR. PETERS: At least an ACRS endorsement of the methodology.

CHAIRMAN STETKAR: Yes. So, for our planning purposes, we may not need to set a formal date today at the meeting, but we should probably start thinking about sometime here in the first calendar quarter, or it depends on your schedules and desires.

MR. PETERS: Sure. And I think some of this can depend upon the feedback we get today.

CHAIRMAN STETKAR: Yes.

MR. PETERS: We are at that level of maturity.

(Laughter.)

CHAIRMAN STETKAR: We are always hopeful of that, but --

MR. PETERS: Yes, exactly.

CHAIRMAN STETKAR: -- sometimes we are ignored.

Thanks.

With that, Xing, it is yours.

MS. XING: Okay. Thanks, ladies and gentlemen.

So, this morning we will talk about the SRM. By this time, we all refer to this project as IDHEAS, which is the name of the new HRA method we have been developing over the last couple of years.

So, for this morning, I will first will give an overview of our staff's overall response to the SRM. And then, Gareth, on behalf of EPRI, will briefly talk about the EPRI's involvement in this project.

And after that, we will brief on the methodology with development and HRA methods specifically for internal at-power event.

So, that is our morning's agenda.

And here are the contributors for this project over the five years. If you like, I can read over the list. Okay, I see.

So, just a quick review what was the research easiest answer, research need back five years ago. So, the majority of the HRA work has been done for our internal procedural events. We have many measures developed for that. So, each of these are very developed in the results from method-to-method and analyst-to-analyst variability.

So, back a few years ago, our team considered that we needed integrated methods to reduce the variability. That is what we were approaching so far.

And the precise internal procedure event, there are many other considerations in NRC for HRA practices. So, the question still also exists to are the measures applicable and adequate for other domains, and does each application have to have its own measure.

So, it was desired that we would to develop a generic methodology for all the applicable cases.

So, with this need, the goal of our project was to develop a new HRA methodology to reduce the variabilities and also apply to all HRA applications. I can judge from your face this is a basic skill.

And the requirements for the new methodology is basic requirement, and we would like it to conform to the existing PRA/HRA standard and HRA good practices. So, we can have a smooth transition.

And we would like to retain and integrate the strengths of the existing methods.

As for the new development, though, we want to enhance the capabilities to address the key weaknesses in the state of practice, particularly for those areas the cause the variability.

And we want to have a state-of-the-art technical basis and be generic enough for all the HRA applications.

So, we have learned the major weaknesses from the HRA benchmarking studies. Most of the methods do not have an explicit cognitive basis on why and how human fails to perform tasks, and, essentially, also all the methods need a stronger basis.

And the methods either lack adequate guidance for performing qualitative analysis or they don't have an adequate interface to use the qualitative analysis results for quantification of human error probabilities.

And the methods have inadequate guidance on how to assess and use performance-shaping factors, which is a quantification we rely on.

So, IDHEAS, this new method, it tries to improve basically and make improvements in these areas.

So, this slide is a little bit messy, but you can focus on the big boxes. So, this slide shows our strategic approach over the last five years.

As a first step, we developed our cognitive basis for human error analysis by synthesizing state-of-the-art technology and scientific literature to understand why humans make errors and how.

With that basis and the integrating stress that exist in HRA methods, such as CBDT or ATHEANA good practices, we developed the IDHEAS method for internal at-power events. We want to build something that in the moment that we have more confidence and knowledge first.

So, after that process of building this specific method for internal events and using an extended cognitive basis, and also integrated with HRA practices in other domains. So, starting at the beginning of 2012, we have been developing the generic IDHEAS methodology for all the NPP applications.

So, what you are looking at here, you can think this is a generic methodology. In a way, it is an extension for the internal at-power method of review

And another way to look at the internal

at-power method is an example implementation of the generic methodology. So, with this same process, in the future, if we need it, it will compute more and from an implementation-specific domain

So, coming to the end of this project, we are going to build it, provided there is a triplet of products. So, first, the cognitive basis can be used with the intent to use it for HRA, of course. But, beyond that, it can also provide a good technical basis for our general human performance and the human factors engineering work at the NRC.

And the next product is a generic IDHEAS methodology. It is intended to use for all kinds of human events in NPP. So, such as the Level-3 PRA low-power shutdown, and some other external events.

And the IDHEAS methodology for internal at-power events is specifically intended to be used for internal at-power events PRA.

So, where we are in the project now, the first product I think we have completed and have the report finalized and we are going to have it published.

CHAIRMAN STETKAR: Jing, has that gone out for any, or will it go out for any, external review?

MS. XING: The staff report has been

reviewed by many people internally and externally in 2011. So, do we not intend to go another round of external review.

CHAIRMAN STETKAR: That is interesting. Because everybody who worked on the report has reviewed the report, you don't need to send it out to get any other opinions?

MS. XING: We have several of our NRC staff, Human Factors staff, who did not work on this project, and they reviewed the report. And, also, our international collaborators, like folks at Halden, they have reviewed the report. So, we have lots of input there.

MEMBER SCHULTZ: You mentioned the report timing of 2011-2012. When was the draft, when was the report finalized for this review that you are describing now?

MS. XING: The report at this moment, it is supposed to be, it is presumably finalized. In the beginning of 2012, we took the comments from the last ACRS meeting and from our reviewers, then made extensive revision to the report. And the report now has been technically edited. So, unless we receive more comments, it is ready to go.

CHAIRMAN STETKAR: The reason I ask, there are two reasons. One is that the last time this Subcommittee saw that report, it had developed to the extent where I believe performance-influencing factors had been reasonably well-defined for the detection and noticing macrocognitive function, and only that one. None of the other macrocognitive functions had effectively any work done on them at that time.

So, since detection and noticing is only one of the five, that is essentially only 20 percent of the ultimate product. Now that is a bit pejorative because the product is really the entire fundamental research.

By the way, for the record, I should note that Harold Ray has joined us.

The recorder is not here, Harold, but we are being recorded, just so you know what is going on.

MEMBER RAY: I was here when you --

CHAIRMAN STETKAR: Were you?

MEMBER RAY: -- went around the table. Yes.

MEMBER SCHULTZ: I can vouch for that.

MEMBER RAY: You came in with a cup of coffee, in fact, sat down.

(Laughter.)

CHAIRMAN STETKAR: Good morning to

everyone.

(Laughter.)

Anyway, back to the NUREG world -- I'm sorry, Harold.

MEMBER RAY: You read my name off.

CHAIRMAN STETKAR: Did I read your name?

MEMBER RAY: Yes, you most certainly did.

MEMBER SCHULTZ: We are going to see how this fits into the model.

(Laughter.)

MEMBER RAY: It's all right, John. It happens to all of us.

CHAIRMAN STETKAR: As I get older, the probability is approaching or the frequency is sort of, you know, "N" per day where "N" is greater than 1.

Back to the seriousness, we, as a Subcommittee, have not really seen the entire report, and we are not having a presentation on that report. I, indeed, read through it. I have about 18 pages of questions regarding it, but we are not going through it. I needed to read through it, so that I had a fundamental understanding of the final version of the report or the nearly-final version of the report.

I had some questions about completeness, especially some of the last tasks that you worked on, communications, supervision, those issues, which are obviously the last efforts that were placed. And the reason I asked about sending it out in its final form for formal feedback is that it is not clear to me what level of review and feedback you have had in the sense of someone being able to objectively sit down, read the entire report, and think about completeness, integration, consistency among the attributes for each of the macrocognitive functions, each of the cognitive mechanisms, and the mapping of the performance-influencing factors. Because that process is a fundamental basis for what we are going to be hearing more about today, the mechanics, if you will.

So, I don't know. It is obviously the staff's report, but I am a bit surprised to hear that it wasn't going out in its final form for any type of review.

MEMBER BLEY: I guess I would toss in on that one. The goal of all this under the SRM was to have something new that would be broadly appreciated and accepted. And since that is the core that underlies everything that is going on, a really well-documented external review seems to me like it would be very useful.

I know there have been international

collaborators. In your, I guess, acknowledgments, you had some real experts in this area participating and reviewing and commenting. But once it is all in a package, it just feels like we have missed something if we don't get a real peer review on it.

MEMBER SCHULTZ: And that would be a peer review by those that are, in fact, going to be applying this. You would like to have a compendium of reviewers that have said, this is a great product and -- and -- I would apply it.

MS. XING: Okay. I will try to address all of your comments together.

So, first of all, for John's comment on completeness, as a Project Manager in the new office of that report, I would say that report is far from outcome to achieve the completeness goal. So, I would more frame it as an initial effort. We tried our best to collect the synthesized information on why humans can make mistakes and the underlying cognitive mechanism. You know, what ways they would fail, which is different causes or we offer approximate causes. That is the focus of that report.

And also, the cognitive function that is addressed, it only addressed the five main cognitive functions. Among those five main cognitive functions, we addressed the very likely decisionmaking and the communication. We didn't address teamwork and the supervision.

Part of the reason was at that time, our team's focus was fully internal. With goals, we are developing this technical basis for internal-event HRA. So, errors, we didn't do a comprehensive related study with the ones that we think are less important.

So, anyway, there is a limitation of resources and time for us to obtain a completeness impact model. But, in 2012, in order to develop this generic methodology for IDHEAS, we did a lot more combination basis study to expand those areas like I just talked, decisionmaking, supervision, teamwork. As you can think about it, these will be critical, either a very complex external event or something like Fukushima.

So, the major part of those extensions, I kind of tried to boil them down in one element here stated in the report called combinative error causal tree.

So, in the long-run and from what I hear, I think what would be really useful would be develop Volume II of the Combination Basis Report to include in this new edition. And, also, PRA, we only address at a very high level in the Early Combination Basis, in the literature review report. And we spent a lot of work in 2012 to dig down in the literature in the PRA, which, actually, that is the basis for the generic methodology. So, one potential plan would be we should have another report in the long run to augment to, basically, the extension.

And for now, I think I can talk to my boss. My book has a project and everything saying that it will need another one before the current Technical Basis Report or versus we can wait for this for the future, when we have the document, as a result that we have external review.

CHAIRMAN STETKAR: Jing, at one level, I hear what you are saying, but at another level this project is now being characterized as coming to some fundamental level of maturity. I hesitate to use the word "closure," but some level of maturity, in the sense that specific methods have been developed, documented, and there is a proposal to start piloting them for use.

This report, this NUREG, is ostensibly the fundamental basis for those methods. If there is something missing here -- and I don't mean completeness in the sense of holistic everything must be perfect and absolutely comprehensive; that will never be achieved at that level of detail -- but it is, as it is written right now, a complete document. There are a couple of little areas saying further research should be done. But, indeed, all of the macrocognitive functions are addressed. There are cognitive mechanisms defined for the macrocognitive functions. There are performance-influencing factors defined as input to the cognitive mechanism.

So, it hangs together as a complete report, which, as I said, the Subcommittee has never seen that report, at least in terms of its being presented formally or discussed in a Subcommittee meeting.

Because it is such a fundamental basis -- I understand that you may in a research sense want to go on and expand and continue to examine things, but, indeed, at this snapshot in time, to support what is being characterized as a fairly-matured, developed methodology, this is what we are hanging our hat on. That is why I personally think that it is important to have the community, the technical community, have a chance to see this in its current form, which is, indeed, a complete report, and provide some feedback. Because if you don't get that, you are running a real risk that somebody is going to, then, take issue with some element of the methodology as being inadequate or not considering certain elements of human performance, which is a real risk, I think.

MS. XING: Yes, I fully understand your consideration. I mean, John already sent you all that report in ADAMS. So, I will certainly appreciate it if you can take a look at that report and provide us a comment.

In March, we plan, our project plan, we plan to start an external review for the method report. And in fact, if we can get all the comments back to incorporate the comments in time, we can start the external review to give this report to the reviewers. That is what they say as a technical basis and that they also see the report of the method of technical basis.

If we cannot make this timeframe, if we like another meeting or discussion on the Technical Basis Report, we can plan on that.

MR. PETERS: So, I think what she is saying is that we will look at getting this into our March external review process, based upon your recommendation. If there is a time that you would like us to set up another Subcommittee meeting --

CHAIRMAN STETKAR: That I think we should discuss among the Subcommittee. We can either do that now or at the end of the day.

But I think we have made the point. Quite honestly, as I was going through the ideas methodology step-by-step, looking at each element, I was using these pages -- these are Appendix B of NUREG-2114 -- to cross-check in terms of completeness and consistency. That essentially was my bible to make sure that the methodology seemed to make sense.

And I would hope that anyone reviewing the methodology, the IDHEAS methodology -- you know, you said you are going to ask for feedback on the methodology -- if they don't have that fundamental basis document and have familiarity with it, and, essentially, have had a chance to provide feedback on it, it is not clear how this fundamental process that you have elaborated here hangs together, because the methodology could have been dreamed from anything.

MEMBER REMPE: Is there any intent to use any aspect of this methodology in the Level-3 activity that is being launched or underway right now?

MR. PETERS: That is the intent of the project. As far as I understand with a Level-3 project, for the parts of the HRA that have already been performed by the utility, we are not going to be redoing with the IDHEAS methodology. We will be doing checks based upon the methodology they used. But for the areas that were not developed, the intent is to look at using this methodology to build out those areas.

The issue is, with the Commission direction being so old, we were given this task in the 2008-2009 timeframe, and that this methodology has reached a level of maturity and has been specifically built for a generic variety of applications, other HRA methodologies have not been built specifically for a Level 2 or any of the aspects of Level 3 that we can model. So, this is the only methodology that we know out there that has at least some of the technical basis for doing that development.

MEMBER REMPE: So, this isn't my area of expertise, but, again, having external review before you apply it for the Level 3 would seem like a good idea to me.

MR. PETERS: I think so, too, but just as a counterpart, we don't know of any HRA methodologies that have been externally reviewed for a Level-2 or -3 application. So, your choices are between one that has been built for that application and others that have not, and none of them have been reviewed internally. So, we are looking into doing an external review also associated with the Level 1.

MEMBER REMPE: Okay.

CHAIRMAN STETKAR: I have to say this in every HRA meeting, so I might as well say it now.

(Laughter.)

People are people. People don't care whether you have divided up a PRA into Level 1, Level 2, Level 3. People don't care that, when something happens in a power plant, it is an internal at-power event or a fire during shutdown. People are people.

So, HRA methods should be able to evaluate people performance regardless of how we, as PRA analysts, have decided to artificially dissect and reconstruct our view of the world. In that sense, I loathe the notion of HRA for at-power internal events, HRA for fire, HRA for seismic, HRA for Level 2, for shutdown, HRA for all of that. It is all human reliability analysis.

And in that sense, indeed, all of the HRA methods have been reviewed for Level 2 and 3 because all of the HRA methods purport to model human performance, period. They don't ever say this is only for when I use this particular procedure for this particular event.

MR. PETERS: But based on the assumptions they make, of course, there are applicability issues when you go out to other areas. For our methodology here, as John indicated, we took a human-centered approach, that we tried to model it from how does a human behave. And hence, that gives us applicability across all domains.

CHAIRMAN STETKAR: And that is why NUREG-2114 is so important, because it does not differentiate.

MR. PARRY: John?

CHAIRMAN STETKAR: Yes, Gareth?

MR. PARRY: Can I just make a comment that you are right in a sense that, if we had a complete HRA methodology, it should be applicable to everything. But the nature of the tasks that are involved and what we model in a Level-1 PRA could be significantly different from those in the Level 2. So, it depends whether the model we have developed and the technical basis. And as Jing said, some of the things that were not in 2114 are things like decisionmaking and coordination with management.

CHAIRMAN STETKAR: I'm sorry, it is now.

MR. PARRY: No, not in the sense that it

is innovative decisionmaking as opposed to procedure-driven decisionmaking, which is more prescriptive rather than knowledge-based.

CHAIRMAN STETKAR: Well, then -- hold on a second -- then --

MS. XING: Okay. Just to make a correction on that?

CHAIRMAN STETKAR: Okay.

MS. XING: When I said decisionmaking wasn't fully addressed in the current methodology, let's say, through the literature review, I list by myself, and I try to narrow down the decisionmaking component mechanism, around 20-some of them.

And in the report, let's say we can split those 20-plus mechanisms in three major areas. Relatively simple decisionmaking will have highly-defined simplistic and dynamic decisionmaking. This should be okay, the decisionmaking, among many key members.

So, because for the common base, at the time we developed that, our target had only focused for internal events. So, the literature revealed, that chapter, only focused the details in an already relatively-simple decisionmaking, not in clothing dynamic decisionmaking across a long period of time, different information coming up all the time.

And also, it did not include the information on distributed decisionmaking. Just to look at distributed decisionmaking, error-causing results.

So, I understand your point that humans are humans and that thought should be generic enough for all this. But the issue is, it is what Gareth said; the tasks are different. So, if you are going to develop a measure that addresses the universe of all the tasks, yes, we can do that, but the method will take like 10 volumes to address all the potential tasks, all the mechanisms.

Just to give you one example, when we developed the IDHEAS method for the internal at-power event, we identified 14 types of failure modes, and we believe these are good enough to cover the internal at-power operation.

And this year, when we worked on this general IDHEAS methodology for all the NPP applications, we so far got like -- I can't remember the number. Of course, right now it is still a dynamic number. I got like 40-plus different types of failure modes, which we could have all those failure modes in the internal model, but that work, there will be a disparity issue with the amount of effort for all users who go through this component list.

So, my understanding is our approach is we develop this generic IDHEAS methodology, which has all the foundations there to address humans are humans. And from that, for a specific application, we can select/extract the part for that particular type of test, make it easy to use. But that is the whole philosophy about --

CHAIRMAN STETKAR: And I understand all of that.

MS. XING: I know.

CHAIRMAN STETKAR: I'm coming back to NUREG-2114, which is the cognitive basis for regardless of how you are going to slice the pie for some user of a methodology. As I read through that document, I tried very carefully to see whether it seemed to be biased toward a procedure-centric view of the world or a PRA Level-1 internal-events-only view of the world. And with a few minor wording exceptions, I didn't at least get the impression that it had that bias.

So, I was, from what I am hearing this morning perhaps, naively believing that, indeed, that document was reasonably complete and, indeed, it does address decisionmaking, perhaps not in a research sense, a complete assessment of decisionmaking, but it seems to hang together. It seems to address the types of errors that we have seen in practice in not only the nuclear power industry, but in other industries. It makes reference to those types of errors.

And that is the sense that I am saying that, if that is the fundamental underpinning of all of the subcategories of a methodology, if you want to put them that way, it just seems to make sense that before we become irretrievably married to some excruciating detail in one or more of those subsets of the methodology, that we have some consensus among at least the technical community on that basic fundamental underpinning.

If some folks in the technical community take issue with some fundamental element of that cognitive psychology underpinning, we ought to try to address it, not in the sense of, well, it is not complete. We haven't looked at, you know, the paper that was published in 1937 at some conference or we haven't looked at some nuance of some other performance, but in terms of a treatment of the different ways people can make errors.

That is why I keep coming back to that.

That is why I kind of used that thing as I was going through the Level-1 internal-events, procedure-sort-of-oriented subset of the overall methodology.

So, anyway, I think we spent quite a bit of time on this. I think you got the message.

MR. PETERS: John, to respond, yes, we got the message and we will look into it. And just to reiterate what Jing said really early in the discussion, that I do have numerous cognitive psychologists on my staff who aren't part of the project who have weighed-in on the project. So, what I consider technical experts in cognitive psychology have looked through this. And we also have counterparts internationally who have looked through it.

But we will go take it back and see if there is anything that makes sense from this standpoint to get further peer review on it.

CHAIRMAN STETKAR: Before we close out, because we are not going to hear much about 2114, let me take a quick straw poll among the Subcommittee. Should we plan to get a Subcommittee presentation on 2114 in its final current form? We all have it. I mean, as Jing mentioned, it was distributed to us.

But, given the amount of effort that we

have to put in, the thousands of pages that we are asked to read every week, we naturally marshal resources.

MEMBER REMPE: Yes.

CHAIRMAN STETKAR: Okay.

MEMBER BLEY: Unless from what Jing said earlier, they have put together a lot more work and are going to issue a revision soon, then I think, yes, we ought to do that.

CHAIRMAN STETKAR: Offline let's work on that.

MEMBER RAY: John, yes.

CHAIRMAN STETKAR: Okay. Thank you.

MEMBER SCHULTZ: I agree.

CHAIRMAN STETKAR: And I will vote for Corradini and Shack.

(Laughter.)

Now we can get to slide 8, or whatever it is.

MS. XING: Oh, we are still on the same slide. So, just the status and for the generic methodology, because we just started to develop that in 2012, and we have a draft report. At this stage, I call it the compact. We would like to explore next year in Level 3 PRA. The reason I say explore it is because it is just a product of this, not a final product. So, we would like to see at least a test of this idea we put in this methodology. If it makes sense, then we revise the report and for formal testing, then, for external review. So, that is down the road there.

And we are in the momentary status for the IDHEAS methodology for internal at-power events. We have completed the report. And so, we believe that base methodology now, we are ready to put it for testing. And we already started some validation work for that method.

So, coming up -- actually, it is next week -- we are going to have an expert elicitation workshop for listening to IEEE. So, our goal for FY13 is to get this final report ready and put it for use.

CHAIRMAN STETKAR: I will wait, but, just briefly, the test and validation, are you going to pilot it on a specific, one or more specific types of applications?

MS. XING: Yes. We have a very comprehensive --

CHAIRMAN STETKAR: Okay. That is fine. We will get to it. I haven't looked forward in the slides yet.

MS. XING: Actually, I do plan to go to the details of the test and validation; just give you some quick idea.

EPRI is going to test the method in some trial applications.

CHAIRMAN STETKAR: Okay.

MS. XING: And our project team is going to do a test using the same scenario in the U.S. empirical study.

CHAIRMAN STETKAR: Okay.

MS. XING: And what is already going on, we are also piloting this method. April Whaley has been piloting this method with events she selected from our early HERA database. Right now, she is working on applying the method to Indian Point 2 steam generator and tube rupture event. So, that is very interesting feedback to us.

CHAIRMAN STETKAR: Good.

MS. XING: So, then, we have a number of-- we have been talking. I am glad Antonio has come back. So, I am going to talk on how to test this method in SDP and get their sense.

So, there is a bunch of other things going on.

CHAIRMAN STETKAR: Okay. Good. Good.

Thank you.

MS. XING: So, the challenge is to put all of these results together.

MEMBER SCHULTZ: And then, just to go back to where you were on the previous slide in the presentation, then this information associated with the, I will call it "testing and validation," this is what is also going to be included in the final report 2013?

MS. XING: Yes. Well, 80 percent.

MEMBER SCHULTZ: You have a final draft, December --

MS. XING: Yes. Right.

MEMBER SCHULTZ: -- 2012.

MS. XING: We are going to take this, the result, and put it in the final report. But the publication of the final report is not a stopper for more testing and validation.

MEMBER SCHULTZ: I understand.

MS. XING: Yes.

MEMBER SCHULTZ: Do you have a hard-stop date for publication?

MS. XING: A hard-stop date for publication is September 30, 2013.

MEMBER SCHULTZ: Okay.

MS. XING: So, that is why we planned to integrate the testing results and going for external review and all these things. I hope things work out as planned.

MEMBER SCHULTZ: Right.

MS. XING: Yes.

MEMBER SCHULTZ: But I am anxious to hear some more about the workshops and, also, the test application, the examples. It seems that it would be very valuable to integrate that into the report publication by September.

MS. XING: Thank you. I like that suggestion. In the report we gave to you, the December 17th report, we had an appendix which only has the workshop plan, the expert elicitation plan.

MEMBER SCHULTZ: Yes. I saw that.

MS. XING: So, hopefully, by the end of March, when we finish expert elicitation, summarize the results, we will replace that appendix with the actual process, not just the plan.

MEMBER SCHULTZ: If there is a place today to talk more about the workshops -- it may not be in your presentation -- but it would be valuable to the Subcommittee to understand what is happening there.

MS. XING: Yes, we probably don't have

time. But one possibility, since we are already running late, if we skip the generic methodology part, and I can talk about the workshop --

MEMBER SCHULTZ: Perhaps just to be thinking about it during the day, and sometime later, perhaps in the afternoon, to summarize in five minutes what is going to happen in the workshop and where you feel it is going to --

MS. XING: Okay. So, let's give you a five-minute summary of the workshop now.

MEMBER SCHULTZ: That will be fine.

MS. XING: How about that?

MEMBER SCHULTZ: I know it is in your backup slides. So, go ahead.

MS. XING: Because I am not sure we will have time later in the day.

CHAIRMAN STETKAR: No, that is fine. Go on.

MS. XING: So, we are using the formal, the process we use is the SHAAC process, which is a formal, structured expert elicitation process. So, by being formal, we have everything planned out ahead of time and have a procedure how to carry it out. And being structured, we have a different type of expert, too, with different responsibilities and rules.

And for being interactive, we are going to have two workshops that all the experts will have a face-to-face interaction.

And so, the SHAAC process is to elicit the state of knowledge, enter with everybody's knowledge there, and have the evaluators or technical integrators at the end to synthesize all the information.

So, therefore, on these slides, you can see we have three major stages. Well, we have many different detail stages. We have three major types of experts:

Data experts to compile whatever data we can find about HEP from simulation, from previous expert judgment, the results from previous expert judgments, the results from other HRA practices. So, by now, they have compiled a database to their best knowledge on what we have.

And resource experts will be men like the operator trainers, licensing examiners, and the human factors representatives. So, they will present their knowledge, best knowledge.

And we have a technical integration team. Those are the HRA analysts. So, they will take information from the data and the resource expert and make the HEP estimate.

And we have a technical expert leader to supervise all this activity.

So, we have three major stages for this activity. The preparation stage; we have the project plan, as you see in the appendix, and develop the database and develop the worksheets and the procedure for the workshops and training, which happened yesterday in the pilot here. We had, once more, a pilot here, and there will be another practice in the beginning of the next workshop.

So, we will have Workshop No. 1 next week, next Wednesday, where the data experts will present the model and data, and the resource experts will provide their knowledge and the initial assessment of the likelihood of the failure mode and the causes of the failure mode from operational aspects. And the integrators will question, will try to elicit more knowledge.

CHAIRMAN STETKAR: Jing, let me interrupt you quickly.

MS. XING: Yes.

CHAIRMAN STETKAR: Your resource experts, could you give us an overview of the technical disciplines that are represented by your resource experts? In other words, I don't mean names, obviously, but from what disciplines have you drawn that set of people?

MS. XING: Okay. The majority of the research experts, their discipline, they need to have a number of years of experience as an operator and an operator trainer.

(Audio interference.)

MR. PARRY: I am here on behalf of Mary Presley, who is the EPRI Project Manager on this. She couldn't be here today. So, I am just going to briefly go through EPRI's role in this development.

Next slide, please.

Just to give you a bit of background, EPRI has what is called the HRA Users' Group. Its mission statement is stated here. It is basically to develop and maintain tools that allow different analysts to come up with, at least to the extent they can, reproducible and comparable results for similar types of actions at different plants, again, recognizing that there are always going to be some subjective elements, but at least to try to reduce that subjectivity.

Also part of its remit is to develop guidelines for the application of HRA methods.

Remember, looking at the bigger picture, EPRI is probably more concerned with the use of PRA as a tool rather than HRA as a specific discipline. So, it is within the context of what is needed to good risk-informed decisions, for example.

One of their goals is to enable industry to converge from the common methods and, also, toward the same part of the goal is to make sure that those methods, while they are consistent among industry, also have some acceptability with other stakeholders like the NRC and industry, other industry groups.

Currently, the membership of this HRA Users' Group, which met last week in Florida, by the way -- it has its annual meeting every year in Florida in January, which is not a bad place to be.

CHAIRMAN STETKAR: I was going to say I have got to get on that group, yes.

(Laughter.)

MR. PARRY: Yes, it was nice last week. Membership currently is all U.S. utilities and some international and corporate members as well.

Next one, please.

So, just to explain EPRI's involvement in this project, the work that EPRI has been involved in -- and it has been involved since the very beginning.

As you know, the MOU to you asked you to work with the staff and other stakeholders. So, EPRI is one of the other stakeholders. So, it is being done under a Memorandum of Understanding.

And the work that EPRI has been involved in has primarily been on the internal events HRA document. The motivation I think that they had was, as you know, they have developed a number of tools over the years. Really, they haven't been, none of them has really been updated for more than 20 years.

MEMBER BLEY: I hadn't thought about that. Is it really? I guess that is true, isn't it?

MR. PARRY: Oh, yes.

CHAIRMAN STETKAR: It is more than 20.

MR. PARRY: 1980- --

CHAIRMAN STETKAR: '7 or something.

MR. PARRY: '6 or something like that.

CHAIRMAN STETKAR: '6 or '7, yes. Yes.

MR. PARRY: It has been a long --

CHAIRMAN STETKAR: Twenty-five years.

MR. PARRY: Yes. You know, EPRI was also involved in the empirical studies and learned quite a lot from those. But, basically, from other peer reviews and things like that, it is that really the guidance on the qualitative analysis part of HRA is not as good as it could be. I mean, EPRI did develop some way back when with SHARP1 and things like that. But, again, that is even older than CBDT.

So, the motivation, again, is to develop better guidance for qualitative analysis and reproducibility of results, and, hopefully, to come up with one method that is common to both NRC and industry.

So, the path forward that EPRI sees, at least for the next year, is we are going to participate in the expert elicitation in the next couple of months. Also, we are going to do some separate testing, separate from what NRC is going to do.

Some of the things we are going to look at is to confirm that the method -- and this is going to be a testing of the internal-events-at-power method. So, we are going to see whether it meets the requirements of the ASME standard, which I believe it does.

Part of that testing is to demonstrate that the guidance that we have created is clear enough and can be used, and that the time that it takes to actually perform one of these analyses is appropriate, because that could be, I mean, when you look through, when you see, when we describe the method, you see it could be fairly time-intensive. And so, that has to be weighed against the requirements that you need to perform specific PRA applications, for example. So, that is going to be an element of the testing to see how long it really takes.

There is a question of whether the HEPs it produces are reasonable. I think John is probably going to ask me what are reasonable HEPs. Thank you for that.

(Laughter.)

One of the goals is to make sure that there is reproducibility in the results because that has been one of the problems of HRA, is that it does have a lot of subjectivity. Especially when you have rather ill-defined methods, like SPAR-H, for example, you can get answers all over the map.

So, the hope is that this method will give you an increased chance of being reproducible, and reproducible from one set of analysts to another, given the same set of conditions.

CHAIRMAN STETKAR: So, Gareth --

MR. PARRY: Yes?

CHAIRMAN STETKAR: -- is EPRI going to do that element? You are not EPRI, so I have to be careful.

MR. PARRY: Right.

CHAIRMAN STETKAR: Are they going to in their part of this process use different sets of analysts from, let's say, HRA PRA teams at nuclear power plants, so that you actually have practitioners out there and test that reproducibility that way?

MR. PARRY: How many analysts I don't know, but certainly we are going to use different analysts. I know we have at least two.

CHAIRMAN STETKAR: From different utilities?

MR. PARRY: Well, different organizations.

CHAIRMAN STETKAR: Different groups? Okay.

MR. PARRY: Yes, yes.

CHAIRMAN STETKAR: Okay.

MR. PARRY: And we are going to look at a variety of -- I don't think we are going to do, I don't think the intent is to do a complete PRA at this point, but just to choose HFEs from different PRAs that are well-defined and have different characteristics, like things that are time-critical versus non-time-critical, things where we have seen a lot of variability versus not a lot of variability.

CHAIRMAN STETKAR: Okay.

MR. PARRY: And then, based on that testing, I think EPRI intends to go with a recommendation to its constituency as to how to use this new method. If the decision is to go forward, then can we incorporate it quite easily into the HRA calculator? I mean, it is set up almost to be that way.

And then, possibly to adopt the methods beyond the internal-events-at-power HRA. I mean, that is clearly one of the big issues these days, particularly with beyond-design-basis accidents. So, that is the plan going forward, as I understand it.

CHAIRMAN STETKAR: And fires.

MR. PARRY: And fires, yes. Well, there is already work done on fires to some extent.

CHAIRMAN STETKAR: To some --

MR. PARRY: To some extent.

CHAIRMAN STETKAR: Not formalized in the quantification part of the --

MR. PARRY: Not quite in the same way, right.

CHAIRMAN STETKAR: -- process, though.

MR. PARRY: Right. Yes. Exactly.

And with that, I think I will stop there.

MEMBER SCHULTZ: Yes, just some clarification in terms of numbers that are involved. Testing in these areas, it is going to be an application test with different participants doing the same tasks? I don't know what the design is from what you have described so far.

MR. PARRY: No, no. Well, I have got to say I don't know the complete design yet, either.

MEMBER SCHULTZ: Okay.

MR. PARRY: As I say --

MEMBER SCHULTZ: It is under development?

MR. PARRY: I think so. I am not directly involved in that aspect of it. I am here as an EPRI contractor. I am just presenting this. But I imagine that we will have several HFEs to look at and, as I say, at least two, maybe more, different individuals who are HRA experts who will apply this.

MEMBER SCHULTZ: But it will be application? It won't be evaluation by those experts?

MR. PARRY: No, no, no, it will be application to specific defined human failures.

MEMBER SCHULTZ: And then, to look at the results --

MR. PARRY: And then, to look at the results --

MEMBER SCHULTZ: -- to do the evaluation.

MR. PARRY: -- and compare them and see whether we can make sense of the results.

CHAIRMAN STETKAR: One would think that a larger sample of practitioners than two would be useful. Three, for example, might be good.

MR. PARRY: Sounds better.

(Laughter.)

CHAIRMAN STETKAR: It sounds better to see whether you get two out of three rather than two out of two or one in one.

MR. PARRY: Yes. Well, you know, I think that remains to be seen, of how --

CHAIRMAN STETKAR: Understand. You are an EPRI contractor; you can't speak for EPRI.

MR. PARRY: Right. I can't speak for EPRI.

MEMBER SCHULTZ: Well, then, my message back to EPRI would be we can't underestimate the value of this part of the project.

MR. PARRY: No. Well --

MEMBER SCHULTZ: It is very important --

MR. PARRY: Yes, and, clearly --

MEMBER SCHULTZ: -- to the worth in the going-forward application because this is where the questions will be directed associated with all of this technology, is: how did it work and what evidence can you present that demonstrates that it works, and that each of these elements has been evaluated and, therefore, I feel confident that I can both perform the evaluation and accept the results?

CHAIRMAN STETKAR: Well, and in the sense that the assembled technical community being very critical of the current status of methods producing results that are different method-to-method or different analyst-to-analyst within the same method, one would think that the industry would be interested in at least resolving that issue. I mean, I think the last thing that in the interest of both the staff and EPRI would be to publish a methodology that later is subject to the same criticisms that any one of the current methodologies is subject.

So, having that active participation among a broader subset than two, for example, sets of analysts would seem in everyone's best interest, recognizing the constraints that EPRI has --

MR. PARRY: Right.

CHAIRMAN STETKAR: -- you know, to get participation.

MR. PARRY: Well, and I suspect that, as

with anything new, particularly one that takes a lot more effort than current methods, it is going to be a hard sell. So, you have to convince people that it is worth the effort to do it. I think part of the testing is designed to do that.

MEMBER SCHULTZ: I agree. That is the point of communication.

Thank you.

CHAIRMAN STETKAR: What we are going to do, by the way, is I don't know whether Mary Presley is on the bridge line or not.

MR. PARRY: She might well be on the bridge line. She was in --

CHAIRMAN STETKAR: We are going to see if we can open that up, open it up and see if she is out there. And perhaps she might want to give us some feedback.

MR. PARRY: Yes, she was in --

CHAIRMAN STETKAR: She is probably screaming at her phone right now.

MR. PARRY: She is on the West Coast, so her meeting won't be started yet.

CHAIRMAN STETKAR: Yes, it is 6:30 out there. It is the shank of the day.

(Laughter.)

Mary, are you out there?

(No response.)

Well, that answers the --

MR. JULIUS: Hi. This is Jeff Julius. Mary was on earlier and Katie and --

MS. PRESLEY: Yes, I'm on as well. Sorry.

CHAIRMAN STETKAR: Okay. Mary, I don't know if you have been listening to the exchange over the last 10 or 15 minutes regarding EPRI's plans for testing this method and the scope of participation in terms of different utility groups or different, let me just say different sets of practitioners. Do you have any comments on that or feedback? Or is it too premature to ask about that?

MS. PRESLEY: I think it is a big premature to ask about that. Right now, our plan is to get the internal testing complete on kind of a select set of HFEs that span the different sets of human action types.

CHAIRMAN STETKAR: Yes.

MS. PRESLEY: And then, we are going to proceed forward and see how we want to test more than that.

CHAIRMAN STETKAR: I think from my personal perspective -- and I don't know whether

Subcommittee members have different ideas -- I think it would be more useful, rather than having two sets of analysts evaluate 25 different HFEs, to have 25 different analysts evaluate six sets of HFEs. So, in terms of organizing the resources, what I heard you say is that you are initially going to control the scope by the number of different types of HFEs or scenario contexts, and in terms of resource management, to test the reproducibility and analyst-to-analyst variability in terms of applying the method. In terms of resource control, if you can get participation from a broader set of analysts than perhaps two, it might make sense. Again, that is just some feedback.

I don't know if any of the other members have --

MEMBER SCHULTZ: I would agree with that, John, and there are two reasons, the one that you mentioned. The other is that we believe, as Gareth indicated, it seems like yesterday, 20 years, but when this went through its first development, if you will. And now, we are coming back.

We have to believe and we have seen from the products that we have so far that there have been, we believe, good advances to the methodology. And this testing is going to be very important to move it forward and to have more than a few. It doesn't have to be many. But everyone that participates in this testing is going to either buy-in or not.

CHAIRMAN STETKAR: Yes.

MEMBER SCHULTZ: And I believe, based on the work that has been done over the last two years, which has been fairly intensive in the development, and I think the project is at the point where it should prove itself. But if it can be proven to five groups, let's say, rather than two, it would be very valuable.

CHAIRMAN STETKAR: Thanks.

Anybody else?

(No response.)

The only reason I wanted to make sure that we got feedback to Mary, and perhaps feedback from her, is we are going to reclose the bridge line because, for those of you out there, you don't sit here; we get all kinds of pops and crackles and things like that that are really disturbing to us and even worse for the recorder.

So, unless, Mary, you have any other comments you would like to make in that regard, we will reclose the bridge line. We will reopen it periodically during the day to see if people have comments.

Anything else?

MS. PRESLEY: No, but thank you very much for the feedback. I will take that into --

CHAIRMAN STETKAR: Thanks, Mary.

MS. PRESLEY: Yes.

MR. ZOULIS: This is Antonio Zoulis with the NRDRA.

One of the things that I would like to see out of this is how do we address dependency, which I believe is what causes a lot of the variability between analyst-to-analyst. And also, whether or not we believe a sequence that has two or three HEPs in it could be the below 1 to the minus 8, 1 to the minus 9. Again, those are issues that we struggle with all the time.

And also, I think recently NEI has also made some comments to the STP process about conservatisms in the way we handle HFEs. So, I know the guy who came out with an EPRI document recently that addressed some of that, but having more substantial basis for how to address those issues I think would be very beneficial.

Thank you.

CHAIRMAN STETKAR: Yes, thanks.

MR. PARRY: Mary may want to make a comment on that, too, if the bridge line is still open. Is it?

CHAIRMAN STETKAR: I think it is closed again.

MR. PARRY: Okay.

CHAIRMAN STETKAR: We will come back to it. Because I am assuming in your presentation of the methodology you will get to the section that talks about dependence?

MR. PARRY: Yes. Yes, I mean, obviously, it is not something we have fully developed yet.

CHAIRMAN STETKAR: Right.

MR. PARRY: But the ideas are there and we believe that at least it provides a more rational basis for addressing the panel.

CHAIRMAN STETKAR: Yes. My only message is that we will go and revisit the dependency as part of that discussion, anyway.

MR. PARRY: Yes. Okay.

MS. XING: And just from project management's perspective, at present for the method, we adapt the NUREG-1921, dependency method So, that is it.

And we also proposed the different potential ways we can make improvement in the report. And in the project plan -- it is in FY13 -- we are going to further boil down the approach we propose, if we can come up with better treatment.

CHAIRMAN STETKAR: Okay.

MS. XING: Okay. So, I have a question for the Committee here. Since we are behind the schedule -- so, our original plan was that next we would spend one hour to talk about the generic methodology and the later hour for the more mature method for internal event. Now, since already we don't have that one hour, one proposal is I can give you like a 10-minute overview of what the generic methodology looks like.

And since we are interested in another meeting for the technical basis, we can put that technical basis and the generic methodology together. Alternatively, we can forget about generic methodology or just to go to the --

CHAIRMAN STETKAR: I personally think it is a good idea; the 10-minute, quick run-through sounds like a good idea, just to get all of the members oriented toward that framework.

MS. XING: Okay.

CHAIRMAN STETKAR: And I agree that we can delve into the generic methodology document in a subsequent Subcommittee meeting with relationship to the cognizant psychology.

And having read through both of the reports, there is an awful lot of overlap between what is in the generic methodology report and what is, I am assuming, going to be covered under the Level-1 internal-events procedure-driven part of the methodology. So, to avoid duplication, I think we can do that.

MS. XING: Yes.

CHAIRMAN STETKAR: That will get us back on time a little bit.

MS. XING: Okay. Yes.

CHAIRMAN STETKAR: Because I think it is important for us to understand, since the Level-1 internal-events at-power methodology is the furthest along and will start being piloted, that you folks have the benefit of our feedback on perhaps some details on that method. So, we want to make sure we have enough time for that.

So, let's try to do 10 or 15, 20 minutes.

MS. XING: Okay.

CHAIRMAN STETKAR: No more than 20, so we can get a little bit back to --

MS. XING: Okay. I will try to see if I can do it in 10 minutes.

CHAIRMAN STETKAR: And I will try to be quiet.

MS. XING: The staff cover intend to use the other HRA domains

Okay. So, just before we start, the concept is -- why this picture jumped up by itself?

CHAIRMAN STETKAR: And I was going to say I take offense to that because it looks an awful lot like me.

(Laughter.)

MS. XING: So, the generic method can be viewed as an extension of the internal-event method, or vice versa; the internal-event method can be viewed an example of implementation of the general methodology.

Okay. This picture just like jumped up by itself. I don't know why. So, just want to show that humans, that they make an error, even a paid experienced operator.

Okay. So, the generic methodology will focus, we hope the generic methodology is for HRA cases. So, it heavily relies on the combinative basis.

And we still want an independent method. It will conform to the generic, the general HRA process defined in the PRA standard.

You have a qualitative analysis starting from and standing -- well, you all know more than I know about this process.

So, this is what we have for the generic methodology. It has five elements.

The first part is guidance for HFE analysis. Since we have barely done any development, it is mainly stands from existing documents. In particular, it is from the most recent fire HRA document.

And the second element is the cognitive error-causal tree. As we talked earlier, this we have more development, extending the technical basis with data earlier and put it organized in the format that can readily support qualitative analysis and the quantification.

And the last element that we have put in the analysis and the quantification process development.

Then, as an integrating analysis, one target work on there is the dependency.

So, I will skip the generic guidance for HFE analysis because it is mainly back from existing knowledge.

The more extension we did in the cognitive basis -- so, early we gave you those basic cognitive functions. Well, an important extension here is, once with the cognitive literature and the NPP task analysis document, and to try to identify the objectives for the cognitive function by -- not where you say objective; it is a sub-task. So, you do detection. What kind of task do you in the capsule? And so, those objectives, it is the foundation of the basis where we identify failure modes.

And then, the cognitive mechanisms to achieve these functions, that is what we had in the early report, and we added some new there.

And another important development here is this last element of complex characteristics that challenge cognitive mechanisms. This directly goes into the link between here and the cognitive mechanisms. And these characteristics will be the basis for quantification.

So, I will show you how. So, for this overall model, this cognitive basis would benefit, I put as teamwork and the supervision, which is not explicitly included in the previous effort.

Here are some examples of what we mean by the objectives for understanding. So, the major type of task they do and the function. You need to assess and verify the information and maintain situational awareness, prediction for the future, and diagnose problems.

So, for each mechanism, here are some common mechanisms as before.

So, this is how we modeled the PIFs. So, instead of just saying a PIF is a test, we identify a set of the characteristics under this PIF. And this characteristic challenges combination mechanism for cause and error. So, the PIFs we currently model are workload and task demand.

And then, test environment procedures, and training work process and the organizational factors. So, each of these major factors we identify like, I would say, five to ten major characteristics that lead to error.

So, here just gives you a couple of examples. For example, the proper workload, it will not just say workload. The particular workload character will be multitasking, interruption, and these will affect the integration mechanism, the mechanism of integrating information and understanding function.

So, therefore, when you look at this list, you can look at in two ways. By the way, if you look at each function -- so, we grouped this as characteristic, and each PIF is complex. If I say I just want to look at workload, how workload would affect performance, you can group this and the different functions of the workload with the effect of detection. Because of this characteristic, workload would affect understanding because of this in place.

So, this list is the foundation of the quantification model. So, therefore, this technical basis has these four layers of structure, starting from the basic function objective to combinative mechanism, to error causes, and to complex characters.

So, the qualitative analysis is to represent a human event in terms of this human task, a function, and the objectives. And the quantification is to assess these characteristics, and this top and the bottom are linked by underlying cognitive mechanisms. So, this slide is essentially how we developed the generic methodology.

How many? Okay. I have five more minutes.

So, I will skip the qualitative analysis part, which is basically you carry out those analyses, qualitative analyses; identifying the tasks, identifying critical tasks; identify what function to achieve those tasks; identify the objectives you need, and to perform workload analysis. Finally, you can integrate these all together and come up with an operational story before you go to the detail of qualitative analysis.

MEMBER BLEY: I would just like to add a comment to what you have presented here for the members of the Subcommittee. And a lot of the work that preceded where these folks are now, this idea of doing a good qualitative analysis, while it has always been around, isn't heavily supported in most other methods, and it is probably the place most analyses have fallen down in the past. So, there is a real emphasis on that in what they are doing.

MEMBER RAY: But, Dennis, one of the things I have been trying to discern here is -- it came up in something in Gareth said a long time ago -- by narrowing it down to a group, if you are using internal events, of people who are licensed and all pretty much meet a standard because of that, it seems like the task is more tractable than if it is expanded to a group -- God knows who and what qualifications that group has -- that are involved in beyond-design-basis events, for example. And I am trying to figure out, does this methodology recognize the difference between those two situations?

MEMBER BLEY: I will let them talk more about it, but I would say, yes, it does. But, on the other hand, I have got a really nice paper from a few years ago that Jim Reason put together, a former professor of psychology at the University of Manchester, to take to the medical community, who are very highly-trained, very organized, to tell them why they are not immune to human error, because of their brilliance in background, which applies here as well.

MEMBER RAY: No, I am just talking about we apply it to a group that, you know, by comparison with the second group I mentioned, they are all licensed operators or --

MEMBER BLEY: They have similar backgrounds, similar training.

MEMBER RAY: Yes. And then, when you get into management, or God knows what other elements of decisionmaking take place beyond internal events, you get a lot of people engaged that it just seems like a different world to me.

MEMBER BLEY: It is, but they are all people, as John was saying --

CHAIRMAN STETKAR: But that is actually one of the reasons why I think it is a good idea that our Subcommittee have a more formal discussion about the NUREG-2114 --

MEMBER RAY: That is why I agree with you.

CHAIRMAN STETKAR: -- and its relationship to this generic methodology. Because that type of perspective, will the information in the NUREG and its mapping into this methodology handle those situations? That is an important question.

You know, I have my own opinion that I am not going to offer right at the moment, but that is certainly a very valid question and will become much, much more important, not only for severe accidents, but integration for Level 3 under seismic events or flooding or whatever. So, it is a good point.

MS. XING: Yes, at the next meeting, if you would like, for the combinative basis, I will like go over the detail of this diagram, then show how we use this diagram to develop the generic methodology.

Okay, the qualitative analysis. Now I think will be our quantification, is what we have here.

So, in the quantification process, we propose a two-level analysis, inspired by fire HRA. First, a scoping analysis, then a detailed failure-mode analysis.

So, for the scoping analysis, it is to determine the HEP range at the right level. So, it is a process where, say in this diagram we break down, we identify the critical tasks in an event. For each critical task, we identify the cognitive functions that are required to obtain the task goal.

Then, we quantify the estimates, the HEP on the HEP range from fail-based cognitive function. So, the HEP is a function of all the potential cognitive characteristics that can contribute to the failure of this function.

So, this is a graphic showing. So, you have a cognitive function there, and you have all these PIF factors affecting this function. And each factor has a bunch of complex characteristics. So, you assess all these characteristics, see if they are present or not present, or maybe we gave a weight or scale.

And finally, we wish to use the expert elicitation to come up in a function like this.

CHAIRMAN STETKAR: We don't have time to

discuss this. Let me just give you a quick feedback.

MS. XING: Yes.

CHAIRMAN STETKAR: This is silly. That is a statement on the record. We can discuss that more in -- this is just doing a body count of the different performance-influencing factors. And if you have got like seven, you are up at some point, and if you have three, you are down at some other point, which to me doesn't make any sense at all. But perhaps there is some notion behind it.

MS. XING: Yes. That is why I used the word "indication" instead of saying the number. So, it is the most simple way you can think; of course, a more normal factor there, the more likely, actually, people end high.

And in addition to that, most basically, at this moment we have three other proposed org charts, like table 18, table 10--

CHAIRMAN STETKAR: Those other ones make a lot more sense, by the way, to me, anyway.

MS. XING: But just the most to move forward is just count your fingers.

Okay. I am just now going to detail the failure-mode analysis, which the condition for using this is if the task is well-defined. You have more information on the task and the related context factors. Therefore, you can really go to a lower level. So, look at what are the objectives of the cognitive function. And for each objective, there could be one or more ways you would fail this objective, which will be a failure mode.

So, then, you look at, select those characteristics that are most significant incidents in this particular failure mode and build a decision tree. That, you will see the detail in the next presentation.

So, we have provided a set of the generic test failure-modes. And that is our hope, that these generic test failure-modes would cover the general, broad HRA domains.

So, this is just to show how we put everything together, this whole process, which is very much like the diagram in the beginning, the generic HRA process.

Okay. So, 15 minutes.

CHAIRMAN STETKAR: Very good. And I have mostly been quiet.

My only comment, Jing, is I read the reports kind of in my mental notion of the way I think they are organized. So, I read the Cognitive Basis

Report, 2114, and then I read the Generic Methodology Report, and then, finally, I read the Internal-Events At-Power Procedure, whatever, Report, which is kind of the hierarchy that I understand the process should flow.

So, in that sense, the generic methodology is not derived from the internal events stuff. The internal events stuff is a subset of the generic methodology.

MEMBER BLEY: A specialization.

CHAIRMAN STETKAR: Thank you. It is a specialization, a better term. That is why you are here.

(Laughter.)

Something for you to just consider, especially if we are going to have another briefing, both on the NUREG and a little more information on the generic methodology, kind of an overarching concern that I had is, as I read through the Generic Methodology Report, it was really difficult for me to understand the linkages and the decisions that were applied with regard to the Cognitive Research Report.

In other words, you have grouped, simplified things, in some cases expanded things out in the teamwork and communications, and it was really,

really difficult for me to have the NUREG, on the one hand, with its framework and its information, and understand how that was mapped into the generic methodology.

And I think that mapping or the documentation of that mapping, the thought process, is essential. Because without that understanding, then it is really, really difficult to understand how a particular subset of that methodology or a specialization of the methodology to a particular type of set of tasks is traceable back to the fundamental reference.

So, you may want to keep that in mind because, honestly, I found that really, really difficult. I wrote up a lot of notes on it. But, if I look back, that is sort of the overarching concern that I have.

MS. XING: I fully understand that. Two points here. One, in this generic methodology, we did a lot more literature which were not in that original --

CHAIRMAN STETKAR: And it references that, and that was another one of my concerns. It said, well, we did some literature work. Well, how does this relate to the literature work that was done in the overarching document and why is it different?

MS. XING: Yes. As I say, like --

CHAIRMAN STETKAR: Or is it different?

MS. XING: A lot of additional work in addition they did on that wasn't in that Cognitive Basis Report, and I didn't write them out in this report.

CHAIRMAN STETKAR: Yes.

MS. XING: So, it is nowhere shown in the report. Eventually, I would like to think of a strategic way to either have a Volume 2 of a Cognitive Basis Report or update that Cognitive Basis Report before we --

CHAIRMAN STETKAR: I think, to kind of get us back on schedule, this is something certainly we should examine in that next Subcommittee meeting. Because from what I am hearing you say, it is that I am getting the impression that the Cognitive Basis Report is in some sense incomplete. And I didn't appreciate that, having read the Cognitive Basis Report. It seemed to me to be fairly comprehensive and complete, and perhaps that is why I was struggling understanding how its degree of detail and my perception of its completeness related to some of the decisions and how they were implemented in the

Methodology Report. So, if we are looking at something that is different or not fully integrated, that is, I guess, a bit of a concern.

Anyway, that is for the next Subcommittee meeting. And I think because of that, it is probably important for us to try to work on scheduling one. Fortunately, our --

MS. XING: Yes, at the next Subcommittee meeting, it would be cognitive basis and methodology together, or say it better.

CHAIRMAN STETKAR: The good news is our schedule is fairly light, fortunately, over the next two or three months. So, we may want to think about doing that sooner than later.

MS. XING: Okay. As long as it does not conflict with our expert elicitation workshop.

CHAIRMAN STETKAR: Yes. No, obviously, you need to get work done. But the only message is that the last time we got together on this was, I think, several months ago, if not a year, I think. I think it was a year ago. So, I certainly don't want to postpone the discussion that long --

MS. XING: Okay.

CHAIRMAN STETKAR: -- or even six months.

Any of the members have any more comments or questions about what we have discussed so far, the generic methodology? We will get into more details when we hear from Gareth and John.

(No response.)

If not, let's take a break, and we will reconvene at 10:30.

(Whereupon, the foregoing matter went off the record at 10:13 a.m. and went back on the record at 10:32 a.m.)

CHAIRMAN STETKAR: We are back in session.

I don't know who is up. Gareth, I guess?

MR. PARRY: Yes. This is a presentation that John and I put together. I am probably going to do most of the talking, and I will ask John to jump in and correct me when I say something incorrect or help me clarify things.

MEMBER BLEY: We will watch his eyebrows while you talk.

MR. PARRY: Okay.

MR. FORESTER: I did it for the training. It is his honor. He is the lead on it.

(Laughter.)

MR. PARRY: Okay. So, what we are going to talk about is the IDHEAS HRA method for the internal at-power events.

So, in this general slide that Jing showed before, we are at the bottom level there, which is the specific application as opposed to the generic one. And, you know, chronologically, we actually did the specific before we did the generic.

CHAIRMAN STETKAR: By the way, just to interrupt --

MR. PARRY: Really? Sorry.

CHAIRMAN STETKAR: Don't get me started.

(Laughter.)

We do need to end at 12:00.

MR. PARRY: I know.

CHAIRMAN STETKAR: We have another noontime meeting. Well, it is just one of these extenuating circumstances. So, we can't really drag on. So, we just collectively need to make sure that we organize our time that way.

MR. PARRY: Okay. I think we can get through it as long as we don't have prolonged interruptions.

(Laughter.)

MEMBER BLEY: You will find a way.

(Laughter.)

MR. PARRY: Okay. Just in terms of the outline of the presentation, what the purpose is is just to talk you through the draft report, which obviously John has read; I don't know if everybody else has read it.

The focus of that report was on the method itself, rather than the technical basis, as we understand. And the elements of the IDHEAS method basically are the identification of the definition of the human failure events, which are the PRA elements; the feasibility assessment, which is really done in parallel and throughout the whole process. It is not a separate element in itself. It is a continuous activity, if you like, and it is necessary for you to include an HFE in the model. The action has to be feasible for it to be taken credit.

The key technical things about the method are beyond the definition. And now, this is to help with the development of the qualitative analysis. It is the performance of a task analysis and the development of what we call the Crew Response Tree, the CRT.

The idea behind this is to identify the critical tasks that are needed to perform the response that is being modeled and by identifying those critical tasks, the opportunities for failure. Then, once you have done that, we have created a model based on what we call Crew Failure Modes.

So, given that you have identified the opportunities for failure, we need to identify the Crew Failure Modes that are applicable to the activities that might lead, if they are not done correctly, might lead to failure.

We have created a set of decision trees that assess the contextual impact on the human error probabilities. And then, we use those in a quantification formula, which I will describe to you.

We have also included in this package an example decision tree. John asked me during the break whether we were going to discuss each decision tree, because apparently he has questions on them. We really didn't have the time to look at them in detail. So, we have taken one example, just to give you an idea of what they look like.

And we also want to make a comment on dependency. I notice Antonio has left. But it is one of the things that I think that we can come out of this project with, is a more rational approach to dealing with dependency, which, as you know, the traditional methods are somewhat simplistic.

So, this figure, which I won't dwell on, is just essentially the flow path, if you like, going from the top to the bottom of the construction of a PRA model and the incorporation of human failure events in that logic model.

And then, we go through the qualitative analysis, the development of the Crew Response Trees, and then, we will talk about how we use those Crew Response Trees to analyze the human failure event to come up with a human error probability.

So, the first part of this I will talk about is the identification and definition of human failure events. And just to remind you, the extended definition of a human failure event is that it is a basic event that represents a failure of an availability of a component, system, or function that is caused by human inaction or an inappropriate action.

The document that we have described addresses specifically HFEs that represent failures to respond to either failure or non-availability of a system or a function or failure to manually initiate a required function following an initiating event.

So, in terms of the PRA characterization of human failure events, we are dealing specifically with what we called post-initiating event HFEs. We are not talking about the classical pre-initiator event HFEs, which are, you know, leave the valve in the wrong position and that type of thing.

But also included in this -- it is not just the pre-initiator -- sorry -- not just the post-initiator HFEs, but also in some cases for the modeling of things like support-system-initiated, you might construct a fault tree. And that might include a response to a failure which the failure in itself does not lead to the initiating event, but the failure to respond to it could. So, we are dealing with response actions, not routine actions that go wrong, if you like.

CHAIRMAN STETKAR: Something that is triggered by a cue.

MR. PARRY: Something that is triggered by a cue, yes.

CHAIRMAN STETKAR: For a failure.

MR. PARRY: Right. Yes.

So, the guidance for the identification of the HFEs, we haven't done a lot of work on this area specifically in this project. That has not been our focus, although the document does contain some discussion, but it is really based on the work that is in the HRA Good Practices document, the ATHEANA document, SHARP1 -- as we said, it is at least 30 years old -- and the Fire Human Reliability Analysis Guidelines. So, there is nothing particularly new in the document, a little bit of generalization, but nothing new.

So, typically, we are going to include HFEs in the model, this type of HFEs that we are dealing with and ideas, when an operation actually is required. So, either restore, maintain, or initiate a function. This message where you survive core damage.

And the identification is done based on the specific procedural guidance that specifies the operator actions, and the typical procedures were are going to look at are EOPs, AOPs, and initiated response procedures, and possibly some system operating procedures as well.

The identification of these really depends on -- in fact, you just said this, John -- the availability of cues that alert the operator to the need to do something. So, a typical HFE definition that you would start with, given that you have developed a PRA model and you have gone through the process of looking through the procedures to identify what actions to take place, you would also include, look for the PRA-specific timing of cues and other information that you need to assess the plant status.

So, that is something you need to look at.

You also need to look at the plant state, a physical condition by which the operator action must be completed, which defines effectively the corresponding time window. So, it is part of the success criteria.

So, if you want to do like in the case of steam generator tube rupture, if you want to define an event that is isolating the generator before it overfills, it is the overfilling that would define the endpoint of the time window.

You also want to look at any other information or cues that they might need to enable them to perform the response correctly. This is all part of assessing the feasibility of the event, if you like.

And you need to look at the equipment system or systems that the crew uses to achieve the functional goal. And effectively, at least to the high level at this point, the way that equipment is used. For example, initiate an injection using the high-pressure injection system or perform depressurization, not necessarily looking at the details at this point, but we are looking at a general characterization of the HFE.

And just out of the comment here on the assessment of feasibility, because I think for most of the things that you would see in the EOPs, you would expect that they are feasible actions; they have been defined that way. There could be certain circumstances under which they become non-feasible, and those are the things that you need to look for.

So, you need to look at things like timing that is available to complete the response; the availabilities of procedures and/or training, availability and applicability I would say; availability of cues because that is the key element because there may be certain scenarios in which the cues are actually not relevant -- sorry -- are not available, even if the procedures are available.

In certain circumstances, depending on the type of responses needed, you need to worry about the accessibility of the area, where the response has to be performed, and you also have to look at things like the availability of resources, personnel, and equipment. As I say, for most of the EOP-driven actions, this shouldn't be a major concern except for some specific unusual circumstances.

The comment we made earlier was that the feasibility is really an ongoing activity. So, you can make the call fairly early on in some cases, and at other times you have to wait until you have done the more detailed timeline of when the cues occur, when they get to the right point in the procedures, and the time that is available, before you can determine feasibility.

CHAIRMAN STETKAR: Gareth, I am woefully bad about looking ahead in the presentation. Are you going to talk about recovery as an issue as part of it?

MR. PARRY: Yes.

CHAIRMAN STETKAR: Okay.

MR. PARRY: Yes. Yes, that is actually quite an important part of the way we have done it.

CHAIRMAN STETKAR: Okay. Because I wanted to read this, this notion of feasibility in the sense of that issue.

MR. PARRY: Okay.

CHAIRMAN STETKAR: We will discuss it when you get there.

MR. PARRY: Okay. The next section of this I am going to talk about is task analysis and the development of the Crew Response Tree. And I think this is, I would probably say that this is an area, at least the task analysis part is something that HR analysts ought to be doing without even addressing ideas.

But it is clear that it is not done as well as can be expected. So, I think one of the things that we have incorporated to ideas is a more formal approach to performing and documenting the task analysis to try to improve this aspect of HRA.

So, coming into this, probably what we have done is we have defined the human failure at a fairly-high functional level. Okay. So, we have a statement in terms of the success criterion for the response, which would be the operators have to perform X by time P or before the pressure exceeds Y. Okay? So, there is a very general description.

So, what we do in the task analysis is to understand in detail what the activities are that are required to perform the "do X" and, also, at the same time to construct a timeline of the events that are significant to the performance of those activities, such as the occurrence of cues, the expected time to get to a certain point in the procedure, at which they are required to get a piece of information. And also, we probably already defined the time one in this case for this idealized situation.

So, the task analysis that we do is not

just looking at actions, but it is looking at cognitive activity, which is I think perhaps differentiates this from some of the older HRA methods perhaps.

So, we are sort of assuming that we are following a procedure of some type. So that what we do here is a procedural task analysis effectively to identify the essential activities that the operators have to perform and the nature of those activities. So, the types of activities are things like collect data. The nature would be how they do that.

Another activity might be interpretation of that data in the light of some criterion that is specified in the proceeding. There, then, is typically a decision to initiate some execution, and the execution is done in different ways. We have to figure out the ways that that is performed.

So, the first thing is to look for the things that are essential to success; they have to do this. There are other things in the proceedings that are not essential to success necessarily, and these are things like verification of your status. Okay. Then, if you missed that out, as long as the equipment is in the right status, you are okay; you haven't failed anything. But that might take time.

So, that is an important part of assessing the time it takes to work through the procedure, even though, if they didn't do it, it would shorten the time, and it wouldn't fail it. So, what we are looking for is we are looking for the essential tasks.

In the document, we have put this task analysis in three different stages. First, there is the current transition of the expected success path. So, we are looking at the path through the procedures, which may involve transferring to a subsidiary procedure, or whatever, to get to the right place.

So, in that path, we are looking for the significant tasks, which are things like entry into the procedure, the transfer within or to another procedure, the point at which the decision is made, and then the execution.

We have taken those sort of significant breakpoints in the path, if you like, as the nodes that we are going to put on our Crew Response Tree. It is not essential that we do that. We could do it at a lower level.

MEMBER BLEY: So, if there are alternative ways to move into that second-order procedure you want to get to, do they get separate places on the decision tree or do they get somehow thought about as a lump?

MR. PARRY: Good question.

MEMBER BLEY: Thank you.

MR. PARRY: We will come to that in a minute.

MEMBER BLEY: Okay.

MR. PARRY: I will give you an example of where there are alternate ways.

MEMBER BLEY: Okay.

MR. PARRY: The next step, then, is to look at what does it take to effect that transfer correctly. And that is where we start looking at the subtasks, where we start looking at the individual steps in the procedure to see what it is they have to do. Do they have to get this piece of data? Do they have to use it in this criterion? And then, do they have to interpret that correctly to get you to the right place. Okay. So, these are the critical subtasks, if you like, of the proceeding.

And then, this is sort of related to what you were asking, Dennis, I think. And that is we also look for the opportunities to recover. Because I think if there are alternate ways of getting to the place, then, at least the way we have treated it for this -- and I think to some extent this is true -- one of them can be regarded as a recovery of another one.

MEMBER BLEY: Okay. Even if in strict terms it might not be a recovery, but the time sequencing might have happened such that that is the way you would get there to --

MR. PARRY: Yes, yes. If you missed the first opportunity, you might get a second opportunity.

MEMBER BLEY: But you might have missed it not because you missed it, but because the timing wasn't right for it to occur.

MR. PARRY: Well, okay, that should be reflected in the timing.

MEMBER BLEY: It should? Okay.

MR. PARRY: Yes, yes.

So, this is a representation of a Crew Response Tree, the way we have developed it. It is taken directly from the report. And it is obviously for a total loss of feedwater. So, the action we are looking at is really at the high level the HFE is to implement feed-and-bleed before core melt.

So, let's look at the first node. We are assuming here we have gone into the right proceedings. So, the next step is to enter into -- it is a Westinghouse plant -- so we are entering into FRH-1, which is the response to loss of total feedwater.

In a lot of the Westinghouse proceedings

I am familiar with, there are two ways to get into FRH-1. One is through E-0, a step in E-0, and it could be the foldout page or it could be a specific set. Another one is through the Critical Safety Function Tree. It happens at this plant that there is only one way, and that is the Critical Safety Function Tree, but at another plan, through E-0, which is a little different.

But, nevertheless, what we have done is, if you look at node 7 on this, okay, in the plant that would be where the primary would be through E-0 and the secondary would be through the Critical Safety Function Trees, which are two ways of getting to FRH-1, 7 would be modeled in this as a recovery to the first method. Okay? So, that is the idea behind this. And we have similar things elsewhere. So, I am not going to dwell on the tree, but just to make sure that you get the picture of what this tree represents.

Now, so we have these fairly high-level nodes enter into FRH-1.

CHAIRMAN STETKAR: Gareth --

MR. PARRY: Yes?

CHAIRMAN STETKAR: -- while you have the picture up here, something that I guess I didn't appreciate, or perhaps I don't understand at kind of a high level in the whole methodology, and I didn't quite get it, I think, until I got close to the end of the report, is that, although the Crew Response Tree is very important for the analysts to depict the critical tasks in some sort of logical and/or relationship, series, parallel, you know, this picture here.

MR. PARRY: Right.

CHAIRMAN STETKAR: The methodology really doesn't use the Crew Response Tree to support quantification per se.

MR. PARRY: Yes, it does.

CHAIRMAN STETKAR: It does?

MR. PARRY: Yes.

MR. FORESTER: I don't think -- not in the way he is thinking about it, I don't think. You mean explicitly values put into then --

CHAIRMAN STETKAR: Right.

MR. PARRY: Oh, oh, oh, oh. Okay.

MR. FORESTER: Like a node tree.

MR. PARRY: Okay. It is not an event.

CHAIRMAN STETKAR: It is not an event?

MR. PARRY: No.

CHAIRMAN STETKAR: And I wanted to make sure I understood it in that context, and I think it is important for other members --

MR. PARRY: Yes.

CHAIRMAN STETKAR: -- who may not have appreciated that, that this is not an event tree that has a specific numerical value for each of the four depicted sequences here or even a formal -- it might have this thought process logic structure --

MR. PARRY: Yes.

CHAIRMAN STETKAR: -- but it doesn't have the same type of bimodal actual quantified structure, right?

MR. PARRY: Yes, you are absolutely right. It is used in the quantification, but in a very specific way, and it is not used as it gets into the event tree.

CHAIRMAN STETKAR: Which I was struggling with initially because -- and I think it is okay -- but in the initial discussion of the development there is a lot of latitude, if you will. It said, well, you know, in this context you could put three nodes, 4, 6, 6, at the top, or you could put, you know, 37 nodes, depending on--

MR. PARRY: Right.

CHAIRMAN STETKAR: -- how you particularly want to define the critical tasks and draw the boxes.

In a sense, it really doesn't make any difference whether you have 3 or 36 as long as you have the right crew failure modes and decision trees supporting each one.

MR. PARRY: That's right.

CHAIRMAN STETKAR: You can have 12 into 1 --

MR. PARRY: Yes.

CHAIRMAN STETKAR: -- or you can have 1 for each of 36.

MR. PARRY: Yes. You are right.

CHAIRMAN STETKAR: Okay. Okay.

MR. PARRY: Yes.

CHAIRMAN STETKAR: Okay.

MR. PARRY: Yes, this is really just a --

CHAIRMAN STETKAR: As I said, it is something I didn't really fully appreciate until I finally got to the end and said, "Oh, okay, that's what they are doing."

MR. PARRY: We had issues with this internal to the project, too, because I think some of us were not thinking of this event tree, some were, and that led to a lot of confusion.

CHAIRMAN STETKAR: Yes.

MR. PARRY: That is a good point. I think we perhaps need to clarify that.

CHAIRMAN STETKAR: It is made in the report. But, as I said, I don't want to take the time to look back at my notes, but it sort of kind of came together toward the end, when you were talking about integration of the quantification, or something like that.

MR. PARRY: Yes.

CHAIRMAN STETKAR: And a statement is made, well, we don't really use the Crew Response Trees --

MR. PARRY: As an event --

CHAIRMAN STETKAR: -- in a logical sense of a traditional event tree.

MR. PARRY: Yes, it is not a logic model; it is a representation --

CHAIRMAN STETKAR: Yes.

MR. PARRY: -- well, communication.

CHAIRMAN STETKAR: Yes.

MR. PARRY: Okay. So, as I said, those nodes, that node, it is a relatively high level. So, now what we are looking at is what are the subtests required for that node. For that particular node, 4, what is required for this plant that we developed that CRT for is that, basically, there is an operator who is looking at the critical safety function status trees. His job is to look at the steam generator levels and the auxiliary feedwater flow, compare those to the criterion, and then, when the criterion is satisfied, transfer to FRH-1.

MEMBER BLEY: After E-0 allows them to do that?

MR. PARRY: He is in E-0 I think by the time you have gotten to this point.

MEMBER BLEY: You have got to go through a big hunk of the diagnosis in E-0 before you are allowed to use those status trees to transfer out. That is how a couple of plants have defined it differently.

MR. PARRY: Yes. This particular plant gets them looking at the critical safety function trees as soon as they realize that they don't need SI.

MEMBER BLEY: Wow. Okay.

MR. PARRY: That is the way the procedure is structured. I just want to make the comment that the development of the timeline is concurrent with the development of CRT. It is difficult to show it on the same picture, but that is what we did.

Okay. So, here is the overview of the quantification model. As you are aware, the way we have developed this quantification is that we are going to tr to estimate the HEPs on the basis of explanation why the HEP could occur. And it is things like due to various conditions, the crew is dismissed, relevant information that results in their failure to get the right plant status assessment. And therefore, they don't do the required response. So, that is an example of an explanation of why the HFE can occur.

These explanations, we call them crew failure scenarios. They are based on the understanding of the cognitive mechanisms and the related performance-influencing factors that are discussed in the basis document.

The different crew failure scenarios, we have grouped them into groups by the crew failure mode, which is sort of the observable by which you could see a crew had made an error. And within that group of scenarios that are characterized by this common crew failure mode, if you like, the scenarios are differentiated by the various PIF characteristics that can affect the likelihood of that CFM occurring.

On the expanded PIF -- we used PIF rather than PSF because it is really a bigger set than the traditional PSFs that you get with things like SPAR-H or the other methods, because it involves scenario characteristics and plant conditions.

So, how do we determine the crew failure modes? Well, basically, we looked at three major elements of the way the crew operates, and we characterized them as things like plant status assessment, response planning and execution.

CHAIRMAN STETKAR: Just something that I had to kind of do, because I did, as I went through this, the same as the generic methodology, I kept a set of pictures from the Cognitive Basis Report. In the sense if I do the mapping, what you are calling plant status assessment I believe includes what the NUREG calls detection and noticing and understanding --

MR. PARRY: Right.

CHAIRMAN STETKAR: -- and sense-making. All of those are rolled into that first bullet?

MR. PARRY: Yes. All of those are rolled into that, yes.

CHAIRMAN STETKAR: Or let's say all, both of those --

MR. PARRY: Yes.

CHAIRMAN STETKAR: -- macrocognitive functions.

MR. PARRY: Right.

CHAIRMAN STETKAR: Okay.

MR. PARRY: They are rolled into that.

CHAIRMAN STETKAR: Because that is the way I started finally --

MR. PARRY: Yes.

CHAIRMAN STETKAR: -- thinking about that in terms of completeness.

MR. PARRY: Right.

CHAIRMAN STETKAR: Okay. And then, the second bullet is what they call decision, and the third one is pretty self-explanatory.

MR. PARRY: Yes. Right. Right.

And when we created these CFMs, we looked at them in the context of the type of subtasks that we think the operators need to do to respond to an upset condition in the plant, in the given plant proceeding. So, with things like responding to an alarm, getting information somehow, evaluating this information, and then performing the execution. So, we are looking at the types of activities that they have to do when we developed the crew failure modes.

And the crew failure modes we came up with are the following: in the plant status assessment -- there is a lot of them, okay, because they are associated really with looking at data and understanding that data, and deciding what it means to an extent.

Now, on the bottom of this, there are two -- okay, let me point it this way; that is probably a little easier. This one has --

CHAIRMAN STETKAR: Because the microphone may not pick you up, Gareth --

MR. PARRY: Oh, okay.

CHAIRMAN STETKAR: -- use the mouse, if you can.

MR. PARRY: Yes, I can do that, I guess.

CHAIRMAN STETKAR: There you go.

MR. PARRY: Okay, this one. All right.

Misread or skipped step in proceeding. I was looking at this the other day. I think at various parts of the report it states it differently. I think, actually, they should go over all three.

CHAIRMAN STETKAR: Yes, and I was going to ask you about that. You need to check the report to make sure it is treated consistently and completely.

MR. PARRY: Yes. No.

CHAIRMAN STETKAR: Because it sort of pops up in the decision trees. Although it is shown on this slide here, it is not shown in I believe the real table in the report.

MR. PARRY: Right. So, I think this really ought to go across all three, and the same with critical --

CHAIRMAN STETKAR: Okay. And that is why I had -- huh? Okay. Okay. It wasn't clear when I looked at the decision trees --

MR. PARRY: Right.

CHAIRMAN STETKAR: -- what column or row it might fit into.

MR. PARRY: There is another table later on in the report which I think is related to the selection of CFM, whereas it does actually across all three phases, and it should do that.

CHAIRMAN STETKAR: Okay.

MR. PARRY: So, that is an error. I think there is also an error in the report that one of the CFMs is missing from one of the tables.

CHAIRMAN STETKAR: Yes. Yes, it is, but that is okay.

MR. PARRY: Okay. We caught it. It is probably right in the presentation, but not in the report.

So, as I say, the types of things we looked at: a key alarm not attended to. And, remember, these are supposed to be modes. So, the way that an external observer could see that an error has been made -- it doesn't say why, okay? So, it is not going into the between-the-ears stuff. It is just an observable.

So, you could tell if the data, if they stopped collecting data prematurely or if they didn't check the data with appropriate frequency, so they missed a specific cue.

So, in terms of the response planning, we have things like delay implementation, misinterpret the procedure. We have one here called choose inappropriate strategy. I think in many we found it difficult to actually find an example of where this would occur in the OPs, but it might be relevant for things like SAMGs, for example, where there are different strategies. They are not prescriptive in the same way.

So, anyway, we left it in here even though we possibly would find relatively little use for this in a real PRA of a Level 1 -- sorry -- of a Level 1 at power.

MEMBER BLEY: I would say, though, if you came up with a situation that was really challenging --

MR. PARRY: Uh-hum.

MEMBER BLEY: -- in that old report, Westinghouse report that was done for NRC, Emily Roth and the rest in that bunch wrote it, where they ran operators through Westinghouse simulators on difficult scenarios that the old CDS thing generated thing for them. They found places where people chose to deviate from the procedures, and many plants give operators a pathway by which they can decide to deviate from procedures. Once you do that, this can crop up.

MR. PARRY: Yes. Okay. But it would be an unusual circumstance, probably a very low-frequency initiating scenario --

MEMBER BLEY: For most plants.

MR. PARRY: -- for most plants, yes.

CHAIRMAN STETKAR: And I thought about, you know, and I agree that it should be in there because, especially when you start looking at fairly-pervasive support system failures. I will keep this to the context of the internal events at-power.

MR. PARRY: Right.

CHAIRMAN STETKAR: But failures of DC power, failures of instrument and control power, you know, could, indeed, I think, prompt that type because the procedures really don't handle them all that well.

MR. PARRY: And I think to some extent, if you have come to the conclusion that the procedure doesn't handle it well, then you may be into a different regime which is more really a diagnostic regime. What the hell is going on?

CHAIRMAN STETKAR: That is okay, but, I mean, this framework has to be able to treat those --

MR. PARRY: Yes.

CHAIRMAN STETKAR: -- because we are in the internal-events at-power box.

MR. PARRY: Right, right.

CHAIRMAN STETKAR: And that is an internal event at-power.

MR. PARRY: It is, and in the language of standards I think it would be more in the recovery portion which is --

CHAIRMAN STETKAR: Well, it is --

MR. PARRY: No, but that is credit, John, because, you know, typically, if you can't find a good reason for the recovery, it is not given any credit.

CHAIRMAN STETKAR: That's right. But, I mean, but from an analyst's point of view, an analyst walking through a particular challenging event scenario and trying to make a decision about, for my evaluation of this scenario, which of these crew failure modes might be appropriate --

MR. PARRY: Yes, it might be appropriate, yes.

CHAIRMAN STETKAR: You know, it certainly, I think, should be on the list.

MR. PARRY: Okay.

CHAIRMAN STETKAR: You might discount the ability of people to effectively respond for other reasons --

MR. PARRY: Yes.

CHAIRMAN STETKAR: -- but I think it should be considered.

MR. PARRY: It is on the list.

CHAIRMAN STETKAR: Yes.

MR. PARRY: So, we are okay with that.

In terms of the execution, they are simple ones. They are just failure to initiate execution and failure to execute it correctly. There could be differences there.

And we have to do a little bit of expansion of the decision trees in this area.

CHAIRMAN STETKAR: Are you, since we are not going to go through all the decision trees, are you going to develop a separate decision tree for the failure to control crew failure mode?

MR. PARRY: That is a good question. That one I think is a little tricky because it is continuous. And no method that I know of has done a very good job on that up until now.

We made some comments on it, and I can't remember --

CHAIRMAN STETKAR: Well, the comments basically say that it is a little tricky and no method has really treated this all that well, and maybe you could kind of think of it in the context of the complex task --

MR. PARRY: Yes.

CHAIRMAN STETKAR: -- you know, kind of discrete, but complex task structure, but I don't think it fits that structure all that well. And the problem is in real PRAs there are typically a number, not many, but a number of these control --

MR. PARRY: Yes.

CHAIRMAN STETKAR: You know, cool down and depressurize.

MR. PARRY: Right.

CHAIRMAN STETKAR: It could be tube rupture. It could be cool down and depressurize, low-pressure injection, initiate RHR cooling, you know. So, it is not something that is a rare type of --

MR. PARRY: No, it isn't.

CHAIRMAN STETKAR: -- human activity in the context of PRA, where you can say, well, the frequency is so rare that we will just ignore it and fail it.

And there are other types of processes going on. That is why I wanted to ask you whether you were going to develop a separate tree for it, because in the report it is kind of just left hanging.

MR. PARRY: Yes, I know. I think it is something we certainly need to look into, because many of these are continual with a continual chance for self-correction. And that is what makes it a little difficult to model it.

CHAIRMAN STETKAR: I mean, you know, we don't want to get into details of the model structure. It doesn't strike me that there are completely different or new performance-influencing factors, or however you want to characterize the branch point names. But how they are presented and how an analyst thinks about the questions for up/down might be very different.

MR. PARRY: Yes.

CHAIRMAN STETKAR: Their perceptions of time available, perceptions of --

MR. PARRY: Right.

CHAIRMAN STETKAR: -- ability to achieve --

MR. PARRY: Right.

CHAIRMAN STETKAR: -- a fixed cool-down rate that may not be constant --

MR. PARRY: Yes.

CHAIRMAN STETKAR: -- and as you said, potential mechanisms of recovery for that.

MR. PARRY: I think one of the issues that we face there, too, is that it is probably hard to define failure because --

CHAIRMAN STETKAR: Well --

MR. PARRY: -- if you don't follow a line exactly, it probably doesn't matter.

CHAIRMAN STETKAR: Failure is not getting to some temperature and pressure by the time you need to get to some temperature and pressure. That is a failure.

MR. PARRY: But if --

CHAIRMAN STETKAR: I mean, in some sense, I don't care if I get there vertically within seven seconds as long as I don't break anything --

MR. PARRY: Right.

CHAIRMAN STETKAR: -- and sit on my hands for two hours, or whether I do it linearly over two hours --

MR. PARRY: Yes.

CHAIRMAN STETKAR: -- or, you know, through some other trajectory. But, indeed, I think you can define failure the same way as you define -- it is failure to meet an objective, a goal of whatever --

MR. PARRY: Yes.

CHAIRMAN STETKAR: -- that process looks -- not necessarily in terms of, gee, I violated a cool-down limit by 6 degrees or something like that.

MR. FORESTER: It just seems to me that this could be treated under the same CFM. It is just part of the conditions for that task. So, you could ask, you know --

CHAIRMAN STETKAR: I think the logic structure might look the same, but I think the questioning and the analyst's sensitivity about what to address in terms of developing a path through the decision tree might be different when you are talking about a control function. I haven't really thought about it. The only reason I brought it up is it was one of the gaps in the report that seemed to be -- it didn't say, well, we are going to develop this. It just sort of was there.

I will let you get back ontrack.

MR. PARRY: Okay.

MR. FORESTER: That needs to be addressed.

MR. PARRY: Yes, we do need to think about that.

Okay. So, in terms of what the quantification model looks like, as you know, we have used a decision tree approach where the decision points relate to the existence of certain PIF categories that have a relationship to the cognitive mechanisms that underlie the CFMs.

The paths taken of each of these branches are dependent on the various specific characteristics that you can use to explore those PIFs.

And then, what we will be doing in the next couple of months is we will be looking at those decision trees and we will be trying to come up with probabilities for the endpoints of those paths through the decision trees.

CHAIRMAN STETKAR: I am going to ask, the methodology presumes that for a particular crew failure mode and a particular PRA scenario I will have one, and only one, path through that decision tree, is that correct?

MR. PARRY: That is correct.

CHAIRMAN STETKAR: In other words, the assessment will lead me on this to one of those four possible outcomes and only one?

MR. PARRY: That is correct. And to that extent, it is potentially conservative.

CHAIRMAN STETKAR: Or potentially optimistic.

MR. PARRY: Or potentially optimistic. Well, yes, I think that the way we try to set up the trees is that on the down branch, which is the good branch, we have chosen to have all the conditions that you need to make it really good.

CHAIRMAN STETKAR: Usually, that is true, but I don't want to get into -- you are getting into "and" and "or" logic under PIFS, and I don't want to go there.

MR. PARRY: Okay. No, no.

CHAIRMAN STETKAR: The reason I brought that up is that, as I went through the early part of the process where you do talk about, very well, the potential pitfalls of that unique, absolute assessment, and as I went through some of the questions and sort of mentally the types of thought process I would go through as an analyst for deciding whether I was on an up or down branch, in many cases it might be really difficult for an analyst to say, "I'm absolutely up" or "absolutely down".

And indeed, as an analyst, recognizing that I want to be conservative, but I don't want to be excessively absurdly conservative, I might have a mental bias that says, well, it is probably good enough to be down, which is where I might be optimistic.

It might be easier for me as an analyst to say, well, you know, I am kind of 90 percent confidence that I am down here and 10 percent confidence that I am up.

MR. PARRY: I can see where you are going.

(Laughter.)

CHAIRMAN STETKAR: And the question is the methodology. I am not advocating that that process -- I think if it is really clear that I ought to be up or down, I think that makes both the analysis process and the quantification more clean, if you will. But in cases where there might be 80/20, 90/10, even 50/50, the methodology doesn't seem to give me as an analyst that latitude. It ties my hands in a sense.

MR. PARRY: I don't think we have got that far down the road to explore that type of situation yet. Because I can think of different ways of doing it. Okay. I can think of saying, "Okay, well, if you are not sure, try both ways." If it doesn't make that big a difference, it doesn't matter.

CHAIRMAN STETKAR: Well, okay. That is one way. But I am presuming that on the bottom path I have like 10 to the minus 100, and on the top path I have 1.

MR. PARRY: Yes.

CHAIRMAN STETKAR: So, it might make a big difference numerically.

MR. PARRY: Yes.

MR. FORESTER: And you have a lot of applications in CBDT where they decide just to average the last branches. They are not really sure where there is a top or the bottom.

CHAIRMAN STETKAR: But I am just thinking in terms of implementing this methodology, should it explicitly allow that type of decision process. Because in terms of unraveling the results, I mean a lot of the traceability through the decision trees and the integration of the decision trees is oriented toward understanding -- I hate to call it "root cause" -- but understanding the causes for a particular error, so that you can go back and examine that.

I don't think you lose that. You now have a set of weighted causes, if you will, but you don't lose that discrimination ability.

MR. PARRY: I don't think we have gone that far down the path of implementation.

CHAIRMAN STETKAR: Okay.

MS. XING: Yes.

MR. PARRY: I think it is a question that we need to think about.

CHAIRMAN STETKAR: Think about it.

MS. XING: Yes, let me make a quick comment on this topic. We actually, from last year, at our meeting, they already asked this.

CHAIRMAN STETKAR: Oh.

MS. XING: It is the very same situation, question that you asked. So, our plan is for now for this measure, the way using this binary simplification, and down the road there are two things that we are planning to do, if we have the budget and resources.

The one, per our users' request, they have being asked if we could put a computer version of this methodology. So, in the computer if we are going to implement in that -- well, even not a computer -- we would like to give the flexibility like using some fuzzy logic for you to decide up-or-down branch and say I am only 90 percent confident.

CHAIRMAN STETKAR: You are getting way ahead of me. I am just a poor, simple, you know, used to be a consultant trying to make a living in this world.

(Laughter.)

I am just talking -- I don't care about computerized bases. I don't care about fuzzy logic. I care about an analyst faced with perhaps a difficult up/down decision, whether the methodology forces the analyst -- forces -- to make an absolute up/down decision or whether the methodology provides some flexibility to say I am 80 percent up and 20 percent down, or vice versa.

MR. PARRY: Yes, as I say, I don't think we have gone that far yet.

CHAIRMAN STETKAR: Yes.

MR. PARRY: The methodology is here to enable that decision to be made. And if there is some uncertainty about it, I think you have to treat it that way, that you are not sure.

MS. XING: Well, maybe you should think about that in the testing, in the upcoming testing.

MR. PARRY: Right. Yes.

MS. XING: So, it gives the analyst --

CHAIRMAN STETKAR: As I went through, I didn't find a lot, the way the decision trees are structured, the way the branch points are structured, and at least the current versions of the questions, in I think a lot of cases the structure is organized such that it is relatively easy to make a binary absolute decision.

But I did find probably a handful of cases where I said, gee, this is a situation where, under some scenarios, it might not at all be clear. And to force my hand to say I am definitely up or definitely down might overly concern me.

MEMBER BLEY: You just hit on a key, and I would like to put Dr. Forester on the spot. He has been suspiciously quiet in this discussion.

(Laughter.)

There was a time I remember you saying, "Gee, if you consider the situation very carefully, the context of what is going on, the status of these PIFs are almost defined by the situation." Do you still think that is true? And I know we flash the word "context" all through Chapters 2, 3, and 4 and 5. Do we get to the point that we try to define that such that, in fact, it makes these things more clear than they appear to be?

And I guess the third part of that question, I would say, if I get in a spot like John was just describing where, under some situations, this could go different ways, does that tell me we ought to break this thing into a couple of pieces, depending on possible situations?

MR. FORESTER: Well, I guess the first part of that is that our goal is to make these binary judgments easy in the sense of the types of judgments you have to make, it should be clear. What John is showing is that this may not always be the case, but that was our goal, and that is what we saw in the empirical studies, was that people had problems making judgments about levels of PSFs. So, we wanted to address that, again, to try to reduce the variability involved in the results that you see.

Yes, I mean, nothing is saying that you couldn't quantify this making assumptions about different contexts and doing a weighted average or something like that. Is that what you were getting at?

MEMBER BLEY: I was wondering, you know, right, no, I don't think about that. I haven't thought about that perspective in this detail, no. I mean, we have defined one of these for each of the crew failure modes, but those crew failure modes can apply in a wide variety of situations, which may make it easier to do this. And I don't know that we have walked people through that very well. But maybe the drill, the test will help us.

MS. XING: Yes, you know, in the upcoming expert elicitation workshop, one test for the expert is to give their judgment on the significance of these individual factors. It turns out most workload is a very dominant, significant factor. We may think about like we want a graded approach.

CHAIRMAN STETKAR: I think that is a little bit different because the workshops are being done, essentially, in isolation from a particular scenario context, right? I mean, you are not developing 10 million different HEP estimates to cover everything.

What I am talking about is within the context of a particular scenario. For example, if there are competing mental models of the plant status, which is an element of one of those crew failure modes, I say, well, gee, you know, from our experience, you probably apply this mental model under these conditions of workload, or whatever, but there might be some possibility that we would apply the wrong mental model. And it is just not clear.

That is not going to come from your workshops, your near-term workshops. The only way that would come is from trying to examine how people struggle in a real implementation of this methodology with making a decision about am I up or down on this particular branch.

You may, when you structure your testing process, try to actively -- I don't know how you would do that a priori -- but actively select some conditions that might be difficult to determine, given the logic structure and the guidance about whether an analyst would assign an up-or-down.

MR. PARRY: I think possibly the more difficult of the branches is going to be the one related to recovery because we don't have specific questions there. The requirement is that they have a convincing case that there is a path for recovery within the tree. That, I think, is possibly the one branch that will give --

CHAIRMAN STETKAR: I don't know. You know, as I said, I kind of tripped over probably four or five just sort of quickly going through things saying -- I came to it initially when I read your methodology discussion. I said, well, gee, let me keep this in the back of my mind and see if I can sort of test that theory. And there were three or four or five or six places where I said, well, gee, this is a case where I think some weighting might apply.

But go on. We will see if we can back ontrack.

MR. FORESTER: Yes, that is a fair assessment because there are situations where competing mental models might be certainly possible, yes.

MR. PARRY: Well, that is something I think we have to take into account as we test it, and particularly as we refine the questions and the trees, too, see whether can be more --

CHAIRMAN STETKAR: Yes. I mean, you know, part of the process might be refining the questions, but, even there, it is not clear to me that you can specify the completeness and the mutual exclusivity of those questions well enough to handle everything. If you can, you are --

MR. FORESTER: Yes, that is going to be -- well, we definitely need to think about it because that is going to be still a potential source of variability in the results.

MR. PARRY: Yes, it is. I mean, I think

we are never going to get rid of all --

CHAIRMAN STETKAR: You will never get rid of all of the variability. But if you run into this problem with higher frequency than perhaps you are anticipating, or even higher frequency than I am anticipating, at least removing the absolute up/down variability and forcing -- you still have to force somebody to say, "Well, I assigned 20 percent down or 80 percent down because...." That would at least -- you still might have the variability that the factors would be reversed, but at least you would understand that and the rationale for that source of variability.

MR. PARRY: All right. Let's go ahead.

Okay. Here is the quantification equation. And this sort of gets to your question about the use of the CRT, John.

Basically, what we are going to do is we are going to do, it is a nested sum. The outer sum is over the CRT, leads to the HFE. So, let me flip to the next slide just to illustrate what I mean by those.

It is the red lines of the CRT sequences. Okay? So, it is the starting with node 4 going through node 7. So, that is the CRT sequence we are talking about.

What we are going to look at is the CFMs related to this first node because this one we are going to treat as a recovery. And it is a recovery that is based on what -- recovery has to be assessed on the failure mode that you are coming into this for because of the potential for dependence there.

Okay. So, for each of those sequences, we are going to look at that initial mode and, then, begin to identify they are all the relevant CFMs for that node, and the relevant CFMs are associated with what are the subtasks that are required to achieve that mode.

And then, given the context implied by the scenario, we will be able to choose a path through the decision tree for each of those CFMs, and we will choose the probability. And that is what we will substitute into the equation. The quantification itself is relatively simple.

Okay? Is that clear enough?

MR. FORESTER: There are some examples as you get --

CHAIRMAN STETKAR: It is clear to me.

(Laughter.)

It took me a while to sort it out, but --

MR. PARRY: Yes.

CHAIRMAN STETKAR: You haven't refuted my mental picture. So, I'm okay.

MR. PARRY: Okay.

MR. FORESTER: A few pages ahead, there is a little example that shows this in a little more detail.

MR. PARRY: Yes. So, the steps in the quantification then, we come into this having created the CRT and the timeline.

The next step, we have sort of understood the critical tasks that are needed. And based on those tasks and the critical subtasks, we are going to select the appropriate crew failure modes and, then, select the DT paths and assign the HEPs. And then, we substitute them in the equation.

The selection of the CFMs, this has been an issue that we have been asked a lot about. Personally, I think it is relatively straightforward as long as you understand what the tasks are and you understand what the CFMs mean.

So, for the CRTs that we are going to quantify, we are going to look at that initial node. As I said in the example that I gave, it was node 4, and we will talk about that one in particular.

We understand what the critical subtasks are, what they have to do to respond correctly. And then, based on the nature of the subtasks, we can identify from the list of CFMs which ones are relevant. Okay. And as implied by the equation, you could have several CFMs for each mode, depending on how detailed a task you have to do.

So, what we have done is created a table that asks certain questions, which is a screening table.

CHAIRMAN STETKAR: Now let me just make sure I have got it straight also.

MR. PARRY: Okay.

CHAIRMAN STETKAR: You just said you could have several CFMs for each node --

MR. PARRY: Right.

CHAIRMAN STETKAR: -- depending on how you have defined the tasks.

MR. PARRY: Right.

CHAIRMAN STETKAR: However, in principle, for a particular crew response scenario, there should be -- should, in principle -- be a defined, in the context of this methodology, a defined set of crew failure modes, right?

Let me go back to my example of three

nodes with 12 crew failure modes for each node versus 36 nodes with one crew failure mode for each node.

MR. PARRY: Right

CHAIRMAN STETKAR: In each case, I have 36 crew failure modes.

MR. PARRY: Yes.

CHAIRMAN STETKAR: I don't have 36 decision trees.

MR. PARRY: Right, right, right.

CHAIRMAN STETKAR: But just to make sure that I understand the process.

MR. PARRY: Yes, yes.

CHAIRMAN STETKAR: That 36 in some sense characterizes the crew response scenario.

MR. PARRY: Right.

CHAIRMAN STETKAR: Regardless of how I have agglomerated them into nodes in the CRT.

MR. PARRY: Yes. There is a law of conservation-of-scale thing.

CHAIRMAN STETKAR: Okay.

MR. PARRY: But it depends on the detailed nature of what has to be done to get success.

CHAIRMAN STETKAR: Yes. Okay.

MR. PARRY: Right.

CHAIRMAN STETKAR: Okay.

MR. PARRY: Yes. Yes, you are correct.

CHAIRMAN STETKAR: Okay. Thanks.

MR. PARRY: So, we ask questions like, for example, in the plant status assessment, "Does the success require response to an alarm?" If it doesn't, then we won't worry about that CFM.

There is another one. I will give you an example in terms of, if the response doesn't require monitoring data, but just requires looking at a moment in time, then the critical data not checked with appropriate frequency might not be a CFM that you would worry about.

If the node that we are talking about is purely a decision node or a transfer to a different procedure, then we are not going to be looking at execution.

If there is a dedicated operator that is dealing with a specific activity, then perhaps communication is not something that we need to worry about.

And as you will notice on this table, we have a mystery, the skipped steps in the procedure being for all phases; whereas, the previous one, we had it only for response planning. That was one of the errors that we discovered in there.

CHAIRMAN STETKAR: Let me ask you one quick one here because I will just forget it otherwise. In terms of, again, back to a link to the cognitive basis, one of the proximate causes that is addressed in that NUREG is effectiveness -- I have forgotten how it is characterized -- but effectiveness of supervision and leadership. You do identify miscommunication or the communication, team communication proximate cause.

MR. PARRY: Yes.

CHAIRMAN STETKAR: There is nothing in this methodology that addresses, at least directly, that supervision. The only reason I brought it up here is that you have got this list of CFMs on the righthand side --

MR. PARRY: Right.

CHAIRMAN STETKAR: -- of the table.

MR. PARRY: Right.

CHAIRMAN STETKAR: Was that an active decision to avoid that or to omit it or --

MR. PARRY: I believe it was, but I would have to go back and check.

CHAIRMAN STETKAR: It is something you may want to think about.

MR. PARRY: Yes.

CHAIRMAN STETKAR: And I didn't have a chance to study the details in appendix -- whatever it is -- E, I think, the last one that has a little bit more of that mapping.

MR. PARRY: I am not even sure we would know how to assess that.

MEMBER BLEY: I was involved a bit in the other report. Emily Roth got pressed on that a bit early on in that area. Since they were developing crew failure modes, it is to include the anticipated supervision and possible failures in the supervision. She felt that the way to define things adequately and embedded that all in that process, but it was discussed. I remember that and she defended it a bit.

CHAIRMAN STETKAR: It might. I didn't think of it, actually, until I got to the end and started circling back. And the concerns are the sense of an autocratic supervisor who might squelch communication, not explicitly, but implicitly squelch communication, or questioning among teamwork.

So, it might be handled in the miscommunication. It might be more important in the context that Harold raised earlier this morning, and that is where you have diffuse or not-well-defined lines of communication among folks who are at an emergency operating center, and the technical support center and the crew in the control room trying to make the decision about what do we do in the next 15 or 20 minutes, with varying opinions.

MR. PARRY: Let me just make a comment in general. The way this methodology has been developed, what you are coming up with is really a very crew-specific characteristic probably. And we are really not looking for crew-to-crew variability at that level.

But I think one of the things that you can use this system for is as an aid to root-cause analysis of actual events. And that might be a question you could ask there. I am not sure that it is a very useful question in a prospective PRA-type analysis.

CHAIRMAN STETKAR: My example of the autocratic shift supervisor, certainly crew-to-crew variability -- I detest the term "safety culture," but I will use it. The safety culture of an organization and the organization hierarchy in terms of affecting decisionmaking could be something that overarches crew-to-crew variability.

MR. PARRY: Yes.

CHAIRMAN STETKAR: And again, I think it

applies a little bit more in the context of Harold's examples, but it is not clear that it wouldn't apply within a smaller context or specific scenarios.

MEMBER BLEY: I would like to toss something in on here in both areas. I think they are two separate things. The one Harold brought up I think is really important. And certainly, when we begin to look at moving into the SAMGs, the operating procedures become crucial. If, in fact, this methodology is going to be used for the Level-3 PRA, I believe they told us they were going to look at SAMGs, and somebody had better be giving that some thought because that is going to be a key issue.

On the other point, the way I have seen people handle the first issue raised, and all of us have run into those people in power plants, there are processes now, some INPO-driven, that are trying to reduce the chance of that, but it still happens. And you still see events occurring because of that.

The only way I can think of handling it is in the uncertainty analysis. You haven't gotten to that yet, but when I look at that chapter and this report, it kind of says, well, there's some stuff around we can look at and some things that need to be developed. Well, yes.

(Laughter.)

I am going to jump ahead because in that integration part there were a couple of introductory things, I think. The dependency part I thought was really nice and covered a lot of ground. The uncertainty part was woefully inadequate and just kind of punted. I hope, before you are all done, that doesn't stay punted.

And there are references there to things existing and going on that could certainly be pulled in to be considered, and the stuff Nathan did way back on the PTS stuff certainly gives you ways to catalog and think about those things.

But, yes, you could certainly cover it under the uncertainty when you look at the crew response because those kinds of people do exist. We hope they are getting less and less prevalent out there, but they are still around.

CHAIRMAN STETKAR: Yes, there are a couple of ways I thought of being able to treat it where it could be handled or treated.

MS. XING: And in the generic methodology, we did a lot of work for the teamwork and the supervision. So, we had objectives that identified for these two functions, and we are still working on the PIFs that would contribute to these two functions.

MEMBER BLEY: There is a whole world of literature on how crews work. You could do something elaborate, but I don't think you need that. I think knowing how these things are handled, could occur, you can cover it in the uncertainty analysis.

CHAIRMAN STETKAR: It is also in terms of, as I said, some of the things I was thinking about as I read through the report is instilling a sensitivity at least to the analyst. How it is handled numerically, whether it is part of the uncertainty analysis, whether it is part of a weight on an up/down because, you know, accounting for does crew-to-crew variability affect your decisions about are you up or down at a particular branch point.

MR. FORESTER: Yes, in principle, it could affect --

CHAIRMAN STETKAR: I don't know, but I think there are a couple of different ways it could be addressed, but it didn't seem to be addressed.

MR. FORESTER: It is sort of that general issue in HRA that everybody is sort of aware that there are crew characteristics that probably need to be considered, but it has to be a systematic effect. You can like say six crews do this and six crews do that. And also, you have to have the ability to be able to find out how the crews behave. So, it is one of those real challenges of HRAs. It is a very difficult thing to evaluate, too.

MR. PARRY: Certainly a priori.

CHAIRMAN STETKAR: Yes, certainly.

MEMBER RAY: Well, look, at the end of the day, this to me is important because, hopefully, it will let people choose between relying on human performance versus a modification to the plant. And somehow, I mean, I think the uncertainty that Dennis spoke to is probably more useful in my mind in that regard, to focus on, well, it may work out, but it may not. And we ought to fix the thing with a design change instead of simply issuing another SAMG in case this happens.

And it is that decisionmaking that is what we ought, in my judgment, be intending to facilitate, to be able to make decisions whether about the plant should be modified, so as to not rely upon human performance because of this possibility.

Now, trying to quantify what that uncertainty is I appreciate is very difficulty, as you say. But recognizing that it exists in some way seems to me to be essential.

MR. PARRY: I think you raise a good point. I think it raises another sort of application for this method. Because the way I look at what we have developed, it is like a qualitative description of the way people make mistakes. And you could use it in lots of different ways to, as you say, look for under these circumstances is there anything I can see in this model that would flag itself as a source of uncertainty that maybe I need to explore to see whether I need to do something different.

So, I think it is true. I mean, we have been sort of intending to focus on this as a tool for HRA for use in PRAs that are, you know, used for risk-informed decisions and all that sort of thing. But I think we can get more out of the set of decision trees and the construction there, the CRTs, than just numbers. I certainly hope so.

MEMBER RAY: Yes. All right.

MR. PARRY: Should I carry on for this one?

CHAIRMAN STETKAR: You may.

MR. PARRY: Just to give you an example, again to go back to this node for a CRT we have, the essential activities there are monitoring the narrow range levels, total feedwater flow, comparison of the criteria, and making a decision, based on an understanding of what that is.

But all the CFMs that we identified in that table are relevant except for the ones that clearly aren't. There are no alarms involved. So, we are not going to worry about that one.

There is not inappropriate strategy in this case. You go into FRH-1 or you don't.

It is not an execution step. So, we don't need to worry about those.

And we have put down here miscommunication because we have assumed it is a dedicated operator who is doing this, and therefore, he doesn't have to communicate with other people. But, other than that, all the CFMs are potentially relevant for that node.

I want to say a couple of comments about the treatment of recovery. On many of the decision trees, we have a branch related to recovery. What these are is these are potential opportunities for the crew, given that they have made an initial error in a particular failure mode, for them to recover before the time window has expired and that the function has failed.

I don't know if I have gone into any detail. No, I haven't gone into detail on this. I think there is some on the specific decisions.

But the idea behind this recovery branch is to say, okay, now I am going to follow this path. I am assuming they have missed this one. I am going to see what happens as time progresses. Do they get new cues? Where are they going to be seated? Do those new cues lead them to rethink what they are doing and get back on track? I mean, in a very broad sense, that is what that means.

So, we discussed this whole issue of recovery separately, not individually on each of the decision trees. And the idea behind it is that, for somebody to take credit for recovery, then they have to be pretty convinced that, even if the first mistake is made, they have got opportunities for them to get back on track and that those opportunities allow them to do so before the time window expires, and that it has to be convincing enough.

Like, for example, if they formed the mental model early on, which is one of the cognitive mechanisms underlying some of the decision trees, the new information that they get and the guidance that they have has to be sufficient to get them to change that mental model. So, it has to be really a credible recovery path.

This is a little different from the way recovery is treated in something like CBDT, for example, where it is treated as an add-on at the end and, basically, it is self-review, an additional cue, an additional crew member. There are a number of different recovery factors that you can apply.

This, I think, because it is more related to the cause, there is a dependence between the recovery mechanism that in terms of the HFE and the initial failure can be handled a little more coherently.

MEMBER BLEY: I have read your top bullet and read things about this and heard a discussion a couple of times in the past. And I like the concept, if I have got it right. To me, the concept is you don't just come along and say, "Oh, there are these ways to recover. We will multiple by .1" or something, the chance of fail to recover.

You go back and you consider the whole scenario --

MR. PARRY: Right.

MEMBER BLEY: -- and embed this in it. And to me, that allows you to do many things, including what I raised earlier, the thing I talked about with timing. Even if you don't do detailed analysis on all the timings, procedures themselves have multiple paths to get you back.

MR. PARRY: Right.

MEMBER BLEY: And now, you consider that thing as an integrated piece --

MR. PARRY: Right.

MEMBER BLEY: -- and say, given all these possibilities and the range of times, and that might take some work. I mean, it is not just, yes, it looks like you can do it in time. It gives you a holistic way to do that analysis. It makes sense to me.

CHAIRMAN STETKAR: And I agree. I kind of like this construct.

MR. PARRY: Okay.

CHAIRMAN STETKAR: Except --

MR. PARRY: There had to be an "except".

CHAIRMAN STETKAR: There had to be an "except".

(Laughter.)

When I got to the end, to the action execution crew failure mode decision trees, there is a discussion about time available, and there is a discussion that essentially says, well, we recognize that the way we are treating recovery in, let's call it the front-end cognitive part of the process --

MR. PARRY: Right.

CHAIRMAN STETKAR: -- has inherently these recovery issues built into it, and, yes, the recovery could take some time, but we think that the frequency of those conditions is really small, and we are going to ignore the effects of that time. So, therefore, don't worry about the amount of time that might have been consumed during recovery in the first part of the thing. Just use the nominal available implementation time for your assessment of execution. And at that point, it feel apart for me.

MR. PARRY: Well, yes.

CHAIRMAN STETKAR: Because the timing information, as Dennis mentioned, is really important.

MR. PARRY: Yes, and proper.

CHAIRMAN STETKAR: And for a particular crew response scenario, if I accept this notion that I will construct and evaluate these trees in an absolute up/down fashion, I will have a particular path through each of those decision trees.

MR. PARRY: Right.

CHAIRMAN STETKAR: And therefore, I will know whether or not that path includes recovery. I am either on a recovery up or a recovery down branch. And recovery is in many of those trees.

So, I could construct a timeline accounting for how much time I have consumed on everything up until the point of execution, including whatever recovery I have assessed, and then re-evaluate do I have sufficient time available for implementation. This is this, oh, my God, by the time I have taken credit for all the recovery I wanted to take credit for in the cognitive part of the process, I have now suddenly gotten to the point where execution is not feasible anymore, or that I have much higher time pressure. And I didn't see that notion.

MR. PARRY: Well, okay. I think I know where you are coming from. I think at one point I created like an event-tree version of the CRT which tried to demonstrate the fact that, if you start worrying about, to get to this endpoint, I have already eaten up all my recovery time earlier on, that ought to be a pretty low -- remember, that is preconditioned on having committed an error in the first place. So, you are already down low in probability.

The dominant path would be where everything works fine until you are at the execution, and that is when you screwup.

CHAIRMAN STETKAR: Well, that is one

mental model, the way the world works. Another mental model is that people under certain scenarios might not behave all that well for their initial cues and responses until, 30 minutes into the scenario, the big claxon goes off and reminds them and they eat up another 15 or 20 minutes figuring out what to do for the claxon.

So, I drew myself a little event tree also.

MR. PARRY: Yes.

CHAIRMAN STETKAR: And I think it comes down to a scenario-specific assessment, and that if, indeed, your evaluation of the crew failure modes -- I will call it up in the cognitive part of the model -- for whatever reason, has forced you to, because of the scenario context, has forced you to invoke the recovery parts of those models, and it is in a large fraction of those decision trees, somehow by the time that you get to that execution, you ought to account for that, I think, because they do take time.

MR. PARRY: They do take time. Again, I fall back on, though, the relative frequency with which you are going to breeze through the first ones to get to the execution, and the case where you have committed an error and recovered it and then reduced that amount of time. It is a balancing act.

I mean, I agree it is --

CHAIRMAN STETKAR: And they overfilled their steam generators because they didn't have enough time to do something by the time they figured out -- you know, it happens.

MR. PARRY: Well, in that case, they should have, I think the problem is --

CHAIRMAN STETKAR: It is old, but, you know, it does happen.

MEMBER BLEY: Like the railroad. I will be quiet after that.

MR. FORESTER: Yes, I don't know. The only way, it is going to increase the complexity, but sometimes it may require building another one of those time criterion --

CHAIRMAN STETKAR: I actually don't think it is -- I thought about this, and it is not clear to me that it increases the complexity of the evaluation because you require an analyst for each crew response scenario to establish the path through each relevant decision tree that applies for that scenario.

And in your absolute up/down context, it is a specific path. It is one, and only one, path.

You know whether or not you have invoked recovery, and you should be able to assess the amount of time that is allocated for that recovery.

If it were a fully-linked event tree, a sequentially-linked event tree where you had probabilities, up/down probabilities, at every branch point in that event tree, that would become a very, very difficult assessment because you would have a separate available time window for the execution for, in principle, each of those invoked scenarios.

MR. PARRY: This is an account of that, though, John. If we are adding an HEP to this scenario, we are not even getting to the execution because it didn't get there. That is done only with this one.

CHAIRMAN STETKAR: That is right.

MR. PARRY: So, we are only really adding it on this one, where, in fact, we are assuming they got to that point in the proceeding. So, I think even if they don't take credit for recovery, or even if they do, we would still be failing them at that point because there is still a likelihood, even if they have a recovery path, that there is a failure that they --

CHAIRMAN STETKAR: Absolutely. I have got that. My little event-tree structure --

MR. PARRY: Okay.

CHAIRMAN STETKAR: -- has that. What I am concerned about -- and I don't have the mouse here -- is that, take your down branch on No. 5.

MR. PARRY: Okay.

CHAIRMAN STETKAR: And let's say there is a 90 percent chance that that happens, for whatever reason, because I have got bad karma or I have got bad indications or I have got bad something or other, and that there is a 90 percent change that the up branch on the 8 applies, but it takes me 45 minutes by the time whatever prompts successive 8 comes in.

Now, by the time that I get out to the end of 1, I am 45 minutes later than if I had been on the up branch of 5. Those are the concerns. I don't know, in practice, how many times that will happen in the real world.

MEMBER BLEY: It is possible. But I don't see any reason why if you, then, treat it integrally, you wouldn't account for that probabilistically. Maybe I missed your point.

CHAIRMAN STETKAR: My point is that, because the decision trees have, right now in the construct, unique branches through them -- and we are going to have to break here -- I know for a particular crew response scenario that I am on the path down five, up eight.

MR. PARRY: Well, up eight doesn't take you anywhere in this case. But as far as the trees go, there is no -- it was going to be just adding a probability of failure, even if there is a possibility to recover, right? What we take credit for is that HEP is low from that failure mode, not that it is zero and not that it is success.

I think you are thinking mechanistically about what is going on as opposed to the way we are using this to come up with a fundamental basis.

CHAIRMAN STETKAR: I am not sure, and we will probably have to pick this up after lunch because we do have to break.

MEMBER SCHULTZ: It would be helpful to get that answer.

CHAIRMAN STETKAR: Yes.

MEMBER SCHULTZ: Or a lot better understanding.

CHAIRMAN STETKAR: Because we do have another meeting that we have to attend, several of us, so let's recess for lunch, and we will pick this up. We will try to finish this up within the first 30 minutes after lunch, the rest of your -- can you do that?

MR. PARRY: Well, it depends on how much you want me to go through it. If you want me to go through the decision tree as well --

CHAIRMAN STETKAR: I want to make sure that we allocate enough time for the other topics that we need to discuss, which maybe you folks can sort of figure out roughly how much time that will take. I mean, we can run a little bit late. Fortunately, we have all day. But I don't want to run too late.

Anyway, let's recess for lunch and reconvene at 1:00.

(Whereupon, the foregoing matter went off the record for lunch at 12:02 p.m. and went back on the record at 1:04 p.m.)

A-F-T-E-R-N-O-O-N S-E-S-S-I-O-N

1:04 p.m.

CHAIRMAN STETKAR: We are back in session.

Let's see if we can try, if possible, to finish up the discussion of the decision trees in about a half an hour or so, if that works. I know that requires discipline on this side of the table.

MR. PARRY: Just to revisit the issue that we were talking about, I don't know that it is the right forum for us to discuss it in any technical detail.

CHAIRMAN STETKAR: Okay.

MR. PARRY: I would just like to make a comment, though. I think that, to some extent, what you are focusing on is what I would classify as a second-order effect on the first-order approximation. And that is what I would hope to prove to you.

CHAIRMAN STETKAR: And I think I agree with you that we should probably pursue it outside the context of this meeting because we might get into such details --

MR. PARRY: Right.

CHAIRMAN STETKAR: -- that it will be a real time issue.

I am not convinced that it is a second-order effect.

MR. PARRY: Okay. Well, then, it would be our job to convince you.

CHAIRMAN STETKAR: That is right.

MR. PARRY: Okay.

CHAIRMAN STETKAR: That is right.

And I wasn't convinced, you know, just for the simple -- there is like only a paragraph. It at least addresses the issue, but that is why I brought it up.

MR. PARRY: Yes, there was at one time a longer discussion of it, but I think we took it out deliberately from the report.

CHAIRMAN STETKAR: Okay.

MR. PARRY: We thought it might be confusing.

Okay. So, I think we went through the quantification process. Just a couple of comments on the integration into the overall PRA model.

As I said, what we have done is created a model that we can use to estimate the HEP for a single defined human failure event. And so, obviously, since we are doing it one HFE at a time, we have to be concerned about the issue of dependency, which Tony has brought up this morning.

There is some discussion in the document about the issues related to dependency and what they are and how they should be dealt with. We haven't really fully developed this in any way, but I think we just want to make the comment that, because we are looking at causes, and particularly because we are looking at underlying causes like cognitive mechanism, I think this method has the potential to deal with dependency in a much more comprehensive way than is typically done by just looking and seeing if it is the same crew, the same timeframe, and, you know, proximity and time, those sorts of things.

So, I think it is a promise right now, but I think we can see ways of making it work. Well, just a few additional comments on that, and I think they are reflected in the document.

So, the way we were doing it is looking at why and how the first HFE could occur in terms of the crew failure modes and the specific PIF characteristics that drive them. And then, develop a complete picture of the overall scenario, much as we were doing when we were talking about recovery within the HFE. We are talking about developing the whole picture of what is going on in terms of the responses subsequent to the first failure.

The idea behind using this particular construct to address dependency is to really look deeper at the cognitive processes that underlie the CFMs, both the initial one and the subsequent ones, and use the information that you can extract about that to examine the potential for, say -- perhaps the simplest example is if, having an incorrect mental model is the cause of the first HFE, is that likely to continue into the scenario when you are up to the second response and, therefore, will cause an increased likelihood of failure in that regard?

But, as I say, we haven't really developed this to any degree yet. We are still working on finalizing the model to the single HFE. But I think we see how it can be assumed in that regard.

Okay. So, let me, since we didn't present all the decision trees, we wanted to at least present one example and try to explain our philosophy to some extent.

The example we chose for this was the critical failure mode of critical data dismissed or discounted. And the definition of the CFM is that, basically, the crew knows that they are supposed to get this piece of information, and they have got it, and the piece of information could be the value of parameter of status of a pump or a valve or something.

But they have discounted it from the assessment of the plant status and, therefore, got an incorrect synthesis of the information they have, got the wrong mental model of what is going on in the plant.

And we would apply this CFM when that information that they have dismissed is an essential part of getting the right concept. So, you look at the scenario. You look at what is needed. You can identify a piece of information that, if they dismiss, they will get the wrong picture.

The tree that we have developed for this, it looks somewhat busy. What I want to do is to go through each of the branches here individually to explain the philosophy behind these.

The first one, which is whether there is a valid alternative or a deviation scenario, on this event tree I will say in some cases, like with the second branch point, expectations are not formed. It is not an obvious title. Some of these titles were changed to try to turn the trees into a uniform no/yes-type structure. And I think we have a few double-negatives in there somewhere. So, it is a little confusing.

CHAIRMAN STETKAR: Double-negatives I think is something --

MR. PARRY: Well, certain negatives would be good.

(Laughter.)

Okay. Let's look at the first branch point, then, which asks if there is a valid alternative or deviation scenario. And basically, this is really a systems analyst question that he has to answer. It is not really an HRA one.

It is an understanding of the signatures of the various scenarios that could exist. So, what we would be looking at is, is there a scenario that if this particular piece of data were dismissed, is still a valid plant status? Okay. And if we can't find something like that, this tree basically would not apply.

Okay. So, a precondition for this almost has to be that there has to be a rational -- it is a rational decision to do this, but it has to be a scenario that is an appropriate scenario, if that piece of information is dismissed.

The second branch point asks whether -- it is written, "Expectations not formed or irrelevant." So, what this does is, it is really trying to look at the bias that you get from training or knowledge and experience with respect to the plant status.

And the question is designed to determine whether the training is -- if there is a competing scenario whose signature is the signature that you have got minus this piece of information, is the training sufficient to bias you in favor of the incorrect scenario? In other words, can you find a justification, based on your training and experience, that this data is probably irrelevant? And that is the basis for that. And we have seen that type of method or bias, and you will see it in the literature report.

Okay. The third branch point we have is whether the indications are reliable. And again, the purpose of this is to address the fact that, if there is some feeling that that indication might be unreliable, then it is more likely to reinforce the decision to reject that piece of information than if it seen to be a very highly-reliable piece of information.

CHAIRMAN STETKAR: Gareth, I know we agreed we aren't going to go through all of the trees, but you happened to select one that I had two questions on. So, I can't avoid this.

MR. PARRY: Okay.

CHAIRMAN STETKAR: And it is more of my trying to understand some of the rationale rather than --

MR. PARRY: Okay.

CHAIRMAN STETKAR: -- rather than, in this sense, the specifics. For the decision regarding this particular branch point No. 3, the first question as an analyst that I am asked to answer is, "Is the indication potentially ambiguous and a reason can be postulated why the indication is not accurate?"

I don't understand why that is an "and".

MR. PARRY: Perhaps it should be an "or", yes.

CHAIRMAN STETKAR: For example, the reason I highlight this is I had several questions regarding the "and" and "or" logic and the questioning, and at a high level, rather than trying to go through each of those --

MR. PARRY: Right.

CHAIRMAN STETKAR: -- which I agree we don't have the time and it is not the forum to do that.

But was there any type of conscious decision in terms of creating the "and" and "or" logic to avoid inappropriately excessively-conservative decisions? Do you follow me?

Because, see, this particular "and" logic would tend to force people into the good part of the tree --

MR. PARRY: Right, yes.

CHAIRMAN STETKAR: -- rather than the bad part of the tree if it were an "or".

And the only reason I brought it up here is because, if there is some subtle or explicit reasoning in that process to point people toward good things, because you didn't want them to be excessively conservative --

MR. PARRY: No.

CHAIRMAN STETKAR: -- we could really examine that stuff.

MR. PARRY: No, I don't think that is the intent, and this may be a case that, by changing the logic and changing the way the null TS, we may have forgotten to change the "or" to an "and" or an "and" to --

CHAIRMAN STETKAR: Okay. Okay. Yes.

MR. PARRY: It is possible. But I would have to look back at that.

CHAIRMAN STETKAR: There were a number of these where I had questions about "and" and "or" logic, and it was primarily "and's" --

MR. PARRY: Right.

CHAIRMAN STETKAR: -- that to satisfy all of the "and's" would make it really difficult for me to go in the bad direction, to go in the up direction.

MR. PARRY: Yes, we have tended to, I think, lean more towards saying that a lot of the "and's" go down. In a sense, you need all these to make it --

CHAIRMAN STETKAR: Yes. In many, many cases, that is absolutely true.

MR. PARRY: And that is what we tried to do.

CHAIRMAN STETKAR: Yes. Okay.

MEMBER BLEY: I will bet you are right about when you flip it.

CHAIRMAN STETKAR: It might be --

MEMBER BLEY: Yes, the logic.

CHAIRMAN STETKAR: -- if you were flipping the stuff around, these might be residuals.

MR. PARRY: And that happened late in the process --

CHAIRMAN STETKAR: Okay.

MR. PARRY: -- while we were putting it

together. So, it is quite likely that that has happened.

CHAIRMAN STETKAR: Let me just say, you know, march through those things pretty --

MR. PARRY: You know, we will do.

CHAIRMAN STETKAR: -- carefully to see --

MR. PARRY: Right.

CHAIRMAN STETKAR: -- to check that. Okay.

MR. PARRY: No, the intent, actually, is to have to prove pretty definitively that you are going down. And if you don't meet any of the conditions to go down, you go up.

CHAIRMAN STETKAR: Yes, yes.

MR. PARRY: Which is conservative, but, on the other hand, it is probably the cleanest way of doing it.

CHAIRMAN STETKAR: Sure. Sure.

MEMBER REMPE: I have to ask -- again, I don't do this at all as a living -- but how would an analyst assess the ability of the crew to understand that something is unreliable? I mean, do they go through and assess their knowledge of the state of the procedure? I mean, how does an analyst do this with any sort of certainty and repeatability? You know, it is very fuzzy.

(Laughter.)

CHAIRMAN STETKAR: Now you are coming back to the 2080 stuff.

MEMBER REMPE: I have been reading in these documents. You know what I am saying? And is that a reasonable question?

MR. PARRY: It is a reasonable question. And what we are saying here, all I am doing here is explaining the branch parts. Okay? Underlying that there are questions that the analysts are supposed to ask that enables them to get to that point. And we are trying to make those as objective as possible.

CHAIRMAN STETKAR: Yes. And those questions, by and large, or the vast majority of them, seem to be much more objective criteria.

MR. PARRY: Right.

CHAIRMAN STETKAR: Although I admit they still don't -- I can still see us getting into the 2080 confidence --

MR. PARRY: Yes.

CHAIRMAN STETKAR: -- issue.

MR. PARRY: Around those questions, really, what the analyst is expected to do is to talk to plant staff and the operations staff to get the answers to those questions. But they are an attempt to be objective ways of doing this. Whether we have succeeded or not is --

MEMBER REMPE: It is a well-documented science. Okay.

(Laughter.)

MEMBER BLEY: If you go into any plant and talk to the training crew and the senior operators at that plant, they will tell you the stuff that they always have trouble with and the stuff that their own people have had trouble with, the instruments that --

MEMBER REMPE: So, it is very plant-specific is what I am hearing.

CHAIRMAN STETKAR: No, it is plant- and scenario-specific.

MEMBER REMPE: And scenario-specific.

CHAIRMAN STETKAR: Yes.

MEMBER REMPE: So, when you do a tabletop exercise, which I have heard is going to be coming down the pike here --

MEMBER BLEY: With all people who aren't at the plant.

MEMBER REMPE: -- who are a generic plant --

MEMBER BLEY: You won't get that stuff

right.

MEMBER REMPE: And how would that give you guys information? Am I --

MR. PARRY: Well, we don't know that we won't involve people at the plant. I mean, it is --

MEMBER REMPE: But you will pick a particular plant when you have this exercise?

MR. PARRY: Yes, but that is to see whether you can get that type of situation.

MEMBER REMPE: Okay. It will be interesting. Again, I am from a distance in my field. And so, I am kind of just trying to understand and comprehend this stuff.

MR. PARRY: In some cases, I mean, I think what Dennis says is right. You talk to people and they will know what is reliable and what is not.

CHAIRMAN STETKAR: I have got pages of things here. But one of the things I tried to do is to, at a fairly high level, measure the objectivity of the questions --

MR. PARRY: Right.

CHAIRMAN STETKAR: -- and whether or not they would elicit, first of all, the desired sort of absolute yes/no, up/down kind of response and, second of all, were there any inherent biases in their completeness and things like that. And they are not. They are never going to be perfect.

MR. PARRY: No.

MEMBER BLEY: You follow DOE a lot. Did you follow the time the folks got into Oak Ridge recently and cut through the fence and got inside?

MEMBER REMPE: It is actually another part of the Y12 complex.

MEMBER BLEY: Did you hear what the guys in the security office said about the alarms?

MEMBER REMPE: Yes, I know, and it is very --

MEMBER BLEY: They said, "We stopped paying attention to them years ago. They go off all the time."

MEMBER REMPE: It is probably not the only place in the DOE complex that that occurred, too.

MEMBER BLEY: But the same thing, if you were analyzing that complex and got in and talked to them, you would have learned that. And if you go into the power plants, you are working with them.

MEMBER REMPE: Okay. I believe it is very specific to the plant.

MEMBER BLEY: If you don't go into the plant, most of that kind of stuff you can't get.

MEMBER REMPE: But we are going to have an exercise here with people from different plants and non-plant people, and I am not quite sure how it is going to help.

MEMBER BLEY: Well, one would hope that some of those resource experts are associated with the plant if you are doing their plant.

MEMBER REMPE: If their plant --

MEMBER BLEY: If they are not, then they will miss that part.

MEMBER REMPE: But they will be doing it for one plant, and another group will be doing it for another plant?

MEMBER BLEY: I don't know the details in that. Are you talking about the expert elicitation?

CHAIRMAN STETKAR: I think we are mixing exercises here. The resource experts are at the expert elicitation, which has nothing to do with evaluating a particular set of instrumentation for a particular scenario.

MR. FORESTER: No, it would just be the question being given that if the crews think this is unreliable, then what would the HEP be kind of thing.

CHAIRMAN STETKAR: Right. Yes. That is right. Yes. Given a particular trajectory through that particular tree, what would the HEP be?

Joy is talking, I think, more about the EPRI exercises, I think, that you were talking about --

MR. PARRY: The piloting of that.

CHAIRMAN STETKAR: Piloting, or whatever you want to call it.

MEMBER REMPE: I am just kind of wondering how you will get useful information.

MR. PARRY: Well, you just see if it works. I mean, the one question we will have is, can we objectively answer this question? And if not, then maybe we can clarify the question.

None of these questions are the final word yet. I mean, I think we need to work on them to make them -- we need more input, I think, on how to make them do what they are supposed to do, which is help you make objective decisions.

MEMBER REMPE: It will be interesting. I just had to express some cynicism.

CHAIRMAN STETKAR: It will be interesting.

MR. PARRY: Yes.

CHAIRMAN STETKAR: See, until they got to the point where we are, where you have -- I hesitate to use the word "coherent" -- but a coherent, complete methodology that can actually be tested --

MR. PARRY: Right.

CHAIRMAN STETKAR: You could never finish if you tried to make everything perfect.

MR. PARRY: Right.

CHAIRMAN STETKAR: So, I think it is primetime to get it out and ferret out some of the concerns, whether they are valid or can they be mitigated.

MR. PARRY: Right.

Okay. Well, let me go onto the next branch point then. And this one is sort of like a recovery-type approach, a specific one, though.

And this asks, really, it is about the work practices; first of all, more than the work practices, but is there confirmatory indications that would confirm the piece of information that is being dismissed? And is it customary to look for it before you make a decision? And if neither of those exist, then you can say you don't have confirmatory indications. There is no reason why they should revisit their decisions to dismiss the piece of information.

CHAIRMAN STETKAR: I think I got that. Can you go back to the tree structure?

MR. PARRY: Sure.

CHAIRMAN STETKAR: Within that context and the more detailed explanation and the questions for that particular branch point, I think I understand why on sequence 15 there is no branch at recovery potential, because of the things you just said.

MR. PARRY: Right.

CHAIRMAN STETKAR: That we have goodness, goodness, goodness.

MR. PARRY: Yes.

CHAIRMAN STETKAR: And this is a de facto recovery.

MR. PARRY: Right.

CHAIRMAN STETKAR: But the question is, logically, why do I ask recovery potential again now on the 3, 4, 7, 8, 11, 12 cases, where I have now this wonderful confirmatory information that I have that I can use to make the world wonderful, and yet, I can even make the world more wonderful with additional recovery?

I didn't get that. And the reason I bring it up, it is a detail question for this tree, but I also had several questions of the branching logic in several, not several, a few other trees where it wasn't clear to me what the rationale was used for either omitting or including a particular branch point.

MR. PARRY: Okay.

CHAIRMAN STETKAR: And so, what I am asking is not -- I don't want an explanation of this one because of the time --

MR. PARRY: No.

CHAIRMAN STETKAR: -- but what type of -- I hesitate to use the word "peer review" because you probably don't have a peer (laughter) -- but what type of review has been done in terms of the basic logic of these decision trees?

MR. PARRY: At the moment, it is really a small group of us that looked at it: John, Stacie, myself.

And I think you are raising a very good point. Because while I was preparing a presentation for EPRI last week, I was looking at a couple of these trees, and I was thinking to myself, you know, I might delete some of those branches.

And I think the exercise we are going to go through next week, which is the expert elicitation, where these trees have to be explained, it will give us another opportunity to --

CHAIRMAN STETKAR: I would hope that is

the case, as long as they are presented in that context, rather than these are firm, well-reviewed, accepted decision trees; now you do your job in the context of this model.

MR. PARRY: No, I mean, because I think what we may find out is that, and what I hope we find out from the resource experts is that, well, given this and that, I don't even care about that issue.

CHAIRMAN STETKAR: Yes. Okay.

MR. PARRY: In which case, we can delete --

CHAIRMAN STETKAR: That is sort of the nature of this question --

MR. PARRY: Right.

CHAIRMAN STETKAR: -- for this particular tree.

MR. PARRY: Yes, yes.

CHAIRMAN STETKAR: In other cases, there might be, "Gee, why isn't there a branch," you know --

MR. PARRY: Why isn't there, yes.

CHAIRMAN STETKAR: -- "on sequence 15," for example?

MR. PARRY: Right.

CHAIRMAN STETKAR: Okay.

MR. PARRY: No, that is a good point. I

think I wouldn't necessarily portray these as, the detailed tree structure as absolute right now. I think the PIFs that we want in there are the right ones, given that we have used the -- according to the cognitive literature -- But how they work together is not necessarily science.

CHAIRMAN STETKAR: I mean, as a practical matter, a large fraction of the trees -- and I didn't do the body count -- have essentially what I call a branch everywhere logic anywhere.

MR. PARRY: Yes, yes.

CHAIRMAN STETKAR: So, in terms of these types of logical relationships --

MR. PARRY: Yes.

CHAIRMAN STETKAR: -- you don't see the potential for omissions or extra inclusions.

MR. PARRY: Right.

CHAIRMAN STETKAR: I was just trying to apply some of the other, the basic cognitive stuff, to understand if there was a particular mental model. In the creation of the tree logic, you know, has that model been challenged by other reviewers, other people?

MR. PARRY: Not really.

CHAIRMAN STETKAR: What I am hearing is

you are going to rely --

MR. PARRY: Yes.

CHAIRMAN STETKAR: -- on whatever you call them, the workshops that are coming up.

MR. PARRY: Yes.

CHAIRMAN STETKAR: Okay. I hope that, again, the participants in those exercises recognize that they are also critiquing the structure of the logic model --

MR. PARRY: Yes.

CHAIRMAN STETKAR: -- clearly understand that, rather than just using this as the fixed logic to guide their process.

MR. FORESTER: Yes, there are some worksheets and there are some explicit questions about --

CHAIRMAN STETKAR: Okay.

MR. FORESTER: -- is there anything missing here.

CHAIRMAN STETKAR: Okay.

MR. FORESTER: And we can reinforce that.

MR. PARRY: Okay. Let me just briefly, then, talk about the recovery potential, although we have talked about it a little bit before.

I have got nothing more about, any more

detail about that. But, essentially, this is going to be a tricky one. I think it is the branch point that can be misused --

CHAIRMAN STETKAR: Yes.

MR. PARRY: -- to the greatest effect, I think.

But the idea is to take the "yes" branch on this -- and it doesn't mean to say that there won't be a human error probability; it just means that it is lower than the one where there is no recovery -- basically, what the analyst has to do is to do a complete analyst of the HFE from beginning to end, following the failure paths and seeing whether there are opportunities to recover, as I mentioned earlier. And they could come from following procedure. They could come from new alarms. They could come from any number of things. But, to be convincing, the story has to be, well, the story has to be convincing to take credit for the recovery.

And with that, I think we could probably draw to a close.

MR. FORESTER: I think so, if there are not any more questions.

CHAIRMAN STETKAR: I don't have any. Any of the other members have any questions?

(No response.)

If not, thanks. We got through a lot of material in a good amount of time. We appreciate that.

I think that we will be really interested to see how you do with the application of this, the test stuff to see.

(Laughter.)

But you have come an awful long way since we last got together in terms of pulling things together.

MR. PARRY: Yes, I think we have got a relatively-coherent story up to this point.

CHAIRMAN STETKAR: Yes. You know, I tend to agree. It is to the point where somebody should go out and really test it.

MR. PARRY: Yes.

CHAIRMAN STETKAR: Good.

MR. LAI: Shall we open the line?

CHAIRMAN STETKAR: Yes, we probably should -- thanks, John -- just because we are transitioning here.

While we are doing that, for the record, we have been joined by Bill Shack, and Dr. Michael Corradini has joined us. And we appreciate both of you gracing us with your presence.

(Noise on telephone line.)

Okay. We know the line is open.

(Laughter.)

Thank you for saying something. We actually have no indication here whether the line is open or closed.

So, what I would like to do is ask if anyone who has been listening in on the bridge line has any comments, in particular. If you have any questions, that is fine, but I am more interested in feedback and comments that anyone out there may have.

(Noise on telephone line.)

If someone is trying to speak, you are really garbled. It sounds like a female voice, if there are several out there.

MR. PETERS: This is Jing Xing on the other end.

CHAIRMAN STETKAR: Oh, that's is Jing.

MS. XING: Yes, this is me on the phone.

CHAIRMAN STETKAR: Okay, Jing. You are a lot louder now. So, we can hear you.

MS. XING: Okay. I didn't realize I wasn't on mute.

(Laughter.)

CHAIRMAN STETKAR: That is one of the dangers, isn't it?

MS. XING: Yes.

CHAIRMAN STETKAR: Okay. Do you have anything to add, if you have been listening in?

MS. XING: No, I don't have anything to add. I just want to thank the Committee and attendees. I appreciate all your comments.

CHAIRMAN STETKAR: Thank you very much.

Anybody else out there? Mary? I don't know. Anybody from EPRI?

(No response.)

If not, thank you very much. We will reclose the bridge line.

And again, thanks a lot for the presenters. I really appreciate your pulling a lot of information together into something that actually hung together in a presentation format pretty well. I was a little concerned. Thank you.

And with that, we will transition to the next topic regarding the empirical studies, both the international and U.S. empirical studies.

MS. LOIS: My name is Erasmia Lois. I am the Project Manager for the International HRA Empirical Study, which was your Office of Research.

And also, I am the Project Manager for the final presentation on the domestic study.

I think some other participants on the study, Dr. Dang from Paul Scherrer Institute and Jeff Julius from Scientech, may be on the line. If questions come up, which has occurred, we could probably unmute them to ask specific questions.

The objective of this briefing is to inform the ACRS on the overall results and lessons from the international study. Since the study is on the tail-end, actually, we are kind of documenting the bottom-line lessons in NUREG-2127. So, we would appreciate feedback to incorporate in this draft report.

What are the study's objectives? To test/evaluate methods in light of control room simulator data, with the intent of characterizing the methods, identifying strengths and weaknesses; develop a technical basis for improving the methods and method implementation, and, in general, to improve the HRA practices. And these activities are focusing on addressing ACRS and Commission direction over the years.

What is the motivation of the study? Mainly, it is the fact that human reliability and PRA has been used in risk-informed regulatory decisionmaking, and the recognition that differences in underlying frameworks and data, and quantification algorithms of HRA methods result in differences in HEP evaluations and in human error probabilities, and the recognition that HRA methods do not have the benefit of being tested with empirical data.

I would like to briefly remind the Committee that we have done work on improving HRA since early 2000. We did the good practices and, then, evaluated methods with respect to the good practices, supporting directly the PRA standards. And through these activities and in direction with the ACRS and the Commission, that we need to move forward with and embrace HRA model differences became strong, and the Commission directed, as we showed this morning, the staff to address HRA model differences.

So, the status supports largely what has been done for developing a hybrid method. The status: the study started in November of 2006. Phase 1 was to do the pilot, which was reviewed, the methodology for this study was reviewed and revised accordingly. And then, Phase 2 was to do what we call the Steam Generator Tube Scenarios Study. So, the pilot was published in NUREG/IA-0216, Volume 1; the SGTR in

Volume 2 of NUREG-0216. Phase 3, which is Loss-of-Feedwater Study, is the NUREG, Volume 3, will be published probably by February. And as I noted before, NUREG-2127 is in draft form and is going to be, hopefully, finished by March of 2013.

The list of authors here comprise mainly all of these people that contributed to the study through the years, and it comprises what we call the Evaluation Team, which was responsible for designing the experiment and executing the study throughout. It is different versions of NUREG-0216; volumes have different combinations of these names.

This is the list of methods that we evaluated, 13. You can see the large participation of organizations, and EPRI supported these methods, I am noting here, as well as so many other countries.

This chart quickly provides another view of the study. As I noted, it started as the experiments were performed at Halden facilities. Fourteen crews participated, which are actual crews. And this side of the figure presents the development of the empirical evidence that was used to evaluate the methods, and on the other side, participate in HRA analyses, performed the analyses, and the analyses were summarized by what we call the Assessment Group.

And then, the empirical evidence and the HRA analysis results were compared in order to understand how methods really performed in predicting crew performance, and then, to develop insights and improve the HRA methods.

I note that I am not going to spend a lot of time talking about how we developed the empirical evidence. However, it is a significant contribution of the study because it involved a lot of actually interdisciplinary teamwork on that. It involved to take the raw data from redo's, et cetera, et cetera, and develop an understanding of what has happened in each one of those analyses by each of the original crew, and then roll it up to truer level, so that we have a story of what happened during this analysis by each of the original crew, and then analyze that in a causal format, like HRA needs.

So, development of methodology to identify performance and empiricals for those methods that are based on PSFs or to develop what we call operational stories for those methods that are depending on narratives to identify what crews could do and when.

In terms of the HRA analysis, the teams did not have the benefit to go and visit the Halden facilities or the reference plant, but they were given an information package. And also, they had the benefits to ask questions, and all teams received the answers to the questions. So, on the one hand, they did not do what typically has been done, to visit the plant, but, then, all teams had all the information shared. So, we had the opportunity to see how actually methods used this information, the same information.

I think that is it.

MEMBER CORRADINI: Can I just make sure I understand. So, you took one of these crews and you ran them through these scenarios and observed them?

MS. LOIS: Uh-hum.

MEMBER CORRADINI: And then, essentially, analyzed their response to the scenario?

MS. LOIS: Yes.

MEMBER CORRADINI: And then, compared it to what the HRA analysis said might have, should have, might have occurred?

MR. FORESTER: That is correct.

MS. LOIS: And then, the HRA, we examined the HRA analysis and were able to really predict whose would fail or succeed and why.

MEMBER CORRADINI: Okay.

MS. LOIS: And these narratives that I

talked about on the empirical evidence development, we needed to analyze true behavior --

MEMBER CORRADINI: Right.

MS. LOIS: -- from an HRA perspective, which is a causal analysis.

MEMBER CORRADINI: And you say there were 14 crews?

MS. LOIS: Fourteen crews.

MEMBER CORRADINI: And how were they picked?

MS. LOIS: It was volunteers to have --

MEMBER CORRADINI: So, volunteers --

MS. LOIS: Yes.

MEMBER CORRADINI: -- just showed up, and you took whoever they chose to be the crew?

MR. FORESTER: I think that is pretty much all their crew. Well, yes. Yes.

MS. LOIS: All crews, all crews from that residence plant.

MEMBER CORRADINI: Oh, excuse me. I'm sorry. So, it is a set of crews from a plant? I misunderstood. I thought it was like 14 different plants. No?

MR. FORESTER: No, no. One plant, all the crews from that plant or most, roughly all crews.

MEMBER CORRADINI: Okay. Are we allowed to know who that is?

MR. FORESTER: It is a European plant. I don't know whether we can name the plant or not.

MEMBER CORRADINI: That is fine. And then, essentially, it is a PWR because you were identifying steam generator?

MS. LOIS: Yes.

MEMBER CORRADINI: So, it is a PWR?

MS. LOIS: A PWR, yes.

CHAIRMAN STETKAR: There were 14 crews, a plant?

MEMBER CORRADINI: Wow.

CHAIRMAN STETKAR: That is a strange plant.

MEMBER REMPE: Was it mix and match. And so, it was like maybe --

MS. LOIS: Two sides, right?

MR. FORESTER: Yes, I am not sure.

CHAIRMAN STETKAR: Okay. I was going to say a typical plant will have six or seven operators.

MR. FORESTER: It wasn't five or six people per crew. I think there was only three.

MEMBER CORRADINI: They were reduced crews?

MR. FORESTER: Reduced crews, up to three, I believe.

CHAIRMAN STETKAR: You mean split crews or?

MEMBER BLEY: Three people in the control room for each event.

CHAIRMAN STETKAR: Yes, but, I mean, it is strange that a single unit, for example, would have that many groups of even three operators. You typically have six or seven crews for a typical rotation.

MEMBER REMPE: Yes, if you had one operator be part of two crews, because it is different shifts?

MS. LOIS: No.

CHAIRMAN STETKAR: No, you don't want to do that.

MEMBER CORRADINI: Maybe we are offtrack. But, to me, I know you told me you are only going to tell us about the left branch, but I am very intrigued about the right branch.

So, 14 crews, one plant, PWR, no overlapping of people? So, John wasn't on two crews?

MS. LOIS: No.

MEMBER CORRADINI: And Mary wasn't on two crews?

MS. LOIS: It does happen that this one has 14 crews.

MEMBER CORRADINI: Okay. Do you know -- they must have, I am assuming, they must have an internal rating of how these crews behave at their plant in responding to unusual events. Do you know how these people grade-out when they came in? In other words --

MEMBER BLEY: No.

MEMBER CORRADINI: No? And that wouldn't help or that biases the data?

MEMBER BLEY: We didn't know. That wasn't something they brought to the project.

MEMBER CORRADINI: All right.

MS. LOIS: No. But, on the other hand, when the experiment finished, they took advantage of the lessons learned and they have done some changes in their procedures and training and even plant modification. And it is one of the plants that is using reporting possibilities frequently to improve their processes.

MEMBER BLEY: One thing you kind of flew by when Erasmia said it, I think, was a big part of the evaluation was what the crews did and why. I mean, that is all she said, was "and why," but a lot of their analysis goes through all the qualitative things the people doing the analysis said would affect the operators, and what actually affected the operators, and what the operators affected.

MEMBER CORRADINI: But I remember she said that, but does that mean you interviewed the operators after the fact and said, "Why did you" --

MEMBER BLEY: Halden did.

MEMBER CORRADINI: Halden did? Okay.

MEMBER BLEY: That is what they do. That is their business.

MEMBER CORRADINI: Okay.

MS. LOIS: And another, since you have brought us back to this picture, another note I make, we had workshops after, for every phase of this study, with the HRA teams. So, the method evaluators presented their evaluation to the HRA teams, and the Halden staff presented the crew performance, and there were very thorough discussions, and we had the opportunity to do feedback from the HRA --

MEMBER CORRADINI: So, I have a different question now that is more even off-topic. So, does the aviation industry do the same sort of thing about HRA analysis for their flight crews?

MS. LOIS: John could answer that.

MR. FORESTER: Yes, I don't know whether they use HRA in the aviation industry.

MEMBER CORRADINI: Really?

MR. FORESTER: Not very much, if any. I would say NASA has used HRA, but not the commercial industry.

MEMBER CORRADINI: When I was on the INPO Accreditation Board for Training, we had two representatives, one from Airbus and one from Boeing. They claimed they do this. So, I am curious --

CHAIRMAN STETKAR: Do they do HRA or human factors engineering, which is a completely different discipline?

MEMBER CORRADINI: Well, I am too pedestrian to know the difference.

CHAIRMAN STETKAR: I am sure they do human factors engineering.

MEMBER BLEY: They certainly do simulator training and observations.

MEMBER ARMIJO: But National Transportation Safety Board does all these analyses of accidents --

MEMBER BLEY: Right.

MEMBER ARMIJO: -- many of which are

attributed to pilot error.

CHAIRMAN STETKAR: Yes, but that is forensics.

MEMBER CORRADINI: That is forensics. That is after the fact.

MEMBER ARMIJO: That is after the fact, but --

MEMBER CORRADINI: I understand the training part of it.

MEMBER ARMIJO: -- but the HRA analysis is part of that.

CHAIRMAN STETKAR: HRA is human reliability. It is not root cause of a specific event.

MEMBER ARMIJO: If it is pilot error that caused it, the pilot reliability wasn't very good.

CHAIRMAN STETKAR: Well, this slide -- oops.

MEMBER CORRADINI: Well, I will stop now because I am way offbase. But where I was going with the question was, is this approach indicative of other approaches of other crews that you are trying to understand the reliability and how they would interact in off-normal events? And the only one that pops into my head, because at least -- and again, I don't really know, John, if it was human factors or what, but I remember distinctly the Boeing and the Airbus people were saying that they sent their crews through simulators.

CHAIRMAN STETKAR: Absolutely. Absolutely.

MEMBER CORRADINI: And they, essentially, try to see how they perform in off-normal events, which seems similar to this. That is why I was curious if they have done studies like that.

MEMBER BLEY: There were two parts to this. You got it. But the one part was what Halden has traditionally done, bring operators in, run drills on them, analyze what they do, take movies, the whole thing, track every second.

MEMBER CORRADINI: And brief them.

MEMBER BLEY: And brief them and work with them.

The other half of this was a different group of people did HRA analysis ahead of time, and then, they compared those analyses to what the crews did.

MEMBER CORRADINI: Okay. And your point is the lefthand side --

CHAIRMAN STETKAR: Go to the next slide,

Erasmia.

MEMBER CORRADINI: -- may not exist in other worlds? Is that your point?

MEMBER BLEY: Or isn't acknowledged, and I can say that in some industries -- I won't say which ones -- if they do probabilistic risk assessment and associated human reliability analysis, sometimes they do that in like a Skunk Works and don't tell anybody they are doing it because they don't want it in the legal process if they have accidents later.

MEMBER CORRADINI: Okay. Fine.

MEMBER REMPE: When they do this, what kind of signals do they give them? Do they give them erroneous signals? Do they give them data that are incorrect or give them partial data? Because a lot of times in real accidents some of the data wasn't there for the operators. I mean, how far do they go?

MS. LOIS: So, actually, it was I did not do that. For the steam generator tube rupture scenario, we had two variants. One was what we call the easy scenario, the well-practiced scenario, and a variant which was exactly what you said, some behind-the-scenes false indications which made the scenario more difficult, but within their training abilities, if you wish, or experience. It was not as far out --

the trainers were within the scenario.

MEMBER REMPE: Okay.

MS. LOIS: That allowed to see the variability, as I am going to go into later, variability of crew behavior in different scenarios and under different --

MEMBER BLEY: They intentionally set up some scenarios with what they called "masking," trying to hide the scenario through other things going on. That included instruments; it included different things that could go together, but the one would hide the other. They did that intentionally to see if the analysis could deal with that situation and to see how the operators deal with that. And it is a thing that they are starting to do more and more in power plants because they haven't been doing much of that in the past. And the real events that get you in trouble have that --

MEMBER REMPE: Right, and so the training has to be emphasizing having an inquiring mind.

MEMBER BLEY: I mean, we would like to do that, but it is new.

MR. FORESTER: And to the extent possible, they tried to do these so that there are legitimate ways it could actually happen, rather than just a bunch of totally random phases. You are trying to create situations where certain indications might be hidden, so it is more plausible.

MS. LOIS: And actually, it took a while to convince Halden to do these kinds of analysis because the typical human engineering analysis is not so much focused on creating and analyzing failures. It is more focused on, quote/unquote, "success". So, Dennis and John were part of what I call cultural change in the Halden philosophy on how to conduct experiments for human reliability.

So, the method assessment was based on comparing the HRA results to empirical evidence, which is what the crew, actually, from both qualitative and quantitative perspectives. That is the main focus of the study, but also we looked at the traceability of quantitative and qualitative analysis, the usefulness of the qualitative and quantitative analysis with respect to error reduction: what we learn in them; can we improve human performance at the plant after we do this analysis? And also, another major aspect was analyzing the adequacy of the guidance given by the method for both the qualitative and quantitative analysis.

With respect to qualitative analysis,

since we have methods that are used in performance-shaping factors mainly as a way to evaluate human performance, so we looked at PSF, what we call PSF assessments; namely, have the methods identified performance variable drivers in terms of PSFs and how well those drivers were evaluated. And for those methods that are using narratives as a way of predicting performance, we compared how well the method applications predicted the ways crews could fail and the operational situations that would contribute to these failure paths.

In terms of quantitative analysis of predictive power and evaluation, we looked at the potential optimism for the most difficult HFEs, the consistency of the ranking of the HFEs compared on the basis of HFE estimates. So, in actuality, we ranked the HFE difficulty on the basis of empirical evidence, and then, we looked if the human error probabilities produced reflected that HFE ranking. And even if you do have the HFE ranked appropriately, another aspect is, do they differentiate enough, so that the level of difficulty is well-recognized

And also, we looked at the HFE relative to the confidence and uncertainty bounds, although that is more weak criterion, given that the data were limited.

MR. FORESTER: Even with 14 crews, it is still not a lot of data.

MS. LOIS: Okay. So, designing easy and complex scenarios had the effect of observing crew variability, and that crew variability gave us the opportunity to go beyond just failure counting and examine a broader spectrum of performance issues. For example, how crews go about to respond to an event and how to implement procedures. If crews followed a different path, what was the reason for this variability? And allowed to develop an understanding of what issues are building when they are performing a scenario and the underlying reasons.

So, developing the operational descriptions allowed compilations of the critical evidence with method predictions. And actually, site descriptions allowed to explain how tasks were performed, why performed in this way, and the consequences, and exemplified how HRA concepts, such as diagnosis, can be observed in actual crew performance.

And I am going to note that both aspects of the study are very important because now, when we are doing human reliability in general, we talk about diagnosis and develop PSFs, but to have empirical evidence for those kinds of concepts, it was the staff brought us very valuable information, and pointed out that HRA practices in which cognitive demands are not examined well have the impact to miss important performance issues.

So, yes, it was crew variability was observed, and it was observed mainly in what the study brought up significantly, that the crews, when they do respond to an event, even for the easy tasks, the well-trained scenarios, after the initial diagnosis, they perform what we call cognitive tasks continually. They are continually thinking what is going on and looking at the parameters, and they make decisions. They continue to make decisions.

That aspect is frequently not-well-encompassed in human reliability. Analysts do not seem to think in terms of how crews would react and the different ways that crews could follow in responding in an event. But, at the same time, by the same token, the methods themselves do not help analysts to think in these terms.

MR. FORESTER: And there are examples where some methods will say, you know, well, if you have already entered the right procedure, you have diagnosed what the event is, then you are just following procedures; now it is just an execution sort of issue. So, the diagnosis can be ignored under some circumstances.

What you see when people do that, they miss a lot of important information. They miss areas where the operators might make mistakes.

MS. LOIS: Okay. So, the PSF analysis, regularly, we focused on identifying PSF presence in crew performance. And that helped in clarifying some very basic HRA concepts.

For PSF, for example, we created non-overlapping definitions, and then we determined how these PSFs, actually, how you can say it is present, that crew behavior was driven by a particular performance-shaping practice.

And that whole aspect of initial diagnosis versus cognition and the cognitive tasks that go on, it was kind of a lightbulb coming on which was abundantly available/evident in the HRA analysis. People most frequently will not think about it. Some methods, such as ATHEANA or MERMOS are continually thinking about the various paths that crews could take and why, but other methods do not. I don't know. Shall I say ASEP or SPAR-H?

So, in terms of quantification, what are the overall findings? So, we have strong evidence that HEPs are frequently optimistic, even for difficult HEPs. And people would recognize that this human action is difficult. And yet, the number that would come up is very optimistic.

The ranking of HEPs. HEPs do not reflect the relative difficulty levels for the HFEs observed. And that is a -- yes?

MEMBER CORRADINI: So, for your first bullet, you are underpredicting failure rates?

MS. LOIS: Uh-hum.

MEMBER CORRADINI: What you saw and what was predicted always was saying that humans would do a better job at diagnosing and reacting?

MS. LOIS: Not always, but --

MR. FORESTER: Not always, but --

MEMBER CORRADINI: But there was a noticeable trend? Okay.

MS. LOIS: Yes.

People are not thinking in terms of do what we call the reasonableness check afterwards and say, Okay, do these numbers line up in terms of the level of difficulty?", which is one of the basic PRA good practices.

CHAIRMAN STETKAR: Yes. That second bullet is to me, actually, a bit more disturbing than the first one because a lot of people have said, well, we might not have a good understanding of human error probabilities in an absolute sense, but the methods at least should give us confidence that the relative ranking of actions within a particular risk assessment is appropriate, that the highest HEPs apply where you would expect them and the lowest HEPs apply where you would expect them. But that second bullet really calls into question even that degree of confidence.

MS. LOIS: Yes. And, of course, the third bullet, which is the capability of the analysis to differentiate, to discriminate among the human actions in terms of difficulty, we would see some analysis that was difficult human actions and not difficult, and the level of differentiation was minimal. So, that thought process also did not seem to be, you know, pervasive in the analysis.

MR. FORESTER: I would say if you look in the report, not in the final report but for the different scenarios, you know, if you look at the rankings, they missed here and there, but it wasn't horrible. I mean, often, they would get the right kinds of ranking.

CHAIRMAN STETKAR: The right relative --

MR. FORESTER: Yes, the right relative ranking. And maybe they would have one that was like -- they are going along pretty good, and then, they got this one way up here, when it should have been down there. So, again, it wasn't always horrible.

MEMBER BLEY: The "should have been" is another issue. I mean, we have limited data.

MR. FORESTER: Sure.

MEMBER BLEY: And it should have been based on limited data and the assessments of the group of individuals who were evaluating all of this.

CHAIRMAN STETKAR: I was just kind of hanging up on the term "in many cases," which implies perhaps more pervasive experience than what I am hearing now.

MEMBER BLEY: They are going to show you some trends.

MS. LOIS: Yes.

MR. FORESTER: You really have to look at the individual method to see --

MS. LOIS: John is always a good guy. He is not going to dig anyone. But, in my estimate, it is an issue. If we look at my list -- this is a decision-making point of view -- the ranking is an important aspect and the differentiation among the HEPs is also important. It should be good practices, you know, and the PRA standards. These are some really basic things. And my point of view is, before we get to have ideas, if we can have some of these rankings appropriately. Then, we are in a good position.

MR. FORESTER: You are right. It is bad enough that we definitely need to improve the ability to do that.

MS. LOIS: The fourth bullet is also very interesting, which is no method was consistently producing high or low HEPs. So, this idea that one method is just a screening tool and has been used frequently in risk-informed regulatory decisions, I think we have strong evidence that it isn't the case.

CHAIRMAN STETKAR: Point 1 is not universally-conservative?

(Laughter.)

MS. LOIS: No one is universally-conservative. And I am talking about SPAR-H, which has been used, and I heard from the regions, "Oh, this is a screening tool, and I just want to use it and decide whether or not I am going to go deeper into analysis." I mean, you can get 10 to the minus 6 very easily, very, very easily.

And then, the fifth bullet, comparison against confidence and bounds, it was limited. It is a limited evaluation, and I think it holds more for the difficult HEPs, for the high-value HEPs.

So, here is the picture --

CHAIRMAN STETKAR: The uncertainty bounds -- go back to the previous slide. I want to understand that. It says, "The uncertainty bounds of the reference data were broad for the easier HFEs and relatively narrow for the more difficult HFEs,..." That says you saw, from the actual experience, you saw larger, I will call it, variability.

MR. FORESTER: Well, I think it is just if you have 7 out of 14 errors, then you have some information. If you have zero out of 14, then you --

CHAIRMAN STETKAR: Oh, okay. Okay. Got it. Thanks. That helps.

MS. LOIS: Yes. The easy, everybody has --

CHAIRMAN STETKAR: Yes. Zero out of 14, I understand it in that context. I was thinking of it more in terms of measurability variability, if you will.

MS. LOIS: So, this shows variability. It is the human actions from more difficult to most easy ones. So, you can see in the SGTR the variability for the most difficult human actions, the same thing for the loss-of-feedwater.

Also, the variability method-to-method is large.

MEMBER CORRADINI: I am trying to understand. So, the circles are the analytical methodologies, and where is the data?

MR. FORESTER: Well, these are the predicted values from the HEPs.

MEMBER CORRADINI: Oh, independent of the data?

MR. FORESTER: Independent of the data.

MEMBER CORRADINI: Excuse me. I'm sorry. Excuse me.

MR. FORESTER: Team A did that one; B did that one, and so forth, for each of the different HEPs.

MS. LOIS: But the data are the human failure events --

MR. FORESTER: The single event.

MS. LOIS: -- that are ranked from the most difficult to easiest.

MEMBER CORRADINI: Yes, that part I got.

MS. LOIS: Yes.

MEMBER CORRADINI: Yes. Things that are close to 1, you fail more often. What I am trying to understand is that, I am guessing, going back to your bullets, one of the conclusions of your bullets can't be discerned from this. This just shows how the various methodologies performed in comparison to each other?

MR. FORESTER: Correct, except that this was the hardest action. I think all crews failed. You see there is quite a range of predictions about what those crews were going to do.

MEMBER CORRADINI: Well, how do I know that all crews failed on this action?

MR. FORESTER: You don't know that precisely.

(Laughter.)

MEMBER CORRADINI: Okay.

MR. FORESTER: But what you do know --

MEMBER BLEY: From the tightness of the bounds --

CHAIRMAN STETKAR: You notice between .7 and 1, for example, on the blue.

MR. FORESTER: From the hardest to the easiest.

MEMBER CORRADINI: So, that part I got.

MR. FORESTER: Okay.

MS. LOIS: But we do know.

MR. FORESTER: We have the data, but it is not shown here.

MEMBER CORRADINI: That is all I was trying to get at.

MR. FORESTER: Correct.

MEMBER CORRADINI: In other words, I was looking for like a yellow dot that says it is at 1. And so, they just always failed. And there is reality for the 14 crews and everybody else spread through it. But, okay, I am with you now.

And what do the red circles mean?

MS. LOIS: These are the outliers.

MEMBER CORRADINI: Oh, the outliers?

MS. LOIS: And I will explain why.

MEMBER CORRADINI: Okay.

MS. LOIS: And in terms of actually this picture shows how two individual methods did with respect to the ranking that we were talking before. Like the thing that was the most difficult human action to the easiest one, you can see how the result changed. And also, the green line shows the relatively-small differentiation among the numbers produced.

MEMBER CORRADINI: This is the first time this has been done, right?

MS. LOIS: Huh?

MEMBER CORRADINI: This is the first time this has been done?

MS. LOIS: This is the first time, but, then, Harry is going to talk the domestic, what we call the domestic study. So, it is done once more.

MEMBER CORRADINI: So, do I take heart in this?

MR. FORESTER: Well, again, if you look at that green line --

MEMBER CORRADINI: Because I am not taking heart.

MR. FORESTER: No, you are not very happy with that green line. But, again, not all the methods did that. The orange one, even there, I mean, they did nail the two most difficult ones. They were keyed into that. And then, they came down as they should have been, and then, there is some bouncing around.

MEMBER BLEY: Tell us again what the orange and green lines are, John.

MR. FORESTER: Pardon me?

MEMBER BLEY: Tell us again what the

orange and green lines are.

MR. FORESTER: They are just two different method applications. So, this might be -- I don't know what they were. This might be ASEP; this might be --

MEMBER BLEY: Oh, okay.

MR. FORESTER: But one of the differences --

MEMBER BLEY: Well, one of the methods got the same answer for everything.

MR. FORESTER: That is right.

MEMBER BLEY: That is what the green line is telling us.

MR. FORESTER: And so, that is not encouraging because there is no difference. They can't tell the easy ones from the hard ones.

MEMBER BLEY: But, in fairness, something that wasn't talked about here, I mean, this was done to see how methods do, how people do, all that sort of thing. But this was done before the other work you have heard about this morning was done. So, the results that they learned here were used to guide the development of the methodology that was discussed this morning, hopefully, to avoid some of the reasons for the funny spots you see up here.

MR. FORESTER: That is exactly right.

MEMBER BLEY: And I still want to say, in all of these studies, we had limited data. In fact, in one of them, while all of the crews failed, in the pre-run the crew that tested it out to see if the scenario was okay didn't fail. You know, so if you added that into the story, you would have gotten a bigger stretch at the high end of the worst one.

In general, they used this go guide the development of the new methodology, right?

MR. FORESTER: Yes. This is really hard research to do, and it is not going to be perfect. But, in spite of that, there are clear lessons learned. To me at least, in my mind, there are not going to be any arguments about the conclusions. We definitely learned a lot of information about how to improve the methodologies and what needs to be done, in spite of the limitations.

MS. LOIS: So, although we thought at the beginning that quantification would not inform the study as much, we found out retrospectively that we hit some really good issues. And therefore, it is really an important aspect of the study, although originally we thought that we would focus more on the qualitative analysis part of it.

The findings from the qualitative analysis: handling of crew cognition, it is an important aspect, but it is one that has actually not been well-addressed in HRA methods in general. And as we heard this morning, probably it is one of the issues that we don't know how to handle.

But keeping in mind that people are continually thinking and making decisions as they go along the scenario, it is very important. Now I ATHEANA analysis and CBDT, et cetera, because they are causal analyses, they do a much better job on that.

And the point I make here, it is evident that HRA does have the capability to predict crew performance. We actually saw ATHEANA and MERMOS and CESA actually predicted some actual crew failures. They said crews are going to go, looking at these procedures, are going to go this way and that way, and we observed that. So, in a way, it verifies that human reliability has the capability to predict crew performance and identify issues as you do the analysis.

MEMBER CORRADINI: So, you are happier qualitatively than you are quantitatively? That is what I heard you just say.

MR. FORESTER: That is correct because --

CHAIRMAN STETKAR: Qualitatively with some methods.

MEMBER CORRADINI: I understand. I understand.

MR. FORESTER: But even some of those methods, yes, there are issues. They have a much better qualitative analysis, but that doesn't always translate into better quantitative results.

CHAIRMAN STETKAR: I think the key is that most of the PRAs that are being done today don't use any of those methods.

MR. FORESTER: Yes, that is the second piece.

CHAIRMAN STETKAR: That is the second part of the bad news.

MS. LOIS: And on the PSF-based methods, we saw that these methods depend heavily on expert judgment to decide which PSF and the strength of the PSF. And the, quote/unquote, "incorrect judgments" contribute both to over- and underestimation of the HEPs.

And, of course, another issue is that the range of PSFs supported by the method, it is limited, and we saw that the experts can compensate frequently by stretching one particular PSF to account for complexity, for example, the complexity PSF. But if you stick with the method, and if you follow it verbatim, you are going to miss important PSFs.

So, I just want to conclude from the qualitative analysis perspective the issues that we identified in the quantitative analysis do stem from lack of good structural and qualitative analysis and comprehensive structure.

Regarding traceability, we examined it from both the qualitative and quantitative perspectives. And it seems that the PSF, those methods are good with respect to traceability on the quantification aspect. Because you can see, once we have made the decision, what PSFs are appropriate. Then, you can follow up and see what they did, how it came out with the numbers.

The context-based methods are good with qualitative traceability. I said before that we actually saw methods predicted actual crew performance, but are not as good with respect to quantification because it is not easily traceable, the way of translating those insights into HEPs, and reproducibility is an issue, too.

With respect to error reduction, most methods do not have that capability, especially with

PSF-based methods. The newer, narrative-based methods have much better capability because they do describe how elements in the scenario can be evolved and are easily comprehended by operations people. So, actually, ATHEANA is built on a narrative-based analysis, right? That is what is the focus, to identify what we are going to build this out of there and address them.

CHAIRMAN STETKAR: I am curious, regardless of fidelity of the numbers, the second bullet, when you say "capability of error reduction," do you mean by that what I would characterize as a root-cause analysis for the contributor to the error, such that somebody in the plant will decide to write a better procedure or improve training or paint the handwheel red?

But I am curious why the traditional PSF methods are more deficient in that sense than the context-based methods. Certainly, as I said, I want to get away from the fidelity of your ability to predict a human error probability or characterize the qualitative aspects of a scenario.

MR. FORESTER: I guess I would say that, if you had a PSF method that treated the correct range of PSFs, and had an underlying definition for those

PSFs that was usable, then, yes, in principle, it would be --

CHAIRMAN STETKAR: Okay. So, this is more in terms of the scope, the breadth of the number of PSFs that are considered --

MR. FORESTER: I think, yes.

CHAIRMAN STETKAR: -- providing enough fidelity to identify causes? I mean, in principle, if you had 25 PSFs, you could say, okay, I am failing because of PSFs No. 1, 17, and 83, or whatever.

MR. FORESTER: I think that is true, yes. Maybe the operational story maybe gives you a little bit clearer picture in some ways, but not necessarily.

CHAIRMAN STETKAR: Okay.

MR. FORESTER: I agree.

CHAIRMAN STETKAR: That helps. That helps.

MEMBER SHACK: Well, I thought there was some discussion that you felt that the range of the PSF itself had to be larger?

MR. FORESTER: Yes, yes.

MEMBER SHACK: Okay.

MR. FORESTER: Yes, because of you look at like ASEP, and there are just a few PSFs that really get considered. They are treated in a direct way.

So, you are going to get limited information out of it. If it is a situation that happens to address, those happen to address the key aspects of that event, then you may get something, but, otherwise, there is just not enough coverage to really help you identify all the potential issues and, out of that, then someone being able to decide how to fix it. So, just limited information.

CHAIRMAN STETKAR: I hate to ask this, since you were at the table in the previous discussion, but since you were, I will.

(Laughter.)

MR. FORESTER: Okay.

CHAIRMAN STETKAR: The number of PSFs that are being considered, PIFs, PSFs, whatever you call them, branch points, that are being considered in IDHEAS is pretty small, much smaller, if I just use a body count, compared to the number of line items, for example, in the cognitive basis document. Is there a danger that IDHEAS may suffer from that same characteristic?

MR. FORESTER: I guess I hadn't counted them.

CHAIRMAN STETKAR: I mean, typically, in a CBT, in a typical decision tree, you might have anywhere from three to five, you know. Now, those are, in principle, agglomerations of other things, but they are still --

MR. FORESTER: Right, they are, but, then, you are also considering across all the CFMs, though.

CHAIRMAN STETKAR: Yes, but there aren't all that -- there is workload; there is HSI. I mean, there is a lot of commonality there.

MR. FORESTER: I mean bias kinds of issues and the training and setup.

CHAIRMAN STETKAR: Okay.

MR. FORESTER: So, in some ways, maybe the PSFs might not be any broader, but we use the term "PIS" because you look at plant conditions. They are a part of that. So, there is a bigger part of the context I believe getting captured in that, in the decision tree then.

CHAIRMAN STETKAR: Okay.

MR. FORESTER: But you raise a fair point. I mean, we probably should go through the process of confirming that. I mean, we have gone through the process of these are what we think we need to do to address this CFM, but it is fair returning and taking another look to see if there is anything obvious that we are not addressing.

MS. LOIS: Another aspect of it for PSF methods is how these PSFs are used. For example, SPAR-H has a good number of PSFs. However, it relies so much on analyst judgment to come up with what is the PSF. And we saw complexity in HMI being used in actually not an intended way from the method itself.

And then, the strength of the PSF. And so, it is --

CHAIRMAN STETKAR: Yes, you are right, there are other problems with PSFs.

MS. LOIS: Yes, yes.

CHAIRMAN STETKAR: Oh, yes. Yes. I am not advocating PSFs. I am just trying to understand a little bit some of the insights, you know, in the context of this slide, in particular.

MS. LOIS: A lot of insights in the report about improving and, of course, methods. Improve guidance for identifying a comprehensive set of factors. Even if the method passed the factors, if really the guidance on identifying how these factors do show up, it is important. And then, how do you judge the strength of the factor -- I'm sorry -- yes, and how to judge the strength of the factor. And the other thing is, how do you link back your qualitative insights into the quantitative insights?

I do want to make a note here. When we started doing IDHEAS, one thing came out from this study is, in actuality, methods are suffering because they don't have a comprehensive, underlying framework. And that is what tried in IDHEAS with the cognitive report. We went back -- the reason for doing the cognitive research was because we recognized that methods have bits and pieces, and all methods are good, but they have partial truth. And we try now to have more comprehensive technical basis, so that we are not pushed to miss PSFs or judge the PSFs inappropriately.

So, here we say we can improve the guidance, but the question remains, shall we put the effort here? Shall we put effort into individual methods?

Improved guidance for HFE decomposition, for example, is important for considering this issue that we saw, and decide that, once you have the initial diagnosis, you can run with it, just to use the response factor of the method.

So, all of these issues are important, and we saw that by improving the guidance for individual methods, probably some of these issues can be eliminated.

CHAIRMAN STETKAR: That is, quite honestly, one of the things I would be really interested to see, how the IDHEAS pilot applications work. Because although we discussed earlier this morning -- I don't know if you were in the back or not -- the fact that, for a particular human response scenario, let's say there are -- I will use the example I used this morning -- 36 crew failure modes, each of which has associated performance-influencing factors. It, in principle, doesn't make any difference whether I combine those into three nodes in a Crew Response Tree, each with 12 contributions, or whether I develop 36 different nodes, each with one. I should, in principle, have the same assessment.

There is still, I believe, in IDHEAS quite a bit of flexibility left up to the analyst in terms of that level of discrimination.

MR. FORESTER: Yes, I am not so sure any way you do that is going to necessarily be the same --

CHAIRMAN STETKAR: Yes, but that is one of the areas --

MR. FORESTER: I agree.

CHAIRMAN STETKAR: -- that it may be prone to some of this analyst-to-analyst judgment variability that we are trying to address.

MR. FORESTER: I think, because you are looking at the different phases of information processing, that that will help things.

CHAIRMAN STETKAR: It is more structured. It is, in principle, more structured.

MR. FORESTER: That is the hope.

CHAIRMAN STETKAR: In practice, it is not clear yet.

MR. FORESTER: Fair enough, yes.

MS. LOIS: But at least it is traceable, like in the PRA, if you show why you used these failure modes for agreement, you don't agree with that, but you know why. So, I don't think you can reduce variability, but at least you make it objective enough so that people can understand where you are coming from. So, that is two objectives.

CHAIRMAN STETKAR: Yes.

MS. LOIS: So, you can improve guidance for performing this reasonableness check. I think that is important and probably should be implemented before we finish the IDHEAS development.

And perform a thorough assessment of potential failure mechanisms, it goes hand-in-hand in connection with the possible operational context.

And again, I agree that evidence of

qualitative analysis, good qualitative analysis, leads to more comprehensive basis as an input to HFE quantification. However, given the limitations of the methods, it is questionable if we should focus on improving individual methods, as I noted before.

So, in terms of conclusions, the international study performed a comparison of HRA outcomes to crew performance, proved to be very valuable; documented crew performance from an HRA perspective, which is a causal analysis. And actually, it needed efforts, again, what it was envisioned at the beginning, but it was a worthwhile effort; allowed an understanding of how methods are applied, and identified methods' strengths and weaknesses; allowed a comparative analysis of methods to some extent.

And the major conclusion is that no method meets all desirable characteristics. And therefore, probably a hybrid could be the solution. And more studies to further improve the robustness of human reliability. And, of course, this analysis could be used as a way of actually testing IDHEAS, because now a lot of work has been done. When it comes to test the variability or the actual capability of IDHEAS to identify the various paths, I think it will be very useful.

CHAIRMAN STETKAR: Any comments, questions for Erasmia?

(No response.)

If not, Harry, before you launch into your presentation, it is probably prudent for us to take a break now rather than letting you get a third or halfway through.

So, let's do that. Let's recess until 2:45, and we will come back and hear about the U.S. study.

MS. LOIS: Thank you very much.

CHAIRMAN STETKAR: Thank you.

(Whereupon, the foregoing matter went off the record at 2:32 p.m. and went back on the record at 2:47 p.m.)

CHAIRMAN STETKAR: We are back in session.

Harry, it is yours.

MR. LIAO: Thank you.

Well, good afternoon, everybody.

My name is Harry Liao. I a research staff member with Sandia National Laboratories. In the next 30 minutes, I am going to take discuss some of the U.S. HRA study.

So, basically, this is to inform the

Subcommittee on the findings from the U.S. HRA Empirical Study and, also, to obtain feedback from the Committee to incorporate in the documentation of the study.

I just want to give you some background about this study. The U.S. study is a followup study on the international study. It was expected to address the limitations in the international study and some concerns over the results from the international study.

Like Erasmia said in her presentation, the international study used Halden facilities and European crews. So, the Commission was interested in ensuring the applicability of Halden human performance studies in the U.S. applications using U.S. crews. And also, in the U.S. study, the HRA analysts did not have the opportunity to visit the plant, interview staff, and observe simulation runs.

In the international study, only one method was used by multiple teams. So, it was a very difficult to separate method effects from analyst effects.

In the SRM, in February 2009, the NRC staff was asked to pursue testing U.S. operating crews' performance in a variety of situations and keep the Commission informed on benchmarking and HRA database projects. The U.S. study was one of those projects.

Although this was a U.S. study, it was largely supported by an international collaboration: Paul Scherrer Institute, Halden Reactor Project, EPRI, they supported this project and made important contributions without funding from the NRC.

The HRA analysts were comprised of HRA experts from National Laboratories, the NRC, SAIC, EPRI, and the University of Mexico, and the Czech Republic.

Here, we also want to take the opportunity to thank the participating plant and the many other individuals who made this study possible.

The U.S. study achieved the objectives of improving insights from the international study and obtaining some insights on the limitations and concerns over the international study.

This table lists the milestones of the U.S. study. This study, it started in 2009, the design of the study, the scenario, and the human failure events were developed in 2009.

The HRA teams were formed and they completed their predictions in 2010.

The HRA, the crew data was collected and analyzed in March of 2011. Following that, the HRA predictions were compared to the crew data and, also, the teams using the same methods were also compared, and the results were discussed in the workshop in June 2011.

In January 2012, the initial study results were submitted to a PSAM 11 conference.

We are going to complete a NUREG/CR report to document the study in 2013.

This graph, basically, illustrates the methodology of the U.S. study, which is very similar to that used in the international study. At the beginning of the study, three scenarios were developed.

The first scenario is loss-of-feedwater followed by steam generator tube rupture. In the first scenario, three HFEs were defined.

The second scenario is loss of component cooling water and the reactor cooling pump seal water, in which one HFE was defined.

And the third scenario was a standard steam generator tube rupture.

Nine HRA teams used the four HRA methods to analyze the five HFEs defined in this study. Of the nine HRA teams, two used the ASEP, three used the HCR/ORE and the CBDT, two used the SPAR-H, two used ATHEANA. Of the three teams that used the HCR and the CBDT, two of them used the EPRI HRA Calculator software; the other one did not.

The three scenarios were similar with four licensed crews from a participating U.S. nuclear power plant on their full-scope training simulator.

MEMBER CORRADINI: Just something that I understand: the difference between Halden and a full-scope simulator is nothing? It is just a different simulator? Essentially, is that really what it comes down to?

MR. FORESTER: It is a digital simulator in a sense. So, it is not this analog controllers like we have here. So, there are big-screen displays and these computers.

MEMBER CORRADINI: Okay. But, then, maybe I am going back to your presentation, but that isn't what they see in the plants?

MR. FORESTER: That is correct. So, the crews were trained on using that interface.

MEMBER CORRADINI: Okay. All right. Sorry. Sorry to get behind.

MR. FORESTER: No, that is a fair issue.

I mean, they were trained and they felt like they were comfortable doing it, and they did fine.

MEMBER CORRADINI: They were trained?

MR. FORESTER: Yes. I mean, it is a question, would it have been exactly the same? We don't know.

MEMBER CORRADINI: Okay. Thank you.

MR. LIAO: So, after HRA predictions were completed and crew data were reflected and analyzed, the methods in this study were assessed in two parts. The first part is to compare the HRA predictions to crew data, crew performance data. The comparison started with the quantitative results, but the qualitative predictions are weighted more strongly in the comparison.

The second part of the method assessment is intra-method comparisons, meaning comparing the HRA teams using the same method. The comparison focused on differences in their qualitative predictions, in HFE ranking quantification results, on their analysis approaches and assumptions. Also, the potential contributors to their differences were identified.

Crew data were collected for four HFEs. This table lists the difficulty levels of those four HFEs. The difficulty levels were ranked by three human supervisors who participated in this study. The supervisors ranked those four HFEs in the same order. In addition, the supervisor ranking was consistent with the failure rates, observing the study.

For the most difficult HFE, 2A, four of the crews failed. For HFE 1C, three out of the four crews failed. For HFEs 1A and 3A, all of the crews succeeded.

The HEPs of each method are plotted in those four pictures. Actually, they are ordered by their difficulty levels on the horizontal axes. As we can see, compared to the international study, the HFE ranking was moderately improved for most of the methods. And for most of the HFEs, there was one order of magnitude or less of difference across the teams using the same method, especially considering this point; this is HFE 1C by a team that used 6DT or ANEP, or they misunderstood the scenario. So, this could be considered a mistake. Another of the reasons -- ASEP and HCR/ORE and ATHEANA produced more consistent results than SPAR-H.

And for HFE 2A, which is the most typical HFE, except ASEP, most of the teams underestimated for this HFE. This confirmed a finding from the international study that, for difficult HFEs, there is a tendency to underestimate the HEP.

For HFE 3A, all of the teams concluded that this was the easiest, but there was significant variability across the teams within the same method.

MEMBER BLEY: You know, there is one thing I don't think was described the first time or this time. Those bounds that you see on here were put together by some kind of Bayesian calculation that mixed two things. The evaluators put together an estimate of the ordered, which things are most likely, and so on. So, from their judgment, they built that. And then, from the actual data, some of which was zero failures, they fed into this Bayesian machine and out popped the bounds. So, it is just to tell you where they came from. We had a big fight there about one of them.

(Laughter.)

MS. LOIS: Admittedly, the bounds are weak.

MEMBER BLEY: Well, yes, and they are weak on one end when you have zero failures because they are somehow constrained, and they are weak on the other end because we only have four crews for the high number of failures. Well, we had four crews, all of which failed. If we threw in the fifth crew that was the test crew that succeeded, then we would have had four out of five, and that would look different at the beginning. Just a point of information.

But, still, I think all the main conclusions you gave are right on target.

MS. LOIS: A point well-taken.

MR. LIAO: First, this is just a summary of what I talked about of the quantitative findings.

Next I am going to talk about some example findings from intra-method comparisons. For ASEP, Team 1 followed the method very closely. For Team 1, it seems that their team did a very detailed review of procedure paths, based on their experience from the international study on how to do a good HRA. As a result, Team 1 obtained better estimates of the required time for post-attacks and those actions. And also, they identified more issues for HFE success.

And this difference is because ASEP has limited guidance for reviewing procedure paths and estimating time required for actions. In addition, ASEP focused on the evaluation of time reliability curve with a few PSF adjustments. So, this limits the method's ability to identify diagnosis difficulties.

Another observation about ASEP is about, even if diagnosis difficulties are identified, the method is poorly-equipped to quantify those difficulties. Although analysts may compensate based on their experience, it can lead to differences in their quantitative results.

Another difference between the two ASEP teams is that the two teams obtained different HEPs for post-diagnosis actions. One of the reasons for this is the method has limited guidance on what is good in the post-diagnosis analysis. So, different teams can make different decisions.

Another reason is that ASEP allows analysts to use either ASEP or THERP to quantify post-response actions. The choice of a quantification model can cost their ability in their HEPs.

CHAIRMAN STETKAR: Mostly, I think it is the first bullet. It really doesn't pay much attention to the post-diagnosis stuff.

MEMBER CORRADINI: Can I ask you a question --

MR. LIAO: Sure.

MEMBER CORRADINI: -- maybe that I don't understand? Can you go back two clicks, please, to the figure?

MR. LIAO: Sure.

MEMBER CORRADINI: So, how is the

calculated value less than the fifth percentile? How is it outside of the 5-to-95 bound?

MEMBER BLEY: I can speak to that because that one that is the furthest out is mine.

(Laughter.)

What I told you about the way the bounds were done, it was a Bayesian analysis, and I never quite was able to determine what the prior was, but I know it was affected by their ordering. And then, zero out of so many failures, and they came up with that bound. The event itself, I forget. I think they said it was very easy, but there was about many, many, many hours to respond to it. So, there was a difference in the argument.

MEMBER CORRADINI: Okay. Got it.

MS. LOIS: It is this one, right?

CHAIRMAN STETKAR: That might be Dennis, just for example.

MEMBER BLEY: That might be Dennis.

MS. LOIS: But we have --

CHAIRMAN STETKAR: And there is another one up there.

MEMBER BLEY: There is another one up there.

MR. FORESTER: Interestingly, on the

simplest action, the easiest action, we saw the greatest variability in that.

MEMBER CORRADINI: Well, that I read as the conclusion. Then, I noticed they were outside the 5-to-95, and I didn't understand that. Okay. Thanks.

MEMBER BLEY: But correct me. I am not sure if that Bayesian approach for laying out those bounds is really clearly explained in any of the reports.

MS. LOIS: Probably not, but probably we have Vinh Dang on the phone. If we unmute him, he will be able to explain to us.

(Laughter.)

MEMBER BLEY: Well, I have trouble following it. It would be nice to read it somewhere.

MR. FORESTER: There is discussion in the report.

MEMBER BLEY: Well, there is a discussion, I agree with that.

MR. FORESTER: Yes, I'm not sure whether it is enough, but, I mean, there is definitely a discussion. It is described, I believe.

MEMBER BLEY: Vinh's yelling, but we can't hear him.

(Laughter.)

CHAIRMAN STETKAR: John, go get it open, just in case Vinh is actually on there.

MS. LOIS: It is getting late.

CHAIRMAN STETKAR: Well, but it is --

MEMBER BLEY: It is not bedtime quite.

CHAIRMAN STETKAR: It is still only nine o'clock at night. He is still young. He can stay up this late.

(Laughter.)

MEMBER BLEY: Probably not as young as you think anymore.

(Laughter.)

CHAIRMAN STETKAR: He is still a lot younger than I am.

MEMBER BLEY: And I guess if we are waiting for him, on that last one, it would be interesting to understand how the high-end bound is as high as it is. It is really high.

CHAIRMAN STETKAR: It is.

If anyone is out there on the bridge line, could you just say something or make some sort of sound, so we can confirm that it is open?

MR. JULIUS. Hi. This is Julius.

CHAIRMAN STETKAR: Oh, okay. Thanks, Jeff.

Vinh, are you there?

MR. DANG: Yes, I am.

CHAIRMAN STETKAR: Hi, Vinh. Happy new year.

MR. DANG: Thank you.

This is Vinh Dang from PSI.

CHAIRMAN STETKAR: Please explain our Bayesian stuff here.

MR. DANG: Okay. The Bayesian analysis takes only the counts, and the prior to Jeffrey's prior, a non-informative prior. And with the formulas for the Bayesian update, you just plug in the number of failures and the total number of trials, and you will get the bounds.

CHAIRMAN STETKAR: Fair enough.

MEMBER BLEY: Zero out of four isn't much information.

MR. DANG: Yes, exactly, and that is the reason why the bounds are rather large on the righthand side.

CHAIRMAN STETKAR: Because it is a non-informative prior, that is in some sense why it is kind of holding up the lower bound of that, right? Zero out of four is very weak evidence.

MR. DANG: Correct.

CHAIRMAN STETKAR: Okay.

MEMBER BLEY: You would have to have zero out of a thousand probably before you would start to --

CHAIRMAN STETKAR: Before you would start to pull that down.

MEMBER BLEY: Yes.

CHAIRMAN STETKAR: Yes.

MR. DANG: you would start to pull it down, and you would sort of get in there, yes, because you are not saying much about the upper bound as well.

CHAIRMAN STETKAR: Right. That is also --

MR. DANG: It doesn't take too much to start pulling it down, but it does take a while for the upper bound to come down.

CHAIRMAN STETKAR: Yes, yes, yes, yes.

MR. DANG: And the small difference that you see between the second and the third -- so, that is 1A and 3A -- comes from the fact that in the 3A case we only have three observations. I think the fourth crew did not. So, that is zero out of 3, actually.

MEMBER BLEY: I didn't remember that. Oh, okay.

MR. DANG: That is like just ever so

slightly higher.

CHAIRMAN STETKAR: Okay. Good. Anything more, Vinh? We are going to have to mute you because we get all kinds of static and really superfluous noise on our high-tech system here.

(Laughter.)

MR. DANG: No, I have no further comments, if you have no further questions.

CHAIRMAN STETKAR: Okay. Stay awake. We may come back to you. But thank you.

MR. DANG: Okay. You're welcome. Bye.

MR. FORESTER: The Jeffrey's prior should at least be in Volume 3 of the international study.

MS. LOIS: Yes, it is.

MR. FORESTER: I would expect it is in there.

MEMBER BLEY: It might say what he just said. I didn't remember that, that it was only Jeffrey's prior in the counts.

MR. LIAO: Okay. About HCR/ORE and CBDT, I think it was observed in this study that analysts made different judgments in addressing complex scenarios. This is because the method basically assumes operators are able to follow procedures and procedure cues are available in situations where procedures could not have coped well with the complexity, and the available time is short. Analysts have to apply their judgment to make the method of critical situations and to treat HFE-specific aspects of procedural guidance and diagnosis complexity.

Another difference among those teams was those teams obtained different timing estimates for diagnosis and execution. This may be explained by a method needing guidance on what to include in the timing analysis. Different analysts may make different decisions on how to account for factors such as delays in instructions and parallel actions.

CHAIRMAN STETKAR: And that correlation, in particular, is really, really sensitive to fairly small differences --

MR. LIAO: Yes.

CHAIRMAN STETKAR: -- in those time estimates.

MR. LIAO: Yes, when the timing is short, it is very sensitive.

CHAIRMAN STETKAR: Yes.

MR. LIAO: And another difference between the teams is one of the teams decomposed the HFEs into subtasks and quantified their subtasks separately, but the other two teams did not. This difference can be explained by the method needing guidance on when and how to decompose HFEs into subcomponents.

The two SPAR-H teams used very different qualitative analysis approaches. Team 1 treated HFEs as single tasks. In contrast, Team 2 decomposed HFEs into basically events, based on break points in procedures. They also considered transitions in the procedures. This is because the method SPAR-H is a quantitative method. It does not provide much insight on qualitative analysis. So, both approaches are consistent with the method. However, Team 2's decomposed approach contributed to relatively-better qualitative predictions, but the qualitative analysis did not lead to better quantitative results.

One reason for this may be the analysts had limited experience with the method. Another reason is that the method has limited guidance for translating qualitative analysis into quantitative results.

MEMBER BLEY: Harry?

MR. LIAO: Yes?

MEMBER BLEY: I don't know how hard you guys looked at this. I find this one interesting because I usually like the idea of decomposing to get at the detail. And it sounds like the detail they got at they did pretty well. But, somehow, using SPAR-H -- and you get to go through the factors and just pick, as I remember, 1, 2, 3, 4, 5 for the value of it -- it is looking like somehow, you know, you picked what was important one and you give it a bad number, and then the next subtask -- that somehow that method and the way it is easy to just pick things out without paying real attention to the quantification associated with it drives you to get the same answer.

(Laughter.)

CHAIRMAN STETKAR: That is a sort of central limit there.

MEMBER BLEY: We ran across a method like that once a long time ago --

CHAIRMAN STETKAR: Yes, we did.

MEMBER BLEY: -- where it had lots of detail, but it washed out all the effectiveness because, when you gave each one some credit, it just essentially got to the same answer every time.

CHAIRMAN STETKAR: That is right.

MEMBER BLEY: About this one, about 10 to the minus 2. It is interesting. It would be interesting to really understand that better, why that decomposition led to this. I don't know if this was the same group that got the flat line on the one John showed us or Erasmia showed us.

MS. LOIS: Yes, that is the group.

MEMBER BLEY: It is the same group?

MS. LOIS: And actually, that group did this detailed analysis to demonstrate the capability of the TRC concept, for which we were talking this morning. So, it seems that it was applicable analysis, but we really do not -- I mean, probably we should ask. We were not going back; we did not go back to the HRA teams to ask questions after the workshop we had.

MEMBER BLEY: I mean, to do a really good qualitative thing that drives you into a useless quantitative one is curious. It is worth understanding.

MR. FORESTER: Yes, it is a bit of a mystery; it really is.

CHAIRMAN STETKAR: Some of the things that we looked at -- and we are talking about the same thing that --

MEMBER BLEY: We are.

CHAIRMAN STETKAR: -- I came to the conclusion that in many cases the range, in that particular methodology, the range of goodness to badness, if you will, of performance-shaping factors and the unwillingness of an analyst to say that something was absolutely good or absolutely bad tended to --

MEMBER BLEY: Or didn't count.

CHAIRMAN STETKAR: Or didn't count.

MEMBER BLEY: He gave everything a little bit of weight.

CHAIRMAN STETKAR: Well, "I can't completely discount this" or "I can't assign this at either the high end or the low end or even the high-to-low end" differences tended to just, as Dennis said --

MEMBER BLEY: It drove us for a while to say, "Pick out the one thing that is most important and base your quantification on that."

(Laughter.)

MEMBER BLEY: Because if you treat it all, unless you are really willing to be discriminating, you get the same answer. And I am just curious if that is what is happening.

MR. FORESTER: That is a really good point. I hadn't really thought of that in the sense that it could just be that they do this great analysis, but, then, their final conclusion is, well, these crews are going to be able to handle this, so they are all nominal in a sense, even though we are, in fact -- or isn't nominal.

MEMBER BLEY: In each case, you give something a bad mark. Overall, every event gets a bad mark or everything gets middle except for one, and the middle washes it all out. It would be interesting to understand it.

MR. LIAO: I just wanted to add one comment. It seems to me this thing did not go back to do a reasonableness check after they assigned all of these --

MEMBER BLEY: It looks that way, yes.

MR. LIAO: But they could have figured out, changed the HEPs if they had a chance to go back to look at the original --

CHAIRMAN STETKAR: Yes, this notion of at least things ought to line up --

MR. LIAO: Yes.

CHAIRMAN STETKAR: -- in relative terms the way you would expect them to, it doesn't seem they did that.

MEMBER BLEY: Yes.

MR. LIAO: And in this particular scenario, you had --

MEMBER BLEY: The one out here is lower

than the --

(Laughter.)

CHAIRMAN STETKAR: Yes, you are right. I'm sorry.

MEMBER BLEY: It is close, yes.

MR. FORESTER: Yes, overall, I mean, in general, the teams did a little better than this study, I think, at identifying the difficult and the easy. I think you see generally a better trend there. So, they sort of came maybe, and you sort of get the sense that it was more apparent where the difficulties were.

So, then, as Harry points out, the fact that they didn't really do a reasonableness check in a sense and make sure that it fit that is just -- you know, I will say that that group sort of saved us in a sense. We were short of teams. We tried to get as many teams as we could, and then we lost it along the way. And they volunteered to do the analysis with not -- I mean, they were there for the interviews.

CHAIRMAN STETKAR: They were the homeless guys that you went out and paid?

(Laughter.)

MEMBER BLEY: The troubling part is they did the qualitative analysis well.

MR. FORESTER: They did, and they participated in all of it, but they were a little bit rushed compared to some of the other teams, I would say.

MR. LIAO: So, another difference between the two SPAR-H teams is that the two teams, sometimes they account for some factors on different PSFs, which can be explained by the method and limited guidance on PSF selection. For example, one team accounted for lack of plant cues under complexity; another team accounted for this factor under human/machine interface. The two PSFs have different multipliers. So, this led to their ability in the HEP.

Okay. Now I will come to ATHEANA. There were substantial differences between the two ATHEANA teams in how they did their qualitative and quantitative analysis. Team 1's HEPs consistently were lower for many HFEs, but there are too many differences to identify the main cause.

One notable difference between the two teams is how they applied the method. For Team 1, they first developed a scenario map representing all plant crew responses. And then, they used a Monte Carlo simulation approach to quantify the scenario map with probabilities and the task duration distributions. This approach led to a very clear link between qualitative analysis and resultant HEPs.

In contrast, Team 2 used a more holistic approach in identifying operational challenges. They decomposed the HFEs into some unsafe actions, but they did not quantify those unsafe actions separately. In addition, unlike Team 1, Team 2 only produced time point estimates and time analysis. But it should be noted that, although Team 2's analysis was less detailed, it seems to be quite effective.

The differences between the two teams seem to be the implementation of ATHEANA is quite resource-intensive. So, different analysts may make different decisions on how to apply the method.

So, overall, the findings on method strengths and weaknesses from the international study were confirmed in the U.S. study, but the U.S. study produced more findings.

By using multiple teams for a given method, they were able to separate method effects from analyst effects and test consistence issues across HRA predictions.

One finding about a method effect is that we observed in the U.S. study there are different teams using the same method, but formed a qualitative analysis with different scopes. Also, sometimes when analysts identify some issues, they are beyond the scope of a given method. They had difficulties in quantifying the issues. So, those observations indicate more emphasis is needed for structured, consistent qualitative analysis, and a coupling of quantitative method to quantification models needs to be improved.

Another method, in fact, is the HRA methods rely on human analyst judgment, especially when there is a lack of clarity and specificity in method guidance. Different analysts make different judgments in implementation of the method, and leading to their ability in their HRA predictions. One solution to fix this is to extend the guidance on method implementation and test the guidance for analyst-to-analyst interpretation.

Now we come to analyst effect. One finding about analyst effects is that we observed analysts would make different judgments in compensating for method limitations, leading to variability in their results. For example, like I said before, when a factor is beyond the scope of a method, the analysts have to apply their judgments to stretch the method. You know, they may stretch the method in different ways. So, since this is not a guidance issue, resolving this issue may require method modifications.

Another finding about analyst effects is that allowing plant visits and staff interviews may have been a contributor to relatively-better HRA predictions in the U.S. study, but it also caused some problems contributing to variability. Different analysts used different interview skills. They made different assumptions, interpretations, impressions, based on the information from the visits or interviews from operators. So, that leads to the differences in their analysis and also indicates that improved guidance is needed for performing interviews and collecting HRA information.

MEMBER BLEY: Harry, can I take you back to effect 1 with a question? I am remembering something Alan Swain did years ago. He did a comparative study of the available HRA methods for the Germans. This was 20 years ago or more, a little green book.

In the end, he had an appendix where he contacted the developers of each of the methods and asked them about limitations of their methods or how could it apply to certain problems. And he found them a little disingenuous, but I didn't. I found a different story.

Essentially, each person who developed a method could solve any problem with it by adapting their own method, and they were confident about doing that, where all of them said the other methods weren't flexible and they weren't willing to fiddle with somebody else's method.

This effect of changing the model to deal with things, was there a distinction between the folks who had developed the method and those who hadn't developed it in their ability or willingness to compensate for limitations?

MR. LIAO: There is no evidence from this study, I think, in my opinion, from the study regarding difference between developers and not --

MEMBER BLEY: Okay. Because in almost every case, one of the teams was a developer and the other was not, I think.

MR. FORESTER: Not for ASEP.

MEMBER BLEY: No, not for ASEP, that is true.

(Laughter.)

MR. FORESTER: We would have to look back on it. Let's see.

MEMBER BLEY: Well, I just wondered if it stood out to you.

MR. FORESTER: I don't think we got -- no, it didn't stick out.

MEMBER BLEY: Okay. And I hadn't noticed it in the discussions earlier.

MR. FORESTER: That is true; that is something we ought to look at.

MEMBER BLEY: Just curious.

CHAIRMAN STETKAR: Harry?

MR. LIAO: Yes?

CHAIRMAN STETKAR: On the next slide, if we go to the next slide, that differences in interview skills, did you document the interviews?

MR. LIAO: I did.

CHAIRMAN STETKAR: Can you actually correlate -- you said that the interviews, you know, in some sense were good, but in another sense perhaps increased variability. Have you looked at that process to understand why?

MEMBER BLEY: And if you have, have you thought of coming up with guidance for people on how to do the interviews?

CHAIRMAN STETKAR: On how to do the interviews?

MR. FORESTER: We have definitely thought about -- go ahead.

MR. LIAO: Well, I just briefly looked at their interview scripts and their reports. And basically, they asked questions about timing, experience, and training, and some HMI issues, but the fact is how does that impact their response.

And it seems like some teams used a structured method. They did quite a lot of analysis before they went into interviews. Some teams, now before they went into interviews, they had like some questions, a list of questions they wanted to ask the operators. And for some other teams, they were less structured and the analyses were less detailed before they went to structures (sic). And also, they asked different questions in their interviews.

Regarding how to correlate this to the results, I think in my opinion it is very difficult to do a structured, systematic analysis on this.

MEMBER BLEY: Now you did have somebody -- I don't know if they did it for everybody, but I think for many of the interviews you had somebody who took notes about what the questioning was.

And I think what John was getting at was the way they asked the questions, the way they tried to establish rapport with the people they were dealing with, were some more effective at getting out information?

CHAIRMAN STETKAR: And even in some cases, did they ask about the quality of -- did they ask me, as an operator, about the quality of my performance? Or did they ask me to objectively evaluate something else, you know, the goodness or badness of something else, without making a direct judgment. "How easy is it for you to understand this indicator?" rather than "How good is that indicator?" And those are very different. You will get very different responses when you think you are asking about the same attribute.

MS. LOIS: I just wanted to add that Halden had an expert observing the interviews, and they were making notes of how the interviews were conducted. And I believe that they have done a follow-on analysis addressing the issue of how HRA experts are applying the methods at some level. And also, they had some interviews with the NRC staff that participated in the study, et cetera.

I don't know if Andreas is on the phone. If he is, he can let us know. But I do know that there is a follow-on Halden study that is addressing that at some level.

MEMBER BLEY: No drafts yet?

MS. LOIS: They may have, actually. I don't know. We haven't followed up.

(Laughter.)

But I don't know if that would correlate with --

CHAIRMAN STETKAR: It would be interesting.

MS. LOIS: -- the outcomes, but they do have --

CHAIRMAN STETKAR: Let's see if we can get the bridge line open --

MS. LOIS: Yes. Okay.

CHAIRMAN STETKAR: -- and at least see whether he is on the line.

MS. LOIS: Halden is plus six hours or it is 10 o'clock.

CHAIRMAN STETKAR: No, it is 9:30, but --

MS. LOIS: I don't know if --

CHAIRMAN STETKAR: It is wintertime. It is dark there all the time there anyway.

(Laughter.)

MEMBER BLEY: They have got a rock band among them. So, they might be out playing somewhere tonight.

(Laughter.)

CHAIRMAN STETKAR: On a Wednesday night? Why are they being engineers if they have a paying gig on Wednesday night?

(Laughter.)

MEMBER BLEY: I'm not sure it pays.

(Laughter.)

MR. LAI: It is open.

CHAIRMAN STETKAR: It is open?

Okay. Just somebody say something to make sure it is open, anybody.

MR. JULIUS: This test confirms the bridge line is open.

(Laughter.)

CHAIRMAN STETKAR: Thank you very much.

And now, is Andreas on the line?

(No response.)

MEMBER BLEY: I guess not.

CHAIRMAN STETKAR: Well, one out of two is not bad. We will reclose the bridge line.

That would be interesting if, indeed, they are producing a report because --

MEMBER BLEY: I got interviewed. That is right, they are doing that. That would be very interesting.

CHAIRMAN STETKAR: It would.

MEMBER BLEY: I would like to see what they --

CHAIRMAN STETKAR: Because even within the context of IDHEAS, there will be, I am assuming, some form of exchange with operators to set the context of the scenarios.

MEMBER BLEY: And I think we need to have some guidance.

CHAIRMAN STETKAR: Understanding some potential pitfalls in that discussion process might be really, really useful, because I have run into it a lot, you know.

MEMBER BLEY: Yes.

CHAIRMAN STETKAR: Good. Thank you, Harry.

MR. LIAO: Well, thank you.

So, next, compared to the international study, there was similar variability in crew performance in the U.S. study. Although the HRA predictions in the U.S. study were somewhat better than those in the international study, there was no evidence indicating this was due to a crew effect, meaning using U.S. crews versus European crews.

MR. FORESTER: Also, just related to that, there was another issue that came up, that there was some concern that some of the analysts might be biased in terms of how they expect crews behave, based on their experience with U.S. crews. So, that might have hurt their ability to predict what the European crews were going to do. So, this was just another step. Now we have all U.S. crews. So, it wasn't an issue anymore. It is not like we can compare, but at least we have controlled for that issue. So, that was another part of this.

MR. LIAO: So, in summary, the U.S. study results provided clear evidence of method limitations and indicate specific ways or areas to improve individual methods. Essentially, all methods have limitations in qualitative analysis and the interface between qualitative analysis and quantitative analysis.

The U.S. study also identified the limitations in HRA practices. For example, improved guidance is needed for estimating time for post-response actions, reasonableness checks, and how to conduct interviews.

So, since no method, no single one method is better than other methods in all aspects, and all methods have limitations, one important conclusion from this study is that creation of a hybrid method combining effective elements and the features of existing methods should significantly improve HRA and HRA practices.

And that comes to the end of this presentation.

CHAIRMAN STETKAR: Great. Thank you very much.

Any further questions for Harry?

MEMBER ARMIJO: I had one question for these one order of magnitude difference among predictions from a given method for most of these HFEs. Why is that satisfactory? Isn't that reasonable to expect in something like this? No matter what method you use, you are going to have quite a lot of variability? Is that a good result, an order of magnitude?

MR. FORESTER: Well, it is better than what we saw in the international study.

(Laughter.)

MEMBER ARMIJO: But, you know, would you, for example, set a goal? Say, "Look, we want it to be within a factor of two," almost like a machine could do this? I don't think that is realistic, but I am just trying to understand where you are going.

MR. FORESTER: Yes, I understand.

MEMBER BLEY: Did all crews do an uncertainty analysis, all analysts? I mean, that is not shown here.

CHAIRMAN STETKAR: Yes.

MEMBER BLEY: The analyst uncertainty analysis is not shown.

CHAIRMAN STETKAR: That is an important differentiation because these are supposedly, we will call them, best-estimate numbers.

MEMBER ARMIJO: Yes.

CHAIRMAN STETKAR: I mean, one would hope that applications, regardless method-to-method or analyst-to-analyst applications, at least on a best-estimate perspective wouldn't show this degree of variability.

MEMBER ARMIJO: Yes.

CHAIRMAN STETKAR: There may be very large uncertainties in the absolute value of the human error probability, which is what Dennis is getting to, but this does not display those uncertainties. There may be a factor of plus or minus 10 or more around these best estimates, but one would not hope this degree of variability in those best estimates.

MEMBER ARMIJO: Okay. Okay. Right. Just

looking at it, I would say, oh, boy, if I had to buy a method, I would buy ASEP right now. You know, I am just trying to figure out where this is all going and how much variability would be acceptable, even in a best estimate.

CHAIRMAN STETKAR: You actually might buy the HRA Calculator, because I think, John, you said that the 1C low-lying, at that point it was because they had misinterpreted something, right?

MR. FORESTER: Yes.

CHAIRMAN STETKAR: So, the HRA Calculator, removal of that might also show a similar, even a tighter --

MEMBER ARMIJO: And for some reason, ATHEANA was really underpredicting on that very difficult case, yes, for both teams.

CHAIRMAN STETKAR: And SPAR-H, you can either get the same number or widely-different numbers.

(Laughter.)

MEMBER ARMIJO: Everything is 10 percent.

MEMBER SCHULTZ: Harry, could you go back and describe in the conclusion slide that you have the limitations and the practice? The first bullet is estimation of time required. Can you expand on that?

MR. LIAO: Well, for some methods, for example, ASEP, we need to -- I am trying to organize my thoughts -- for ASEP, you need to first estimate the time required to finish some actions based on procedures. And then, based on that, you get an estimate for how long you could get to diagnose a state of events and for using a time reliability curve. So, this is how to estimate the time for response actions.

For example, two ASEP teams made different decisions on what steps to include in analyzing post-diagnosis actions. And also, CBDT or HCR/ORE, they also made different decisions. And some teams did not consider delays, disruptions, or parallel actions that would have caused their ability and their timing analysis. So, those factors, you improve the guidance for these two; the guidance needs to be improved for considering those factors, how to improve their timing analysis.

MEMBER SCHULTZ: The benefit would be to perhaps take that out of the analysts' hands, get it into the methodology? What is the recommendation to improve on the limitation? Is it to provide more guidance to the analysts related to that or to embed it in the methodology?

MR. LIAO: I think, well, in my opinion, both, more guidance or improve the methodology to more structured timing analysis to reduce the variability.

MR. FORESTER: Because some analysts may look at the procedure and very carefully think about how long it is going to take them to do these steps and what is going to be going on in the scenario. And some of the early methods are very simplistic about estimated time required. They say, you know, how long does it take them to walk across a control and then turn this dial. So, it gets very focused on the actual simple execution part rather than a lot of the other aspects that could really increase their time to get to their response. So, a lot of it is a matter of detail.

MEMBER SHACK: And we talked about reasonableness checks as being something that, from the results, you can observe and say, well, clearly, it would have been helpful if the analysts had taken the time to look at the results and make relative comparisons and assure that the results made sense, were reasonable.

MR. LIAO: Yes.

MEMBER SCHULTZ: So, there is that piece.

With regard to the conducting of the

interviews, what I thought I heard you say was that that result or limitation identification came from an observation of the interviewing process, not to look at the interviewing process and then connect it back to the goodness of results. Or was it both?

MR. LIAO: Both. There was evidence indicating different interpretations or assumptions based on information from interviews. They made different judgments or decisions in their analysis, causing their ability and their results.

MEMBER SCHULTZ: Okay.

MR. LIAO: Yes, both. Both. We have evidence related to, correlated to HEPs and, also, evidence in observation of their interviews.

MEMBER SCHULTZ: Okay. Good. Thank you.

CHAIRMAN STETKAR: Anything more?

(No response.)

Thank you very much.

MR. LIAO: Thank you.

CHAIRMAN STETKAR: Now something everybody is always interested in, and have been for 30 years, is data.

(Laughter.)

MR. CHANG: Good afternoon.

My name is Jim Chang. I am the Human

Reliability Engineer. I work at the Human Factors and Reliability Branch of the Office of Research.

My principal responsibility in this Branch is data. So, I am here at this time to brief the Committee about the SACADA data that we developed for HRA.

CHAIRMAN STETKAR: That is not the Tacoma Narrows Bridge, is it?

(Laughter.)

MR. CHANG: This is a bridge; one side is Norway; the other side is Sweden.

The NRC did a number of the HRA data before. But, so far, we still complain about it, insufficient HRA data. And so, we set up this SACADA data hoping that, one thing we wanted to do was, okay, let's think about a way. What is the human performance data that can go directly to the HRA application?

The HRA application here, we targeted three things. That is all in the NRC's risk-informed decisionmaking program. That is a PRA model. Understand that on the HRA program side, that myself, three years in location, four months in location into the trench doing the STP to know how the STP is performed and then the characteristics of the data in terms of the test data.

On the human performance data side, we conducted workshops looking for what is the potential data to NRC, including operating experience data. This, including the operator initial training data and operator recourse and training that is included in the NRC. It is not within the NRC control.

And so, these areas, because we should be practical that we need to consider appearance of these two things. I want to get the Committee's attention here, that there are two sides of bridges here.

When we started this, we say okay, because we look at the various HRA methods. Each method has their own problem. We do not want to have the data support a particular method. Instead, we are looking for, directly inform the HRA application, looking for information, while we believe that the information we collected could be useful for informing the HRA method.

So, under the objectives here, the goals, as we defined it in two objectives. The first thing is that focused, very basic foundation here that HRA is too general. So, how we use data to provide the HEPs, inform the HEPs with good data information.

And in this year, our objective here is to produce anchor HEPs; that we are not going to replace the HRA method. Say, okay, those are to support all kinds of data. No, we recognize that this data has its limitation, so that we provide some interface on what data we collect, and we provide some probabilities, so that can be used for HRA method, to inform this method.

And coming to the end, we focused on this data, licensed operator training data. That is because the data foundation here is the most practical way that we go there. But the NRC does not want the licensed operator.

So, for this, in order to obtain this data, that we work to get this data. And the tool is the SACADA database. I talked about SACADA data in my slides.

Doing the HEP quantification here, we cannot avoid the qualitative analysis part. That means that understanding the operators' behavior, response to the accident or instance situation. We need to have an understanding of that.

On the data side here, we worked in more emphasizing knowledge, instead of going to document operator response in complicated events, things like the empirical studies of the real events. So, on this end, we focused on the event identification and the research-oriented simulator exercise.

We have further improved the event timeline structures, but it took the past two years that we have put in more effort on the SACADA side, and that is effort on the objective 2. The prospective is the countdown to the end, that these two will merge as two SACADAs based on the same theoretical foundation, just a slight differential mix, but they will be in the same framework.

SACADA, here is a scenario also in preparation and the debriefing. Also, in preparation and debriefing this, we would present the three functions of this software. Also, it means that the scenario that characterizes what is the training for operator response to the situation. And the debriefing is the performance; what is the performance result? Our data, my slide will explain this in detail.

So, we collected the licensed operation training data, and not only to provide error probability again, but we also needed to get reasons, understanding. Given all the failures, human failures we observed, what occurs; what is the error most? What are the causes of these errors? So, this is contextual information to support the number, and then, also, the information you are able to use for the HRA method to calculate, to improve their method.

The SACADA here, we deferred under the original agreement between the NRC and the South Texas Project. That agreement was signed in March 2001. We continue under it to work with them.

This agreement, there are two things for this operation. The first thing is we collect data; it will not be used for regulatory activity. Essentially, that enabled us to work with a plant. We defined this as a research activity. We need to have this in the agreement, so STP will work with us.

The second is the data propriety to the data provided that in this case is STP. But the NRC, to reach cooperation, we have access to a good portion of the human performance information data.

The NRC is not able to use the data that is agreeable to the NRC to pinpoint which individual crew made this error. But for the information entered into the HRA data, we pretty much can get it.

Currently, we have the database in operation. This database is maintained by INL. Currently, it is in transition from the experimental space to the production space. Let me say that in the operation mode, when in the full production space, the operation model is that we are looking for a long-term, sustainable data program. It is not just the three years or five years research program.

So, in order to achieve this goal, we defined how this thing would be operated. All data will be entered by the staff. Certainly, in order to achieve this discipline, something needs to be attractive in this software for people to want to use it in their daily operation. Currently, with STP, I think we achieved this goal, this objective.

So, here, the people doing this here, including the training department, the operator training, and then, also, the operations department that is operation.

The NRC and the contractor, we spot-check the data quality. We are not going to enter data.

As I said, they want to put the effort in the long-term in this database, that we need to be able to provide something that is attractive. The data here is not only for us to estimate the error probability, but it also has the information for the plant to improve their human performance.

And so, the database here, this is streamlined. That is from the beginning of the design situation scenario, and then through a group briefing. And then, how this information is used for improving crew performance, that information is in this software. That reduced their effort in especially the redundancy in the data entry. And STP staff was very happy about this aspect. They feel that they have -- in that aspect, they reduced a lot of redundant effort.

CHAIRMAN STETKAR: Harry (sic), may I ask you something? I am trying to thumb -- James, I'm sorry.

The frequency is increasing to, for example, two per day rather than just one per day. As soon as I start babbling completely incoherently, let me know.

I was trying to look ahead in the slides a little bit.

I know, I mostly babble incoherently. I hear the laughing in the background.

You have mentioned that you are transitioning from the developmental stage to the production stage --

MR. CHANG: Yes.

CHAIRMAN STETKAR: -- which is good. However, so far, we have heard STP, STP, STP.

MR. CHANG: Yes.

CHAIRMAN STETKAR: What efforts have you taken, are you taking to engage the remaining 71 sites? And what type of feedback have you obtained from them? Because, quite honestly, mining data from one, and only one, site really has very, very limited usefulness.

MR. CHANG: Yes. So, that I can jump to --

CHAIRMAN STETKAR: Okay. If you are going to address that later, that is fine. I will let you continue.

MR. CHANG: Well, it is just two slides down.

CHAIRMAN STETKAR: Okay.

MR. CHANG: I can explain it to you. We presented my paper to the PSAM 11 to provide this, basically, the methodology. At the end of last year and the beginning of this year, we hosted GoToMeeting, by GoToMeeting, to different domestic and informational organizations that included domestic.

CHAIRMAN STETKAR: Other than STP as being identified owner/operator of a nuclear power plant, I don't see any -- I guess KAERI might be --

MR. CHANG: Yes. Okay.

CHAIRMAN STETKAR: -- I don't see any other owner/operators, at least under those GoToMeetings.

MR. CHANG: Okay. These are the people who attended the GoToMeeting. The other two plants, I didn't -- one is Oconee. Earlier last year, I went to Oconee to show them in the context. And then, I met the Ops Manager and the Training Manager. I explained, showed them some of the screens, and they were very interested. But, at that time, the software was not mature enough. And what I told them was, once we have it mature enough, that we would like to invite you to review/comment. At that time the feedback was very positive. And they not only wanted to stay at Oconee at that level, but also want to, for the methodology.

When we hosted this GoToMeetings, invited then what was the wrong person. It was a simulator that supports a manager coming to the GoToMeeting instead of the manager. That one is in.

The second was the main shift trainer that was there, and then I met this chief trainer. And I showed him, and he expressed quite interested.

The same thing here, that when we host these, the GoToMeetings, the first time we invite him, his plant has an event. He could not attend. And the second time he was in the training session. He could not attend.

But these are the people we would like to go back to get in touch with for individual aspect.

MEMBER BLEY: James, could I throw something in for you?

MR. CHANG: Yes.

MEMBER BLEY: One thing they did that I think was very important, they had the entree to South Texas primarily because South Texas participated in the other project and got interested.

They spent a very long time developing this computer code, this SACADA, to enter the data, but working at the plant with the utility guys, the training guys primarily, getting it to the point that those guys were happy with it and were entering data and collecting their data that way and using it, rather than getting multiple people involved while it was still crude and rough. Because there was an awful lot of "Well, this doesn't work right. This is what we're doing."

So, they got it to the point that it is really good, beyond a prototype, before they have gone out --

CHAIRMAN STETKAR: Yes, that is why certainly before that stage you it could be very counterproductive to have too many people involved. On the other hand, if it is now being essentially rolled out for at least some preliminary use, you know, without fairly broad-based participation and support, it could be the best system in the world and it isn't going to be all that useful.

MEMBER BLEY: If I could throw one more thing in for him? And he will have to tell you more about this.

But the folks at the plant have gotten it integrated with this operations system software that many, many of the plants are using, so that it is part of, it is like an add-in into that software they are using at many plants. So, it ties in with things they are already using in many places.

I don't know if you have a slide about that later, but --

MR. PETERS: This is Sean Peters, the Chief of the Human Factors and Reliability Branch.

We agree with John. At other times before, we had relative maturity of our data-collection software methodologies. We had to try to coordinate with the plant and utilities. And some of the feedback we got, and we had various levels of participation, but the feedback we got was, "Well, until we can use it for our own purposes, we don't want to participate."

But we do agree with John, and we are starting this outreach initiative. James can tell you a little bit more about what outreach we are planning to do.

CHAIRMAN STETKAR: Okay.

MR. CHANG: Yes, we certainly know that, in order to, we need to have it coming in.

CHAIRMAN STETKAR: Yes, I was going to say I think it is essential, and it is really important. I think part of what Dennis was saying is that the utilities need to, for them to buy into this process, they need to understand that it is a benefit to them and that it is not an additional burden imposed on them, and that it somehow integrates pretty seamlessly with their training program.

For example, I don't see a sub-bullet on there that says anything about INPO or, you know, folks that get more integrally involved with training and operations. I mean, EPRI is okay, but it is not the same as INPO.

MEMBER BLEY: They did look at it --

CHAIRMAN STETKAR: Did they?

MEMBER BLEY: -- at their last visit to the site, and the guys were telling us something about that. I don't know if you remember more details.

MR. CHANG: STP wanted to use the SACADA to inform the INPO that they have used these things in order to get acquainted.

CHAIRMAN STETKAR: Okay.

MR. CHANG: So, INPO was aware of this, and plants are aware of this. But how they are going to feedback, we still don't know.

CHAIRMAN STETKAR: Yes, I mean, obviously, the dates on this thing were -- yes, this is real-time. It is just, I think, at least from me personally the message is that it sounds like it is time to essentially start marketing this in the sense, to see whether or not there is going to be better participation and acceptance in sort of the near-term timeframe also. Yes, because proceeding on if you are only going to get 10 percent participation may not necessarily be very productive.

MR. CHANG: Yes.

MR. PETERS: But, as a corollary, I think 10 percent is infinitely more data than we have ever gotten. So, if I could get 10 percent of the fleet to participate with all their operating cycles, it provides a significant amount of data.

CHAIRMAN STETKAR: It is a significant amount of data certainly for understanding scenarios, more the qualitative stuff that James was talking about. In terms of quantitative estimation, it is still pretty doggone limited. I mean, 10 percent of the fleet, 10 units with 5 or 6 claims per unit is still not a very broad sampling. I mean, you still have to look at it in that sense.

MEMBER REMPE: You have data like this, and people change with time and they add a training program. There might be a budget cut, and they might change things. How do you reflect changes in the situation with such a database?

MEMBER BLEY: You are far ahead of us.

(Laughter.)

MEMBER REMPE: Well, if you are going to use this data, I mean --

MEMBER BLEY: Well, you have the same problem with failure rate data when you collect it. And there were periods of time when the various industry data-collecting systems weren't being fed by particular plants, and you just had to come up with a way to deal with that. Keep track of who is in it, who is out. Keep track of the counts and deal with it, yes.

MEMBER REMPE: So, if a plant is entering data and you see trends of things are improving, they could take credit for that? Or if they are not improving -- I mean, I think what I am hearing is you are going to have the plants sending in more data, right?

MEMBER BLEY: Yes. Remember, this fits in that whole scheme of things you have heard before. This got started thinking of can't we get data from the plants. We always said, "Yes, you can. They will never do that. They will never participate." And James has really chased it and got this going.

And I am of the feeling you have got to start somewhere. The way they got in and got it moving was not to collect data for HRA, but to provide a system for the trainers to develop the scenarios they use for training and to keep track of the debriefing information in a way that forced them to be more uniform, more thorough in what they did, and have records they could go back to very easily to find the kind of information. The things you are asking are the pipe dreams of the next 20 years, my opinion.

MEMBER SCHULTZ: What I am saying, looking forward here, is that the improvement, the way that you have made this attractive to STP, and if you talk to Oconee, too, is you are talking to the training department, so they can improve their training process.

MEMBER BLEY: Exactly.

MEMBER SCHULTZ: It is an "Oh, by the way, maybe there will be some output here that will help the PRA side of things."

(Laughter.)

MEMBER BLEY: One of the first hints of this was we looked at the Robinson fire and developed a timeline of that with lots of information on the timeline. And we took that to a couple of the guys from NRC who were involved in the inspection following the Robinson fire. And they went through our stuff, gave us a lot of comments, and said, "You know, if I had something like this when I was running that inspection, it would have really helped me lay out the data and organize it and use it."

So, we started saying, well, how could we build that to make an inspection tool? And then, after the work at South Texas, you know, that same idea fits with what we do in training. And is there a way to make this useful?

MEMBER SCHULTZ: It is the next step for the training department and the simulator staff.

MEMBER BLEY: And that is where it has caught hold. It hasn't caught hold with the inspectors yet.

MEMBER SCHULTZ: It is to help them not only gather the results, but be able to implement and use them to improve their training program.

MEMBER BLEY: Yes.

MR. CHANG: The additional two points that I want to mention in this slide is that the portion of the outreach is that we got more positive and active -- I mean not positive, but they expressed already, expressed interest in working with us for using this method for entering data.

The Institute of Nuclear Energy Research in Taiwan, they worked with Longmen Nuclear Power Plant. It is an APWR, a full digital control room. Even the plant, it is not operating, but they already use a simulator to train their operators.

The INA has a two staff that went to this plant to receive our training and is working with us, looking to use this methodology to document the crew performance in the digital control room.

Korea Atomic Energy Research Institute, a few years ago, they had built a research program that went to, I think, one of their plants about this, collected about 100 scenario data used there, their method, and putting it into their ops the data. Many folks on the time information in this application come to great detail. They appreciate this SACADA method and want to sign a financial agreement with the NRC so they can have that agreement for them to put in the data using SACADA.

MEMBER BLEY: James?

MR. CHANG: Yes?

MEMBER BLEY: Would it be possible -- and I don't know if this Subcommittee is interested -- but at some point it might be good to have a demo of SACADA for the Subcommittee, so they can see how it works, how the data goes in, all that sort of thing.

MR. CHANG: Certainly. Certainly. Yes, yes.

The second point I want to make is in this, if I GoToMeeting there, we tell this is a web-based tool. So, we welcome them to try it for logging in and a password. And once they go through this INL security screening, they will provide them the log-in and password to this tool that we have a real scenario, real good end data and, of course, a sandbox for them to explore this database to get a feeling.

So, at the current stage, through this mechanism for the people to understand this tool, and then, once they get their comment and then waded through there, it is the next step of engaging them.

And then, back here, this line, so the concept of this, every person logged into this, as I said. And NRC staff, logged in there, the information I see, it will be different from, say, the plant trainer logging into there. That has a control here.

For this thing, it is that each station has identified one person as a data administrator. The people come in to go through the personal information go to an INL, pass the screening, and then INL is providing the account.

But the rule -- this must be -- in terms of what data this individual can see, you can add it. Okay, that is controlled by the station data administrator.

This picture is showing INL because this manager of data is able to see all the data. The NRC is able to see all the data, but not all the information, especially that information leading us to identify individual crew. We have some mechanics to mask the information, so that we got human information, but not going to the individual.

CHAIRMAN STETKAR: Can INL access that crew-specific or individual-specific information?

MR. CHANG: Yes. But not all INL staff.

CHAIRMAN STETKAR: That's okay, but I'm putting on my Harold Ray hat now, and is that a good thing in terms of marketing this out to participation of the utilities? Granted, that INL is not NRC, they are still outsiders. Many, many power plants are really, really reluctant about having an outsider, whatever hat they wear, see information about their people, identifiable information about their people.

MR. CHANG: Okay.

CHAIRMAN STETKAR: So, you may, in terms of this acceptability to a wider base of plants, you may want to think carefully about those interfaces in terms of where that information becomes identifiable.

MR. CHANG: Okay. Yes. Yes. Thank you. I will take that.

On the STP side, we haven't experienced this problem. Everything is fine.

CHAIRMAN STETKAR: I understand, but --

MR. CHANG: Yes.

CHAIRMAN STETKAR: -- you know, there is extreme variability out there in the industry in terms of philosophies and cooperation, and so forth. In general, my experience has been that utilities are very -- I say "utilities" -- power plants are very, very protective of identifying certainly individuals, and even crews.

MR. CHANG: Yes.

CHAIRMAN STETKAR: Crew XYZ, Plant A, is fine, but John Smith on Crew No. 7 at Diablo Canyon is not so good necessarily.

MR. CHANG: Yes. STP, the consumer is probably not, to the INL, able to see their individual person's performance. They worked with us, now still limited in the training data, not the exempt data. They are very separate. In the beginning, they say, "Okay, no, we are not going to put that data in SACADA." That still has the trust to go through --

CHAIRMAN STETKAR: Just keep in mind, you know, people who have been collecting hardware data, things like pumps and valves, utilities or plants in the U.S. share that data more openly now than they used to. Plants in Europe still will not release that data. "My God, I'm not going to tell somebody else about how my equipment performs." That is equipment. That is a piece of steel. It is not a human being.

MR. CHANG: Yes.

MEMBER SCHULTZ: As an indication.

MR. CHANG: This slides talks about it is, from the beginning, what we have constructed, if they have a need to have this psychological/theoretical foundation, that is this. We have taken the same framework model to make a function model to structure data, to correct information, looking at a different area like detecting, understanding, deciding the student action, and the teamwork aspect, like teamwork, communication, and supervision.

So, our methodology detailed much earlier than the IDHEAS method. At that time, when we started in this, the IDHEAS still are in the little research at that time. But we have constituted research information to make sure that, once in the future, now that we have current this data, that the information can be easy to inform the IDHEAS method.

In addition, to look at the information foundation here, we also look for what are the common elements in the HRA method here, elements to estimate HEP. How do we use this thing in our data collection, like these six pilot common elements using the HRA method? And then, it is the task; most likely, it is task decomposition, break down the HEP to a more specific task. And the generic tasks, like a spy uses diagnosis and action as a generic task. Error mode, like CBDT, a CBDT-type of approach, and performance-shaping factors, and error recovery, and task dependency.

But these things that we collect this information and embedded in this collection data, when I go through the all-case structure, you will see that these things are inside our data.

So, we are doing this thing that folks, as for now, focus STP under this and made it a scenario. But this slide is showing the conceptual data scenario design. The scenario is that in this initial condition it is a steady-state in the normal changing, like increasing power, decreasing power, et cetera. And then, given this scenario, that the trainer will inject the malfunction in this scenario, and for each main function see how the crew responds to this malfunction.

In the scenario design here, it goes into more detail than just malfunction. You come to this table, for example, there is about 90 minutes in this scenario. Malfunction of the loss of ECW, when it was injected into this scenario, in the scenario design the trainer, they fill out this table beforehand. These are the things, the crew tasks they want to see.

Say that, okay, the first thing they want it recognize loss of EW flow to a train. And because this recognition can be detected by any crew, there is no specific person that sees this. It is simple, just that crew would detect that.

But the third one, that manually trip diesel generator, before any diesel generator is tripped, in this case, a table, okay, that is a strict manager exactly that they will do that. So, they are first defined. A trainer develops and identifies these tasks before allowing its generation.

And all these tasks that it is color-coded for, they are important. Red are the critical tasks. Once the crew fails the critical task, that is more consequence for their remediation followup training. And then, orange, yellow, white, et cetera.

So, the way they design simulation scenarios, it becomes very convenient for us. The most thing was the task, in essence, what is our image of NNSS. And the team that are domain experts -- they are from PI, HI, cognitive scientists, and then also, the operators -- always on our team internally.

After then, we decide, okay, there is discussion whether we should use the malfunction made available for the elements. This table, we decided, okay, that will be the elements. So, each element is our data way that we come into the data-driven that I will explain later.

So, for each element, it is our data point. We say there are two segments of data out there. The first thing is the context. What is the situation to which the crew is responding? That is the SACADA. There is the "C", the characterization. What is the context?

And the second portion of the information is responding results. So, given this situation, how is this specific crew with performance? That is the information that we capture. This portion is debriefing. So, each data point, we have context information and the response/result.

The context information is entered by the scenario designer, the trainer. When the scenario designer, there is a designer scenario, and then have a training crew to try out the scenario. So, at that time, this scenario designer will provide the context information for this scenario. So that one set of context information will apply to all crew. For a different crew, they will have a different response to the same. So that each crew will have their difference of performance results.

This software, it has four main functions that is the SACADA. It aids also, in that we provide that for the trainer to design the simulation scenario. And current, the status is STP testing. We have to implement, and STP is just beginning to use this to design the scenario.

The second part is the characterization. For each element, what is the cognitive challenges? That also the scenario designer will enter the information, and all we see is function and STP is testing.

The debriefing portion is document performance results. It is performing. If their crew finishes their simulation, they get together to discuss their performance. They fail this portion, actually, first because the STP wanted that function to implement first.

We implemented starting in the main last year, that it implemented and was accepted by the crew. And now, they are using this debriefing portion in their debriefing.

CHAIRMAN STETKAR: Just out of curiosity -- and because we are on the record, you don't necessarily have to indicate this information -- was STP doing a formal debriefing, kind of a structured debriefing, before you talked to them about this framework, if I can call it framework instead of a database?

You said that the reason the debriefing is implemented is because they really wanted that.

MR. CHANG: Yes, yes.

CHAIRMAN STETKAR: Is that because there was a gap or a deficiency in their post-training evaluation?

MR. CHANG: No. Really, it is because the training, once the debriefing is done, they want document, proof of performance. So, that is a key piece here, making their function. So, we developed debriefing first for them to use and then get their engagement.

CHAIRMAN STETKAR: Okay. Okay. Thanks.

MR. CHANG: The four portion of the data --

MEMBER SCHULTZ: I would think it would be to develop consistency and ease of going through the debriefing process.

CHAIRMAN STETKAR: That is a little bit of where I was heading.

MEMBER SCHULTZ: Yes.

CHAIRMAN STETKAR: I just think there was some sort of ad-hoc process in the past.

MEMBER SCHULTZ: I think that is the value that they are seeing.

MEMBER BLEY: They had a pretty formal process. They had data sheets, but this also, the thing they like is, as you go through it, using it, it pushes you to dig deeper and identify what went wrong in a consistent way. So, they have been really pushing for that, and they have decided they like it.

MR. CHANG: Yes. In the pilot, there is three people, individuals, including myself. With STP, we give them different debriefing when they are different week because those they observe.

Let me show you this. This was a task, before they used SACADA, they used the Microsoft Word. So, if we can prove in one room in the training there, could protect this screen, and then, they say, okay, how good we do this in this element, that element?

When I was there, one thing I got back the information. The trainer felt that the tool that we provided, they got much better information in the debriefing. Before that, his crews were in this end when we did it. There is nothing that is coming to their plant, but our screen in the debriefing, guiding screen by screen to them. Okay, ask, what is this aspect?

So, we say that they have a performance deficiency in certain elements here. And then, before they were simply typing what happened, and very short. That is not much information from there.

But after they implement this SACADA, we have screens that could lead them systematically to provide information. In the beginning, feedback from Dennis, in some situation, the crew is reluctant to provide information. And then, they say that, well, what type of error? "Yes, I didn't detect it around that indication, okay?" And then, the next screen asked why. And simply, that is right there and go to the next screen.

And then, the trainer discussed with the shift manager on the crew. So, now they, this for STP, would say, okay, if that is printing -- not everything that prints is put into their crew number. If that doesn't know, no information there, they will get it more. And that is what provides, make sure that the creator, that we will have all the information we want in the database.

Let me come back to this point. So, this is a screen from an NRC user here that I see. For the STP, the two units of PWR, they have 10 operating crew, four-staff crew; a total of 14, 14 crew. They talk about the crew 1, 2, 3. But on the main NRC data that you see that is Alabama, Alaska. This is the kind of masking with the crew, that we don't know which crew. But we do know which crew is that crew, for an operating crew, the information.

So, this is also in here, that that screen is for them to design the scenarios. They simply create what is the malfunction, and we didn't the malfunction, the elements, and then enter what is the color and who is the prospective performer of this element, and then what is the description of that element task.

And given the screen there, our trainer will come here to simply create this part and then go into the preparation. What is the tendency, what is the cognitive tendency in performing that element?

So, if the trainer has that preparation, and we will mention this entering screen. At the high level, we start identify what is the common activity there. It is that this is referring to training transport, procedure transport, going to the different, correspond to the other qualities of emergency action. Action there was better. By the end, it is consistent.

All these will have some more small portion of specific information for that particular activity. If not all this activity measured that comes was asking, okay, what is this task. Is this monitoring and detection type of task or the diagnosis, understanding for operation complication. What is the change, from the trainer perspective, what is the most changing? That could be multiple choice.

And also, that is coming to the bottom screen that goes through the safety system to identify if there is a safety system component involved in this element.

Because the monitoring and the diagnosis are the most likely places that cause the code error. So, --

CHAIRMAN STETKAR: I am sorry. Why do we know that?

MR. CHANG: What is typical. That is in the component activity action. It is not much changing.

CHAIRMAN STETKAR: The only reason I asked that question with that cynicism is that I want to make sure that there aren't any inherent biases or mental models of human performance built into this data-collection process.

MR. CHANG: Yes, yes. Yes.

CHAIRMAN STETKAR: If we know that something is most important and we have concentrated on that, we might be wrong.

MR. CHANG: Yes. Yes.

CHAIRMAN STETKAR: So, that is the only reason why I stopped you there.

MR. CHANG: Yes. For this direction, I don't think we --

CHAIRMAN STETKAR: It is just something to keep in mind. You know, whenever you are designing something that you are designing it with a particular mental model of how the stuff will be used and why you are querying certain types of data, that you always need to be a bit careful. I haven't seen anything here that leads me to that, but sometimes oral statements will give away things.

MR. CHANG: Yes. So, this portion, the trainer enters this information before running the crew. And he decided, okay, based on the trainer, in this scenario, what is the context? So, it is assumed that, okay, this is a monitoring type of test. Okay? And just making sure that, okay, the cue, this is more a detection cue.

So, we asked, what is the cue, the

scenario, the idea of the cue, what is it? Which indicator is it? And then, the second question we ask, what is the detection mode? Or to check is this a piece of the cue? Is the procedure directing the operators to do that or is the procedure not. For example, like this former page, there is monitoring for this type of activity. And also, the operator is not even monitoring activity.

And then, an individual indicator, that is not because this indicator is like this. It is what is the change of the indicator in terms of for the crew to detect the change, the differences?

I should draw the Committee's attention here. The question we ask is I believe quite objective, that they are related, that once they have done dry run, that these things, the objectives can be entered by the trainer.

The second portion of information here, that we have some detection mode, individual indicator mimic this. These are specific to detecting an indicator. That is the context.

And the second part here that we also have overarching contexts. This applies to all different makeup of different functions, like workload, time criticality, and then, some of these. That is only a three-item.

One thing we do is the language we use here, we try to make sense, that to the crew, that it is easy, understandable. For example, workload, well, we defined it on three levels, the different types of workload. And then, I asked the trainer, said, okay, not we define; how do you characterize the different type of workload? And these three levels is the trainer gave us. Okay.

Normal is that everyone, you do something, that you have a peer check. That is their definition of normal, normal workload.

And then, the second level will be concurrent demands. That means that one individual is doing something that he is doing his own. And then, the others have a peer check.

The third level is multiple concurrent demands. That is everyone is busy doing their work. There is no peer check.

So, that was the way that they characterized their workload, and we take that. Okay, that operation seems to make sense to us, and then we implement this thing here.

So, every portion of this cognitive function half of this has a certain context here that is twofold, like a cognitive function, and then a second portion that has overarching effects.

The next slide is the context for the diagnosis and response training. So, actually detecting a cue that is knowing the plant has certain problems, has a problem. And now, it is the crew's activity to typically track more information to confirm or to understand what is the plant status, what is the problem, assess this information.

So, here we get, okay, what is the diagnosis basis? Is that procedure-driven or skill-driven or is not a driven activity in this situation?

Even if there are indications, but a normal cue that you perceive, the operator expects a cue. Maybe in the shift turnover, let's say there is onsite maintenance activity. And given the cue of a normal cue coming in, the operator would think that, well, this is because of the maintenance activity, that something worked; it may cause them not to follow up this cue.

Information specificity. So, they have cue alarms. Does that alarm provide concrete information? So, what is the problem? Or is this a kind of test alarm? Say, okay, well, okay, that area has a problem, and then they need to put more cognitive effort to pinpoint to the problem, et cetera.

And again, the question we ask should be -- the trainer enters the objective, that was to increase the data reliability.

And this shows that for a PWR, these are safety components/systems for them to check.

MEMBER POWERS: How long does it take to do this entry?

MR. CHANG: Sir?

MEMBER POWERS: How long does it take to --

MR. CHANG: We just had an exercise that, for one scenario, this trainer took one-and-a-half hours to characterize, just to characterize all these changing elements.

MEMBER POWERS: So, if you have 2,000 elements, data entries per plant per year, that is one full-time person just entering data?

MR. CHANG: Well, the good news is I would say that one year as training cycle. Its training cycle may four scenarios. So, that is 20 scenarios in one year, 20 scenarios. And then, say that each scenario takes one hour. That is 30 hours to put all the characterization information in for one year.

This is one scenario for all crews.

CHAIRMAN STETKAR: Part of the question, also, though, is how much time do they spend. They need to spend time setting up the scenarios now already, whatever tools or lack of tools they are using. The question is, how much time do they spend doing that now?

MEMBER SCHULTZ: Yes, I can understand them loading the information in in an hour or hour and a half, but the thought process that goes before that --

CHAIRMAN STETKAR: Well, but, I mean, they have to set up the training scenarios anyway.

MEMBER SCHULTZ: Right.

CHAIRMAN STETKAR: They have to basically create the training scenarios with objectives and goals, and things like that. So, there is a lot of that --

MEMBER SCHULTZ: Of course.

CHAIRMAN STETKAR: -- planning process that is involved now. And I don't know how they document that process. I am not involved in training.

So, I know your concern is what added burden is this, and the question that I have is, is it any added burden? Is it just a different way to document what they do already?

MEMBER BLEY: Watching it myself, those sheets he showed you earlier are some of the old sheets they used to lay out the scenarios after they plan them. There is a lot of manual work, and then they use those to evaluate them.

They have been intimately involved in this development, and it has hit the point that, whenever it was a burden, they worked with the guys who were developing the code to ease the burden. And right now, the last time I have heard the plant guys talk, it is working about as good as they would have it. They might come up with new ideas in the future. But they had a whole wish list a year or so ago, and that has all been pretty much incorporated and they are using it, as far as I can tell, day-by-day.

CHAIRMAN STETKAR: Or at least from South Texas, you haven't had feedback that, "Oh, my God, this is really neat except it takes us three times as long as it used to just" --

MEMBER BLEY: Well, that is where they started. "If it does that, we won't be interested." But they worked with them to get it to the point it wasn't a burden.

So, I think now they find it, at least

some aspects are easier than what they were doing before, and it does some things better than they were doing before.

CHAIRMAN STETKAR: Well, the other thing is, if you do this --

MEMBER BLEY: And they don't see it as a burden. They see it as useful.

CHAIRMAN STETKAR: If you do it over a long-enough period of time, you now have an electronically-filed set of scenarios that you can call up pretty quickly.

MEMBER BLEY: And the fact that they built this into the computer system they were using for other operational stuff means it is now integrated better than it was before. So, I think all those were real concerns, and I think South Texas is at the point or past it. Other plants, it might not quite fit them. They might have to adapt it further.

MEMBER SCHULTZ: It certainly could be perceived that now you have an integrated database that you can reflect back on and trend.

MEMBER BLEY: And that is what they haven't done yet, but they are planning to do, yes.

MEMBER SCHULTZ: But they will be able to do that.

MEMBER BLEY: Yes, they will be able and they want to.

MEMBER SCHULTZ: With another database, it is very labor-intensive to do something like that --

MEMBER BLEY: Right.

MEMBER SCHULTZ: -- or if possible at all. Here you have consistent ways in which it has been done. So, there is more likelihood that you are going to be able to trend information about it.

MEMBER BLEY: I have to ask James, but my understanding is that is kind of the next phase, is to work at how to get information out of this. And I think Idaho will be still -- I don't know your budget and all that -- will still be involved in developing that side of the software as well --

MR. CHANG: Yes, certainly.

MEMBER BLEY: -- to do the analysis.

MR. CHANG: So, now let's shift to the debriefing. Now that this information was entered, that when the crew, operating crew or staff crew finished this information, they come through to talk about their performance element-by-element.

The way we designed the screen here, there is no problem, and then, that works; that is good; we make a note. But if there is a performance deficiency, how we capture that information. This is showing the Swiss cheese model or the defense-in-depth concept here.

We see the human, that one slice of cheese here. Given that situation, the change in here that the human needs to do the task. And this is a clean environment, so that individual can do this task. And then, there is a crew, possible recovery. But if the individual fails and there is no crew recovery or recovery too late, the scenario keeps operation, and then maybe another defense, not the human, but there are no consequences.

If there is no defense data, we see that, okay, that has an impact to the system scenario. And that is our concept of the design, the screen.

So, the first thing, in the debriefing they use this so well, and that is what projecting on the screen. The first thing, for elements, they can choose from these alternatives either set of their components as satisfactory. Or it is possible they are extremely well, outstanding in performing this activity. That means that there is a performance deficiency, but, overall, it is erroneous; it is satisfactory, or "unsat".

And then, there is early action. That

means that the crew didn't fail or didn't come to this task because of the things that they did in the earlier task. For example, the trainer, in the scenario, the trainer may be expecting that they have to work with many. But, for a certain crew, it has already then manually tripped the reactor. That is essentially the course of the scenario, not expected as designed by the trainer. So, that thing will key in.

And then, interruption, they say they have submitted a problem. If it is an outage, data very simply is not available.

So, they come into here. Once they enter the information, that means there is a performance deficiency. We want to go through the following screen to ask for more information.

And here, we ask for more specific data. That is the debriefing portion here. What is the performance deficiency? Is it a monitoring/detection issue or diagnosis or the procedure/making the decision or manipulation? Or the supervision/teamwork/communication? Any of them that we have followups are screened to do that. So, this is trying to understand what type of performance deficiency we see. And it tells us more specifics about the performance deficiency.

So, in this case, they say it is procedure/decisionmaking. And here, we are checking, was there a procedure available? And there is a problem because that means interpretation.

And the bottom is the notes that are written by the crews. There is a "Step 2.0s RNO not remembered completely." That was what the crew, they write some description about it on this interpretation, the way they checked.

And the next, we ask, what are the causes of that? So, here it is simple. The crew simply checks, okay, they check the unfamiliar and then, also, that it is person-specific, and no obvious causes.

And is this recovered, this screen? Yes, it was immediately recovered by the other team members, recovered that human deficiency.

And then, come to the next screen. As they come to the end, this thing, overall, has that impacted on the scenario? No, in this case.

And this is the final screen. Does this performance deficiency have an associated remediation or follow-through? In this case, it is no. So, I check the follow-through, just to try to show the

Committee about the options we have here.

And coming to the end, this is all the information entered that will be coming back to the main screen and then has that information they just entered, showing at this table, and that is the performance results.

One thing I mentioned about this information, we talk about 60 HRA methods, functions, components here to estimate in ASEP. There are tasks. One thing, it is about task dependency there. Here they provide us the information. Why if something is a PI and is doing the predicted, that, okay, we think that the scenario will go this way. As soon as the trainer thought about it, the scenario will go this way. But in reality, it was not. And this provides information why the scenario is wrong, wrong as expected. And this provides this information, that we know what has happened in all the scenario.

Come to the end, what we expected currently --

CHAIRMAN STETKAR: James, we are tight on time, but not really tight. Go back to that task dependency. I am trying to understand what this particular screen is telling me. Does this mean, for example, if I look at the "RO" line there, the third line, that that the person didn't do that because a preceding activity? Is that what I am trying to understand from here?

MR. CHANG: Right. In this case, I think this "RO" did not go to this task site to do this. It was here, the comment was that because, in this scenario I think that it was the trainer expected the reactor would automatically trip. But this crew has already detection --

CHAIRMAN STETKAR: Okay. So, basically, they just didn't get to this --

MR. CHANG: Right.

CHAIRMAN STETKAR: -- expected action because of some preceding --

MR. CHANG: Yes.

CHAIRMAN STETKAR: Okay. Okay. Understand.

MEMBER SCHULTZ: They weren't just watching. That is what it says here.

CHAIRMAN STETKAR: Yes.

MEMBER SCHULTZ: They were doing.

CHAIRMAN STETKAR: Being proactive. I get that term.

MR. CHANG: This screen is just showing that the prospective data at this point. Data output has three easy levels. The first thing is simple; we just count. These are critical tasks, but specific critical tasks, how many times it was encountered, that we know how many times it succeeded, how many times it had deficiency.

The second is action on the safety component. When this safety component was encountered, how many times did it succeed/fail?

The third is the time information. We haven't implemented; we are trying to implement this data this year. Now, for each element, we identify specific which alarm. And then, we say, okay, what is the action? Given that task, what is the action? So that we can associate which controller is active.

And for this information, we can get it from the generator data log. Read the log and then get the time information to reduce effort to get time information. That information is important to us, but we are certain that the trainer is not going to enter this information in here for us.

(Laughter.)

This is the way that we are able to get this information.

The second is context based HEPs. This is we care about our context, basically. What is the situation over detecting the cue? What is the context of diagnosis? And what is the context of doing a manipulation? And what is the communication for this element?

So, if performing two different tasks has some profile, context profile, in terms of easiest of the detection cue, diagnosis, then from the HRA perspective, we think that they should have a similar HEP, so that we can look at these two data points that together give us more data points to inform the changing of this profile, context profile.

That is our hypothesis. Whether this will work or not, we still don't know.

CHAIRMAN STETKAR: Yes, that one I am not sure about.

MR. CHANG: Right. Yes. Because one thing might be like safety cautionary information we didn't capture. They say, "Well, we have two plants. In essence, the same kind of file. Can we pool these two plants?

Earlier EPRI talked about this IDHEAS method, that if there is a similar type of plant that performs the same type of action, they should have the same HEP. If I am not mistaken, that was the -- this is all hypothesis. We don't know, but we rely on what the data tell us.

The other, for this thing, we correct the context-specific omission. Given this context, what is the performance? And then, what is the causes of that?

We exported information in XML format for the researchers to use to do more analysis of the information. The first bullet, the information, we think that we can have a function in the SACADA to do this job. The third bullet information relates to data export.

Okay. This I already talked about, the context-based HEPs. So, this is kind of a rough estimate of what are the data points that we have got. We have 104 units, and each unit for this plant, included in each, that is five operating crew and two staff crew. So, that is a seven-crew per unit. Five training weeks per crew. Each training week has four scenarios. And each scenario as four malfunctions. And each malfunction has a number of the elements. And so, for the total, every year, we have got 104. That means that we may get almost 300,000 data points.

Here, I want to jump to a separate Excel spreadsheet. This was a specific STP, this case, for last year. Last year, I mentioned that they specified that the top is pink. That was the operating crew. And then, the bottom is light green or light blue. Yes, yes. That is a staff crew. In this database, they separate the operating crew and staff crew.

On the righthand side, the first column here, these are the different scenarios. In the last year, the scenario is the data below because they had one force outage to get almost two months out of the crew, and no more training, and then, the scheduled outages.

It shows on the screen here the crew. And then, we see this operating crew. For this scenario, there are two sets, six data, and then 170 sets. Okay, set positive, that is a real good accident performance. And then, the four scenarios didn't come to them.

If we dig in deeper here, for this particular scenario, there are three malfunctions. And then, this malfunction that has a crew. And there is malfunction that has this element, this number of elements.

And here, okay, we see that there is one of the crew failed and set this isolation, main steam. I know Step 2. I don't know what was it. It is an immediate action failure. For this immediate action, I think this crew is likely to get a recommendation that they need to have some follow-through activity.

This is one of these outputs that was generated from the data, and the trainer liked this data. That is the word that they used. It provided them with the kind of perspective that what is the crew performance for them to follow up.

While we do not intend to use the data generated, we said, okay, well, every data generally we can use through the HRA. There is a gap between the simulator data which are supposed to be a real event.

One of the papers sent to the Committee discussing about the gap that we presented with examples, like a success criteria. Okay, what is the final set? As I said, I am sending those. On failure, there are differences.

And then, a person might say, okay, we know this; they know this simulator thing. Okay. Inject the seawater into the plant. Consider what the instructor will do.

And then, the crew composition. We can see that crew composition, being in this narrative training, the same crew usually did the training together, except that they have a very good reason, like attending a funeral or something. But in the operating crew that is not a situation. From time to time, people take a vacation, that the other crew can come in as a missed crew. So, that is a crew composition one.

And then, there is the other thing. Some of this is in terms of extra people. Entry here that it has a mixed effect. We need to evaluate this in how this data for this particular HEP, how that can be used.

And something that we simply did not address, we would say that we agreed. We have said the main concern is that the crew in the real situation that has a significant economic effect on the plant, well, that factor is not considered in the simulator. Even we got feed-and-breed number from here, but that gap, we do address --

CHAIRMAN STETKAR: In that sense, you are right. I mean, in that sense, if you are trying for some of these things -- for example, feed-and-bleed, putting ocean water into your reactor vessel -- this might give you some lower-bound estimate for real-world HEP. It certainly will not give you, you know, a realistic estimate from those types of reluctance issues. That doesn't mean it is not good. I mean, it is just you are obviously aware of that.

MR. CHANG: Yes.

Final slide. So, this very focused on to address the human performance information. That was to engage with a data provider.

CHAIRMAN STETKAR: The preceding slide, just something that came to mind was you emphasized, and your introduction noted that you have so far emphasized HEPs. From a human reliability analysis perspective, it is not clear to me what early benefits on numerical values of HEPs this database will provide. I am not sure. I would have to think about it.

MR. CHANG: Yes.

CHAIRMAN STETKAR: However, the scenario context I think is really important. You know, if you can look at when during a scenario, under what conditions, people tend to make certain types of errors, that might be much more useful in the long-run to support human reliability analysis than any particular numbers that come out of this.

I know you are focused on trying to provide a few anchor points or calibration values, or whatever you want to call them, for the numerical values. But don't disregard that other benefit --

MR. CHANG: No.

CHAIRMAN STETKAR: -- because the other benefit, that scenario context might be in the long-run as valuable or even more valuable than just strictly the body counts --

MR. CHANG: Yes.

CHAIRMAN STETKAR: -- the "N" out of 291,200, or whatever.

MR. CHANG: Yes, yes, yes.

CHAIRMAN STETKAR: I'm sorry.

MEMBER SCHULTZ: No, I think you have got the right focus in terms of what is the most immediate beneficial aspect of the program. It may not be a match to what original expectations were, in fact. In fact, I think what you are saying, John, is a lot more work needs to go in to figure out how to derive from this --

CHAIRMAN STETKAR: Once you have it, you know, where --

MEMBER SCHULTZ: How do you derive from this the data that everyone expected it to produce in the first place, where "everyone" is the analysts or the practitioner? It is not the operators and crews. Which is more important? I think it is where you have got what you have got right now, but it would be nice to create the opportunity for both to benefit.

CHAIRMAN STETKAR: Right.

MEMBER SCHULTZ: But it is going to take some work.

CHAIRMAN STETKAR: Yes.

Take square one, this next slide, where James was, that second bullet is absolutely essential. As long as you have got a buy-in by the trainers --

MEMBER SCHULTZ: That's right.

CHAIRMAN STETKAR: -- the operations departments and the trainers, that this is useful --

MEMBER SCHULTZ: That is the only way to get good data.

CHAIRMAN STETKAR: That is the only way you get good data. Mining that data and understanding where the most benefit from it is I think remains to be seen a bit.

MR. CHANG: Yes, so there is a certain effort here to complete doing the exercises on these also and the preparation. That feedback of the plant as to how we need to improve on this component and then continue the data output that is important to generate the information that we think is valuable for the entire community.

And then, at the same time, I think with this stage we are ready to outreach to the plant, and we are doing the job. But it takes time. So, signing the MOU, the memoranda agreement with STP took me eight months. Hopefully, the next plant we sign that agreement, it will be a shorter time.

CHAIRMAN STETKAR: Yes. Well, and more importantly, out in the industry, people who speak to one another -- you have STP onboard. In some sense, you need them to be a bit of a champion for you in terms of convincing their fellow owner/operators that this is a good thing, not necessarily a burden or yet one more thing that they need to do that they don't really want to do. I think without that cooperative industry champion, you may face a little bit of an uphill battle.

MR. PETERS: When you talk about the GoToMeetings, James, Harold/STP was an integral part of that presentation and have been supporting us in that. So, we are very excited about that.

MR. CHANG: Yes, that GoToMeeting with the STP trainer doing his software, that we GoToMeeting to see the software, that is all by GoToMeetings. That is from the users' perspective.

MR. PETERS: My big question as a Branch Chief is we don't have unlimited funding; we don't have unlimited travel budgets for all those items.

CHAIRMAN STETKAR: Sure.

MR. PETERS: And to make sure we can get some use for ourselves is that big key.

CHAIRMAN STETKAR: Yes.

MR. PETERS: How much should we fund if we can't find --

CHAIRMAN STETKAR: Sure. Absolutely.

MR. PETERS: -- substantive use for ourselves.

CHAIRMAN STETKAR: No, that's right. That's right. That's right, and I think that is a legitimate concern, actually.

And I am not as familiar with, unfortunately, training. I don't know, Harold, are you? You know, there are workshops. For example, the NEI Workshop on Fire Protection stuff is a really good forum for getting things out in the open. And I don't know whether the industry has similar types of gatherings of folks for training.

MEMBER RAY: Well, INPO in the Accreditation Program --

CHAIRMAN STETKAR: See, that is why I mentioned INPO.

MEMBER RAY: -- is the place to look for

something like that, not NEI.

MEMBER CORRADINI: Could you present this at -- well, the one thing that I was thinking of is the CONTE Conference.

CHAIRMAN STETKAR: Right.

MEMBER CORRADINI: And the other one is that the INPO Accreditation Board has an annual meeting where the CNOs are there, if you can get on that agenda and make a presentation there, particularly if South Texas would assist in the presentation.

CHAIRMAN STETKAR: Although CNO might be too high a level is the problem.

(Laughter.)

MEMBER CORRADINI: Yes. But, at some point, you are going to have --

CHAIRMAN STETKAR: Yes, but you need kind of a groundswell of support from --

MEMBER CORRADINI: But the CONTE Conference is the one that I was thinking about.

MEMBER RAY: Yes, I don't think the motivation by itself is a problem. It is the prioritization with the fact that almost everybody has got more to do than they have the resources to get done.

CHAIRMAN STETKAR: Yes. That is why this can't be seen as --

MEMBER RAY: It is a question of what is more important.

CHAIRMAN STETKAR: -- can't be seen as an increased burden. In fact, it has to be perceived as improved training efficiency or improved training effectiveness.

MEMBER RAY: Yes, and when is it going to pay off?

CHAIRMAN STETKAR: When is it going to pay off?

MEMBER RAY: And I want deliver.

CHAIRMAN STETKAR: That's right. That's right.

MEMBER RAY: And that is pretty hard to commit to.

CHAIRMAN STETKAR: Okay. We have lost Dennis. We are about to lose Sam.

What I would like to do is first ask if any of the other members have any further questions for James.

(No response.)

If not, thank you. You presented an awful lot of information, and we are just about on time.

So, I really appreciate that.

Before we kind of closeout, what I always have to do is ask if there are any comments from members of the public.

(No response.)

I haven't heard any.

I guess we ought to open up the bridge line. I am not sure who else is out there.

Before we all disappear, as we normally do, I would like to go around the table and see if any of the members have any final wrapup comments, and I would like to do two things. No. 1, make sure that we are on the same page about the next Subcommittee meeting, the topics, and I think we are. And the second is to explore among the members when we feel it might be appropriate to bring this general topic, not particularly the SACADA, but the status of the human reliability analysis methodology to a full Committee meeting, because it has been quite a while.

As I understand it, the bridge line is open. So, again, I have to beg somebody to just say something to make sure that it is.

(No response.)

Just anybody who is out there, just say something.

(No response.)

Now I have to trust the technology. I mean, I am not hearing pops. I will assume now that the bridge line is open.

Is there anyone out there who would like to make a comment or ask a question or anything?

(No response.)

Hearing nothing, I am confident now that we have worn down anybody from the outside world. We have done that.

So, now what I would like to do is, first, go around the table and ask each member if you have additional comments or items you would like to add.

And I will start with Steve.

MEMBER SCHULTZ: I would just like to congratulate all the presenters on the discussions that we have had today, both in terms of the analytical focus that we heard about this morning, and then amplified really strongly by the discussions on the empirical evaluations and projects that have built that. And then, this data collection technique is intriguing as well.

It certainly seems from all of that that there is a lot of work left yet to do in order to make this fiscal 2013 targets to come to be. And so, I encourage you to continue your focus related to that because you have got, obviously, a very good point at which the product is. In December, so much work seemed to have gone into the product line that was thrust upon us over the last 30 days, all of which seemed to have completion dates in December just before the holidays.

(Laughter.)

And with that, I am sure there has been a big sigh of relief. But there is, again, great benefit to be gained by the integration in this project over the next nine months. So, I would certainly encourage that.

And with regard to that, I am very interested in hearing the results of the workshops that are upcoming. I think the Subcommittee would benefit from that fairly shortly. As to when to bring it to the full Committee, I would wait for a general discussion on that, John.

CHAIRMAN STETKAR: Okay. Thanks.

MEMBER SCHULTZ: But, again, thank you very much. I appreciate it.

CHAIRMAN STETKAR: Harold?

MEMBER RAY: Well, John, I am in the mode of, as I think I have commented from time to time, of trying to figure out how to use this information in, I won't say practical ways, but in ways that are effective and realistic, given, as I say, the need to prioritize at the end of the day.

There are going to things that are being imposed on the industry I believe that are not going to be able to get done, given the resources available. And yet, we are relying more and more on human actions to respond to scenarios that we are considering or having to evaluate.

So, I think it is really important. The issue is, in my mind, well, okay, how are we going to use it? For example, are we going to be able to say, "No, that is not something we can simply shovel off onto operator action," if it should come to pass, and here's why? Or are we just gathering statistical data that will somehow get ground into a PRA analysis, and we never know what happens to it. It just comes out at the end, and nobody pays any attention to it anyway.

I think that there is a need for us to know what we can rely on human resources to do and what things we should not. And therefore, I think this is important.

CHAIRMAN STETKAR: Good. Thank you.

Sam?

MEMBER ARMIJO: I agree with what Harold and Steve said. I have nothing more.

CHAIRMAN STETKAR: Joy? You can say things here.

(Laughter.)

MEMBER REMPE: I still have questions about quantification. It is not my area of expertise, but I am from Missouri and I like to really believe things, you know. And I would be interested in seeing what happens with the pilot, the upcoming exercise.

And I second what Steve said about waiting and discussing that before we go to the full Committee.

CHAIRMAN STETKAR: Thank you.

Dr. Corradini?

MEMBER CORRADINI: I wanted particularly to come to see the comparison to the empirical data. So, I was very impressed, interested in that. And I guess I would like to see more.

Again, I think the quantification part of it is of interest to me. I do think, though, that since you encourage me, stimulate me, and persuade me to come to these, I am actually learning something for the HRA.

(Laughter.)

CHAIRMAN STETKAR: And we enjoy your presence so much.

(Laughter.)

MEMBER CORRADINI: I know that.

CHAIRMAN STETKAR: And because you have honored us with your presence, I have waited. Do you have anything to add, Dr. Powers?

MEMBER POWERS: I will explicitly not comment about the empirical studies because of an employer conflict of interest.

CHAIRMAN STETKAR: There is a lot of that going around in this Subcommittee, unfortunately.

(Laughter.)

MR. PETERS: And this is part of our strategy; we do employ everybody.

(Laughter.)

CHAIRMAN STETKAR: Yes.

MEMBER POWERS: I do have to say that, on this data collection, I am not persuaded that this is doable, and in two contexts. One, I think it is just an enormous amount of work, and it is not clear to me that everyone will interpret all the definitions identically. And consequently, I think you will get information that doesn't jibe well.

Okay. Now time is perfectly capable of proving me wrong on this, but I think the tool that will interrogate and inspect the database will, then, be of critical importance in getting anything useful out of it, because it is going to have to understand not what the designers of the database intended, but how the interpreters of data interpret it.

CHAIRMAN STETKAR: Right, right.

MEMBER POWERS: And I think that is a very big challenge.

CHAIRMAN STETKAR: That is another reason why I would like to see a few more people involved in it, to test sort of those things.

MR. PETERS: You mean a few of the utilities?

CHAIRMAN STETKAR: I mean, we had a lot of that experience 30 years ago, though just even filling out data forms for pumps and pipes and valves, that different people misinterpreted failure modes or what even was a failure. This is prone to the same -- of course, we have learned, hopefully, better ways to characterize it.

MEMBER POWERS: I mean, I am always just stunned at what seems to be a simple declarative sentence is subject to multiple interpretations.

CHAIRMAN STETKAR: Well, and especially if it has to do with some -- it is one of my concerns about making identity information available outside of the power plant, especially if it has to do with interpretation about how good I might be, not that piece of equipment over there, but me or my fellow brethren or me as a power plant related to everyone else.

MEMBER POWERS: That brings up another issue that I think the agency wants to give some thought to, is the security of the database. Bob Pollard made quite a lot of mileage out of a couple of sentences concerning deficiencies of operators. And a hacker getting access to this particular database, it might give you a chance to see just how comprehensive The Times and The Post can be in exploiting information.

MEMBER RAY: Well, that even raises the issue of what is proprietary data. For example, is this proprietary to somebody? If so, who? Or is it subject to FOIA, and so on?

I mean, my God, you see the registration of handguns in New York State, for example, in the newspapers.

MEMBER POWERS: Well, I mean, that

information, by design, was public. But I think you need to think a lot about the security on this, the information security of this database, because I think maybe you don't want to see this blasted across The New York Times.

MEMBER ARMIJO: Now I think about it, one thing I wanted to mention. At least it struck me that this is the kind of data that, if the trainers are doing their work well, it is going to change with time, as people -- and so these error probabilities, as people learn where the deficiencies are in their training and their response to the events, are going to learn and get better.

So, it may be that the value is right at the very beginning, not the use of a big database for PRA at some later stage. I don't know. It just seems to me that the biggest value is in the initial work done by the training staff.

MEMBER RAY: Trainers train to detest, Sam, first and foremost, period.

MEMBER ARMIJO: Yes.

MEMBER RAY: The pass/fail rate on licensed operator examines is it, as far as they are concerned, at lest from my experience.

MEMBER ARMIJO: Yes. But the trainer, you know, if he goes into a little more detail, may be able to do a much better job.

MEMBER RAY: I agree with you, but I am just telling you it is how --

MEMBER ARMIJO: Agree, you have got to pass first.

MEMBER RAY: Pass/fail is it, man. And so, really, I think the implication of what you are saying is, are the tests really challenging us to improve in this area or not? In my opinion, the answer is not yet, not at least the test I ever saw. They are very much more simplistic.

CHAIRMAN STETKAR: Anyway, since we are apparently losing Dr. Corradini --

MEMBER CORRADINI: I have got to go to a conference call.

(Laughter.)

CHAIRMAN STETKAR: What I would like to do is let me just sum up my comments.

I would like to thank everybody who has left. I think that we covered a lot of information in a one-day meeting. I was a bit concerned that we not be able to get through it.

I think that, from the staff's perspective, the status of the methodology development, as it is now compared to a year ago, is we are well up on the learning curve. Now you hear things like we have a long ways to go by the end of fiscal year 2013. There is a lot of ground that has been covered in the last year, especially when I look at the amount of ground that was covered from 2006 until the end of 2012. The last year you have made very, very good progress. So, I think you deserve to be congratulated about that.

Regarding, since we are losing interest and people, regarding Subcommittee meetings and full Committee presentation, we should probably handle that offline. I do think that the full Committee should hear about this. The timing of when we do that regarding near-term activities is something we will need to work out with the staff and find out when the opportune time is.

I am confident that the Committee will want to write a letter regarding the status because we just haven't weighed-in as a Committee on it.

The next Subcommittee meeting, I would like to get sometime in the next three months or so, if that fits your schedule.

We talked about the cognitive basis document, and perhaps a little bit more information about the general methodology and how it relates to that. So, I think we want to target that kind of timeframe. Again, John and the staff can work that out.

With that, thanks again to everyone, and we are adjourned.

(Whereupon, at 5:20 p.m., the meeting was adjourned.)

Overview of Staff's response to the SRM "HRA Method Differences"

Jing Xing

NRC/RES/DRA/HFRB

Prepared for 1/16/2013 ACRS meeting

A Collaboration of U.S. NRC Office of Nuclear Regulatory Research (RES) & Electric Power Research Institute (EPRI)

Contributors

US NRC - Erasmia Lois, Jing Xing, James Chang, Song-Hua Shen, Nathan Siu

EPRI - Gareth Parry (ERIN), Mary Presley, Stuart Lewis

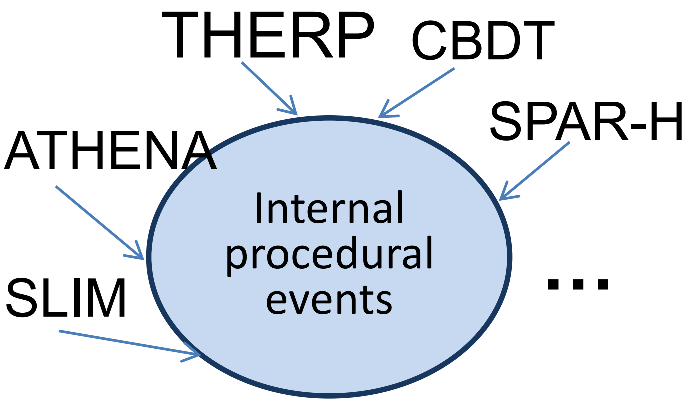
SNL - Stacey Hendrickson, Harry Liao, Susan Stevens-Adams, Katrina Groth

INL - John Forester, April Whaley, Martin Sattison, Ronald Boring, Jeffery Joe, Johanna Oxstrand, Dana Kelly

Paul Scherrer Institute - Vinh Dang

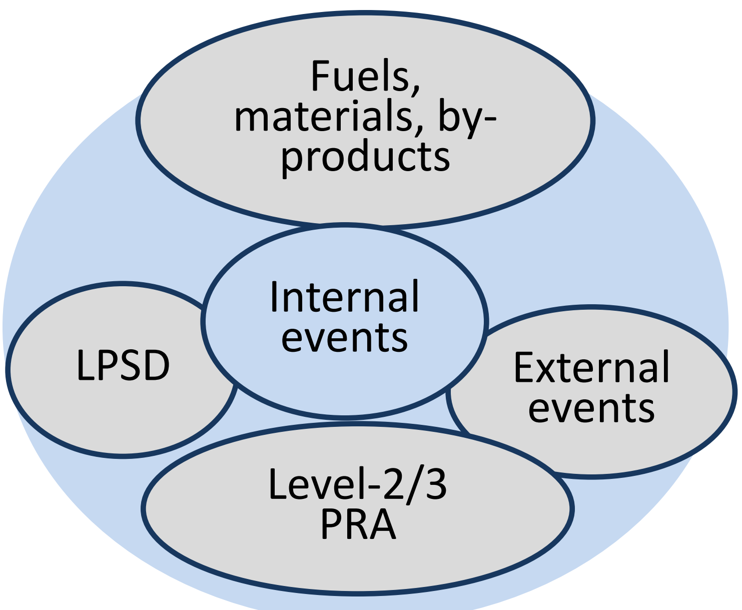
University of Maryland - Ali Mosleh

Research issues and needs



- Method-to-method variability
- Analyst-to-analyst variability

Need an integrated method to reduce variability



- Are the existing methods applicable and adequate?
- Does each application have to have its own method?

Need a generic methodology for all the applications

Research goal and requirements

Goal –

Develop a new HRA methodology to reduce variability and apply to all HRA applications.

Requirements –

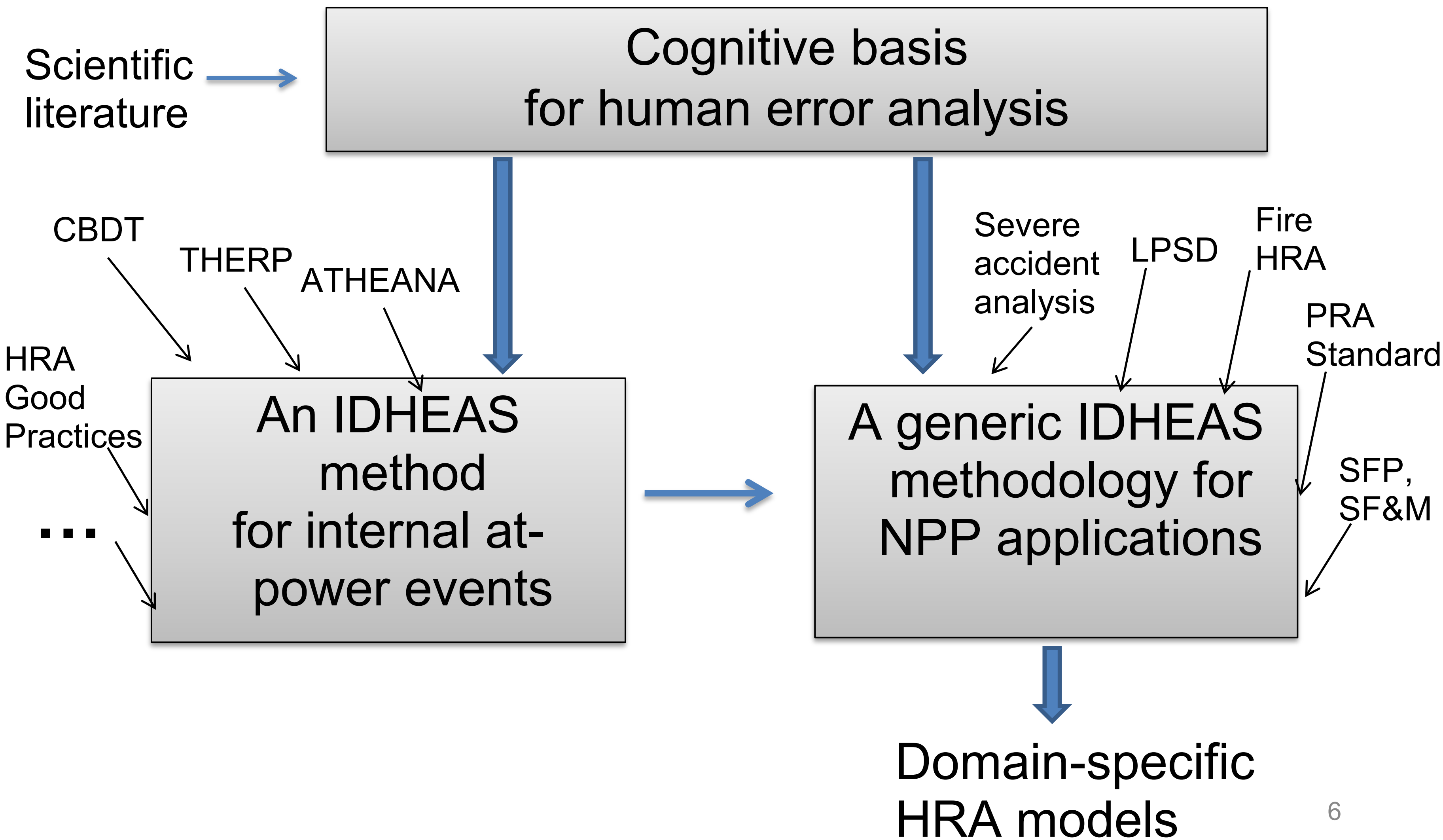
- Conform to the PRA/HRA standard and HRA Good Practices
- Retain and integrate the strengths of existing methods
- Have enhanced capabilities to address the key weaknesses in state-of-practices.
- Have a state-of-art technical basis and be generic enough for all HRA applications in NPPs

Main lessons learned from HRA benchmarking studies

1) Each method evaluated has its own strengths;
2) Most the methods do not have an explicit cognitive basis on why and how human fails to perform tasks, and all the methods need a stronger basis;
3) The methods either lack adequate guidance for performing qualitative analysis or lack an adequate interface for using qualitative analysis results for quantifying human error probabilities (HEPs);
4) The methods have inadequate guidance on how to assess and use performance influencing factors (PIFs).

IDHEAS takes the advantages of 1) and improves 2), 3), 4).

Strategic approach



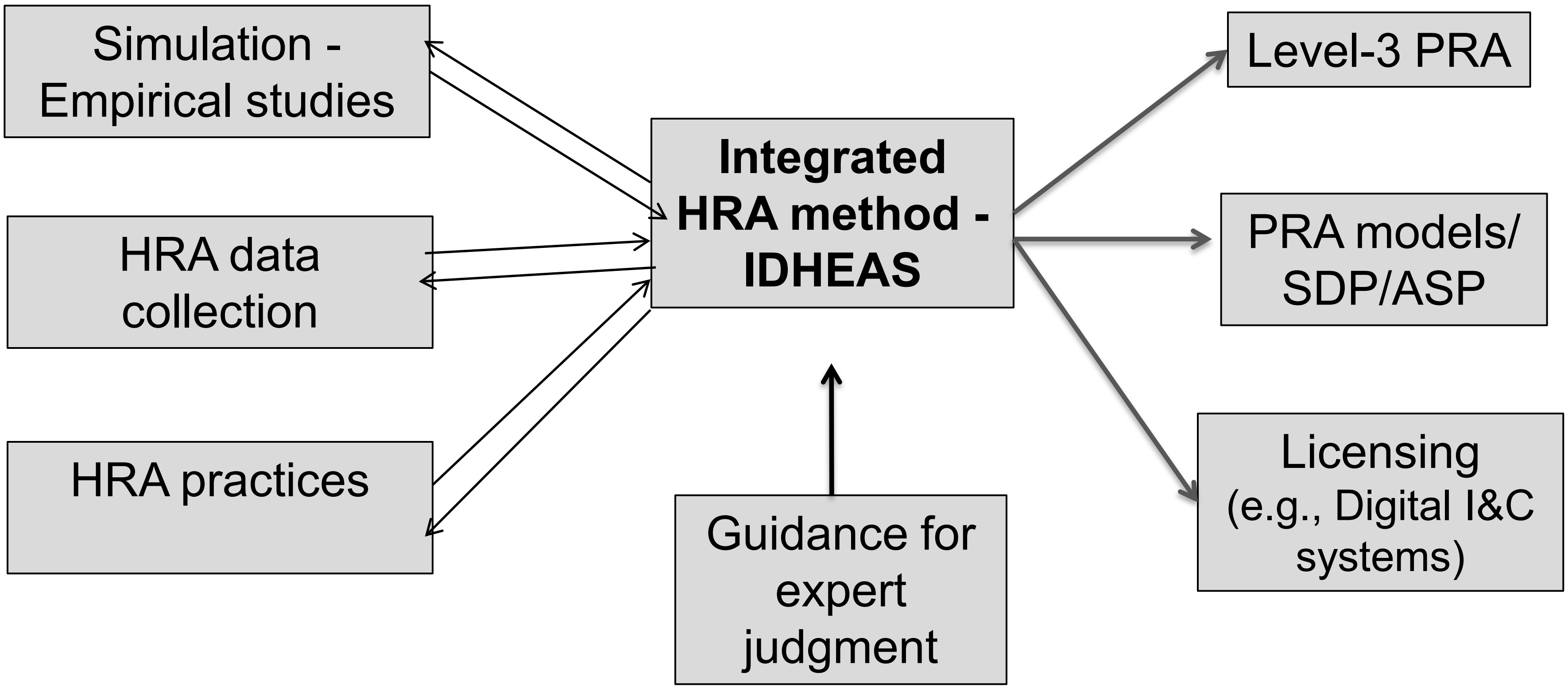
IDHEAS products

Product	Intended applications
Cognitive basis for human error analysis	• HRA • Human performance • Human factors engineering
Generic IDHEAS methodology for NPP applications	• HRA for all kinds of human events in NPP (Level-3 PRA, LPSD, external events, etc.)
An IDHEAS method for internal, at-power events	• Internal, at-power event PRA (PRA models, SDP, ASP, etc.)

IDHEAS Status and planning

Product	Status	FY13 Deliverables
Cognitive basis for human error analysis	Completed	Final report
Generic IDHEAS methodology for NPP applications	•Draft report of the prototype •To be explored in Level-3 PRA	Draft report for formal testing
IDHEAS method for internal, at-power events	• HEP elicitation • Test and validation	Final report

Overall HRA strategy & Path forward



EPRI Role in IDHEAS HRA Methodology Development

ACRS Subcommittee on Risk and Reliability

January 16, 2013

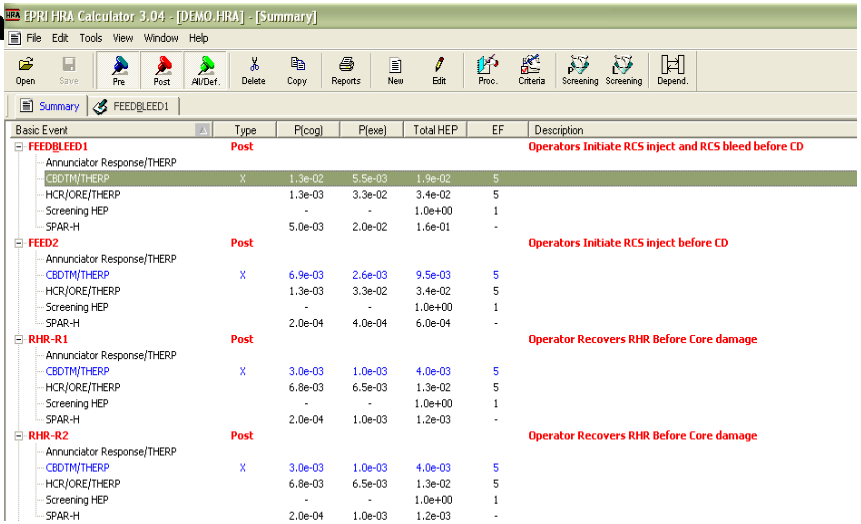
Mary Presley (EPRI)
Gareth Parry (ERIN)

A Collaboration of U.S. NRC Office of Nuclear Regulatory Research (RES) & Electric Power Research Institute (EPRI)

EPRI HRA User's Group

- Mission Statement

1. Develop a tool to enabling different analysts to obtain comparable results for same action & method at similar plants.
2. To develop guidelines for application of HRA methods.
3. Key goal is to enable industry to converge on common methods.
4. Coordinate with industry groups such as USNRC, Owners Groups, & within EPRI to develop guidelines and train



- Membership:
  - All US utilities
  - Several of international and corporate members

EPRI Involvement

- Work to-date has been performed collaboratively between NRC and EPRI under an MOU.
  - EPRI focus on internal events HRA document

- EPRI Motivation:
  - Current methods have not been substantively updated in more than 20 years
  - Lessons learned from Empirical Studies, particularly on importance of qualitative analysis and reproducibility need to be addressed by new

Path Forward

- Expert Elicitation
- EPRI Method Testing
  - Meets ASME/ANS Standard
  - Clear guidance and appropriate examples (qualitative + quantitative analysis)
  - Appropriately time intensive
  - Produces reasonable HEPs
  - Enhances reproducibility
  - Test against variety of IE HFEs
    - Time critical v. non-time critical
    - HFEs traditionally handled well vs. not handled well with existing methods
- Recommendation for use based on testing
- Incorporation into the HRA Calculator [TBD]
- Adapt methods beyond IE HRA [TBD].

Together…Shaping the Future of Electricity

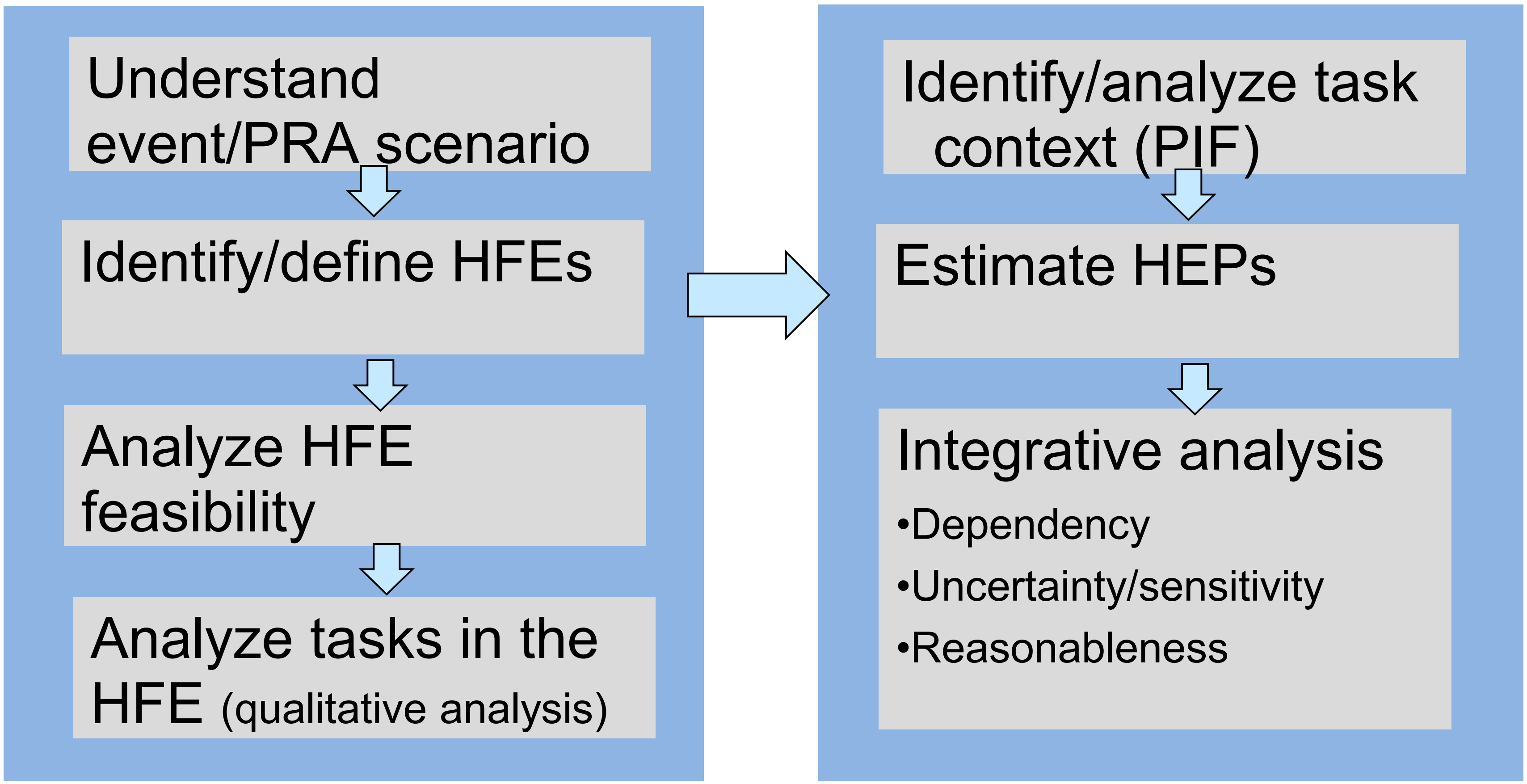
Additional information / backup slides

HRA application areas

Dimension	Specifics
Plant Mode	• At-power • Low power and shutdown
Event Type	• Internal • Spatial (e.g., fire, flood, and seismic)
PRA Phases	• Level 1 • Levels 2 & 3
Radiation Source	• Reactor • Spent fuel pool • Dry cast storage
New/Existing Reactors	• Existing reactors • New & advanced reactors
Temporal Phase	• Pre-initiator • Initiator • Post-initiator
Actor	• Control room • Control room and local combination • Local
Risk-Informed Program	• SPAR • ASP • SDP (of RASP)
Level of analysis	• Detailed • Bounding (screen and scoping)

HRA process

HRA process is defined in the PRA standard and recommended in HRA Good Practices.



Key features of the products

Product	Key features
Cognitive basis for human error analysis	• Cognitive functions supporting NPP tasks • Cognitive mechanisms for reliably performing tasks • Context characters challenging the mechanisms
Generic IDHEAS methodology for NPP applications	• HFE identification / definition / feasibility analysis • Key task identification / representation / analysis • Quantification process (scoping analysis and detailed failure-mode analysis)
IDHEAS method for internal at-power events	• HFE definition & feasibility analysis • Crew response tree for qualitative analysis • Quantification model (crew failure modes ,decision-trees, and HEPs for the tree branches)

IDHEAS status & planning (details)

Product	Status	Plan for FY13	Long-term path -forward
Cognitive basis for human error analysis	•Completed •Finalized the Lit review report	Publish the final report	Update the cognitive basis
Generic IDHEAS methodology for NPP applications	•Completed the prototype; •Developed the draft report	•Test the prototype in Level-2 PRA •Publish the final methodology report	•Validate and calibrate the methodology •Develop HRA models for specific applications
IDHEAS method specific for internal at-power events	•Completed the method development •Delivered the draft report	•Conduct expert elicitation of HEPs •Test the method •Publish the final report	•Roll out to HRA applications •Data-referenced HEP estimation •Improve usability

Expert elicitation of HEPs

Experts: DE – Data experts, RE- Resource experts, TI – Technical integrators

Preparation

DE identify & compile data

PM prepare procedures & worksheets

Training & piloting

Workshop #1

DE present model & data

↓

PE rank DT branches & assess PIFs

↓

TI question data and PE's judgment

Workshop #2

TI estimates HEPs for selected branches

↓

PE question TI's estimation

↓

TI revise & integrate HEPs

IDHEAS validation plan

Validation method	Scheme	Validation scope	Pre-Condition	When (tentative)	Who
HRA test battery	Develop a task/event/scenario battery for HRA method validation/testing			FY12	INL
Scientific validation	1) Team and peers to fill out the questionnaire 2) External review of the reports	Integrated method	IDHEAS method development completed and Report II & III draft ready	FY13	INL
Content validation	1) Team develops content statement/description and peers to fill out questionnaire 2) Check compliance to HRA standard/Good Practices	Integrated method	IDHEAS method development completed and Report II & III draft ready	FY 12-13	Team
Demonstration of working (testing)	Apply the parts and integrated method to selected events/scenarios	All the parts and integrated method	As soon as individual parts are ready for piloting	FY12-14	Self-piloting, EPRI, and NRC staff

IDHEAS validation plan (continued)

Validation method	Scheme	Validation scope	Pre-Condition	When (tentative)	Who
Item validation 1	Perform confirmatory factor analysis to load plant/task/human characteristics to CFM/DTs	CFMs, DT, HEPs	CFMs and DTs are ready for testing; Expert elicitation for HEPs completed	FY13	SNL
Item validation 2	Confirm/modify/calibrate CFMs/DTs/HEPs with the data in the NRC's Simulator Data for HRA	CFM, DT, and HEP	Simulator Data is available	Outside IDHEAS project, Long-term activity	NRC
Benchmarking	Benchmarking the completed method using Halden or US empirical study settings	Integrated method	IDHEAS is completed and the above four validation approaches are completed.	FY13	NRC initiates a new task.
Experimental testing	Experimentally tests the effects of task and PIF characteristics on human performance	DTs and HEPs	1) NRC simulator is ready or 2) agreement with Halden	Outside IDHEAS project, Long-term activity	NRC

The methodology of an Integrated Decision-tree Human Event Analysis System (IDHEAS) – A generic HRA methodology for NPP applications

Jing Xing, James Chang
RES/DRA/HFRB


PULL
MIDVALE
SCHOOL FOR
THE GIFTED

IDHEAS products

Product	Intended applications
Cognitive basis for human error analysis	• HRA • Human performance • Human factors engineering
Generic IDHEAS methodology for NPP applications	• **HRA for all kinds of human events in NPP (Level-3 PRA, LPSD, external events, etc.)**
An IDHEAS method for internal, at-power events	• Internal, at-power event PRA (PRA models, SDP, ASP, etc.)

Research goal and requirements

Goal –

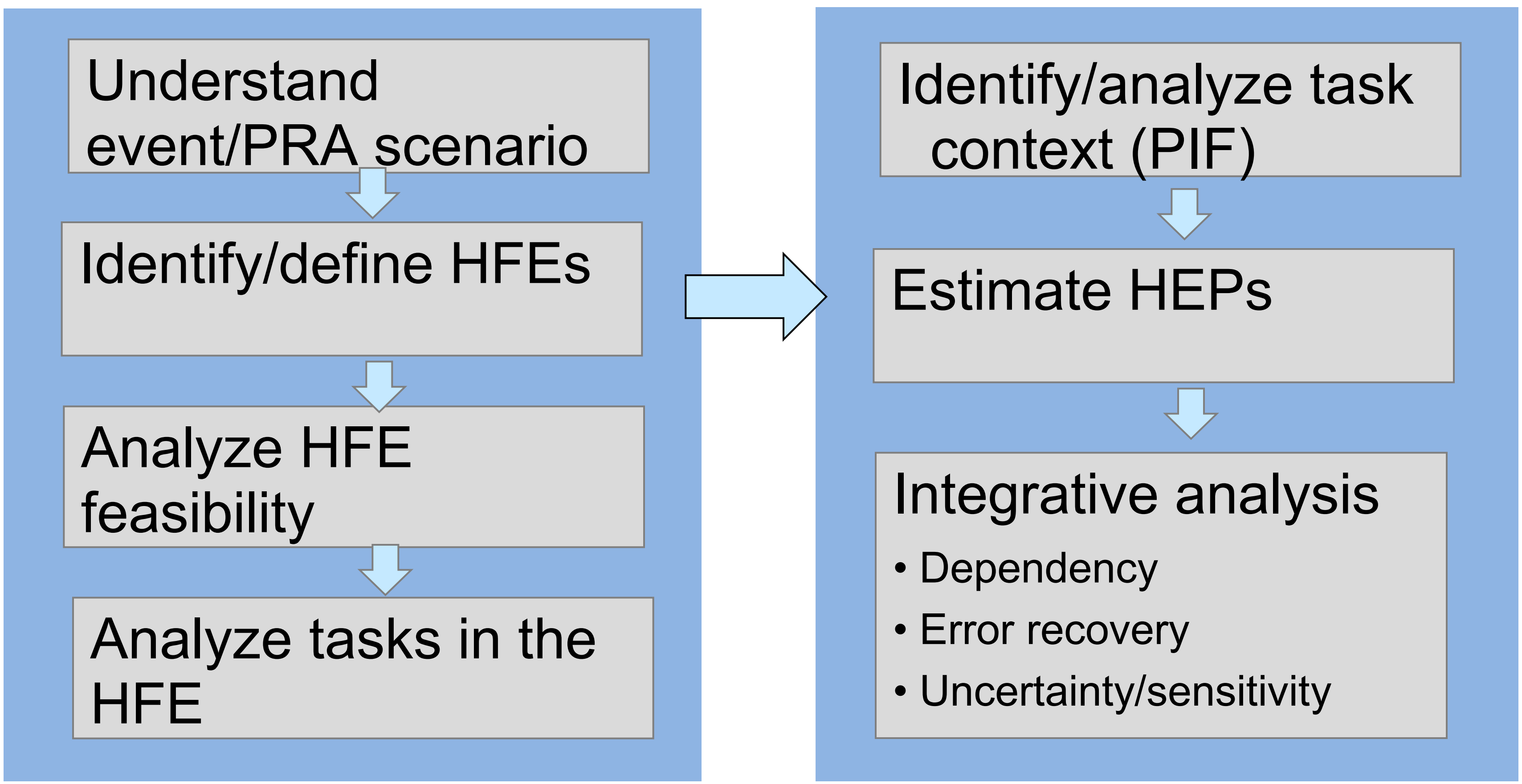
Develop a new HRA methodology applicable to all HRA domains in NPP operation.

Requirements –

- Generic for all HRA applications with state-of-the-art technical basis.
- Conform to the ASME PRA/HRA standard and HRA Good Practices
- Retain and integrate the strengths of existing HRA methods
- Enhance capabilities to address the key weaknesses in state-of-practices.

HRA process

HRA process is defined in the PRA standards and recommended in HRA Good Practices.



IDHEAS methodology

- IDHEAS methodology
  - Guidance for HFE analysis
  - Cognitive error-causal tree
  - Qualitative Analysis Structure
  - Quantification process
  - Integrative analysis

Guidance for HFE analysis

–adapted from HRA Good Practices, PRA standards, Fire HRA, and others

- Guidance for HFE analysis
  - Scenario understanding
  - HFE identification & definition
  - HFE feasibility analysis
- Cognitive error-causal tree
- Qualitative Analysis Structure
- Quantification process
- Integrative analysis

HFE identification

- NUREG-1792 (HRA Good Practices) and PRA Standards provides guidance for general process and considerations of HFE identification.

- NUREG 1921 (Fire HRA) provides detailed guidance for identifying the following three types of post-initiator actions:

- Internal event operator actions
- Operator actions outside of internal events
- Undesired operator responses to spurious alarms, indications, and digital I&C failures

HFE definition

- From HRA Good Practices, PRA standards, and Fire HRA

An HFE should be defined to represent the impact of the human failures at the function, system, train, or component level as appropriate. The definition should include the following:

- Accident sequences, initiating event, and subsequent system and operator action successes and failures preceding the HFE
- Accident sequence-specific procedural guidance
- The cues and other indications for detection and evaluation
- Accident sequence-specific timing of cues and the available time for successful completion
- The available time for action
- The high-level tasks required to achieve the goal of the HFE
- The undesired failure consequences and the likely situations for the failures to occur

HFE feasibility analysis

Feasibility analysis is to assess whether an HFE is feasible.

NUREG-1852 provides guidance for conducting a thorough feasibility assessment of manual actions. It identified the following feasibility criteria:

- Sufficient time to complete the tasks
- Sufficient manpower
- Cues available
- Adequate procedures and training
- Accessible location
- Availability of equipment required for critical tasks
- Operable relevant components

IDHEAS guidance on estimating performance time

Time estimation model in NUREG-1852:

Time Margin = (Available Time- (Cognition Time)
- (Execution Time)

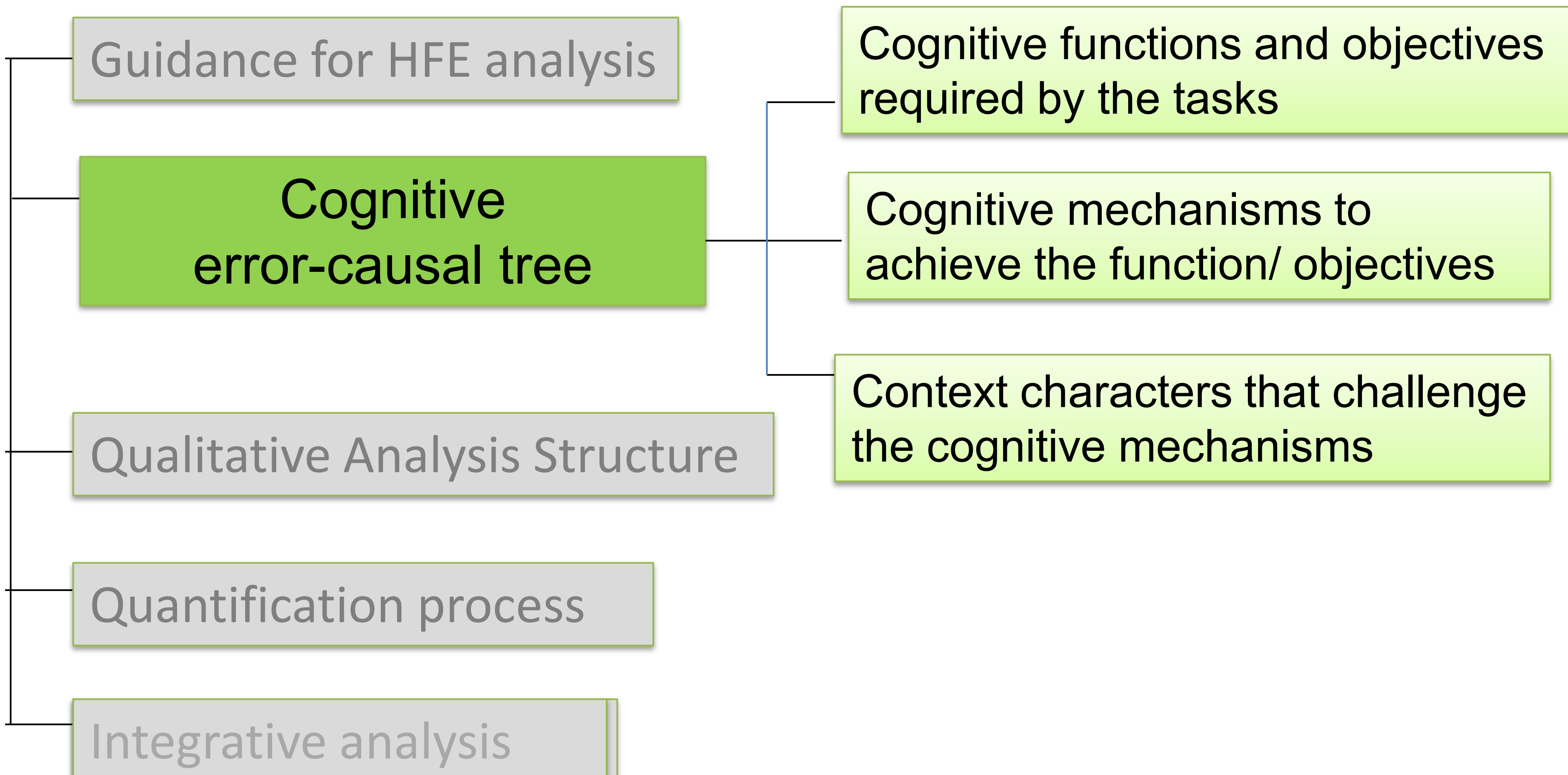
IDHEAS guidance:

Estimating the cognition and execution time is based on the following three sets of factors:

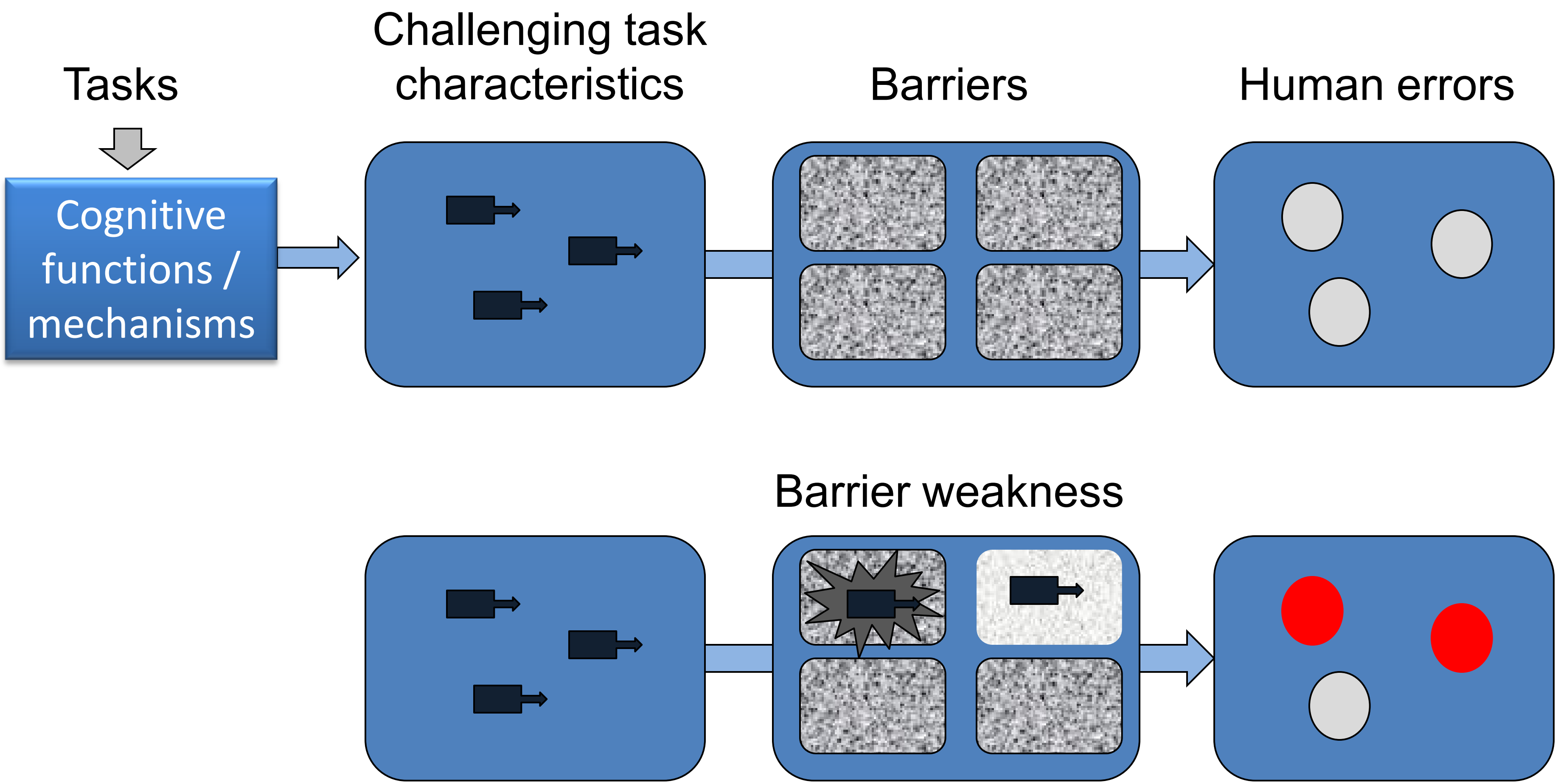
- Contributing factors to estimate time needed
- Modification factors to estimate time range
- Bias factors to calibrate the estimation

Cognitive error-causal tree

– Synthesized from the lit review report, research in decision-making, NPP task analysis and event reports, and HRA practices



Why does an experienced operator fail to respond properly?



Cognitive basis for human error analysis

The cognitive basis is to elucidate the following:

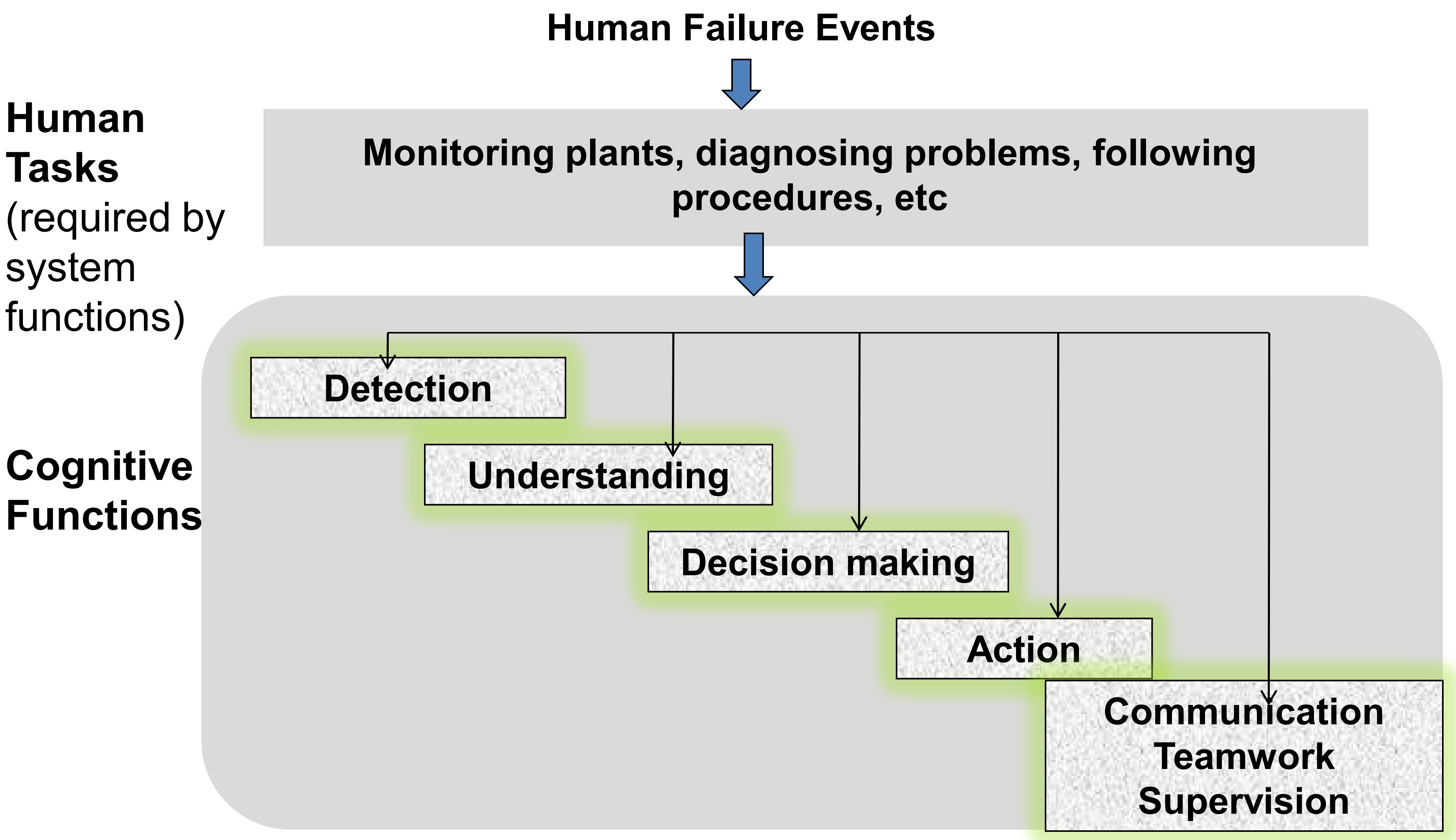
I. Cognitive Functions and objectives
   - How humans perform a cognitive task?

II. Cognitive Mechanisms
   - What makes humans reliably achieve a cognitive function?

III. Error Causes
   - How a cognitive mechanism fails?

IV. Challenging Context Characters
   - What contextual characteristics leads to errors?

Cognitive functions and objectives
- How humans perform a cognitive task?



Objectives of cognitive functions

Objectives of a cognitive function are the types of cognitive subtasks to achieve the goal of the function.

Objectives were identified by classifying human activities required by NPP system functions into generic cognitive tasks (studied in the literature).

Example objectives for *Understanding*

- Assess and verify information
- Develop coherent understanding of the information
- Maintain situational awareness
- Make predictions and expectations for the upcoming situation
- Diagnose problems

Cognitive Mechanisms
- What makes human achieve a function reliably

Example mechanisms for *Understanding*

- Iteration of information assessment, verification, and selection

- Selection of the mental model for the situation

- Integration of selected mental model with information

- Assessment and inhibition of beliefs

Error causes –
How a cognitive mechanism fails?

Examples for the *Understanding* function -

Cognitive mechanism: Selection of the mental model for the situation

Error causes:

- Incorrect mental model selected for the situation
- No mental model matches to the situation
- Mental model is not adjusted to situation evolution

Cognitive mechanism: Inhibition of belief

Error causes:

- Personnel belief is not evaluated against the existing information
- Improper belief is not inhibited.

Challenging Context Characters

What contextual characteristics leads to errors?

Context factor ⊗ Error cause ⇨ Challenging context character

Context factors (PIFs):

Workload and task demands –

demanding cognitive resources, challenging cognitive mechanisms, and leading to errors.

HSI/environment and procedures –

Aggravating the cognitive demands

Training, work process, and organizational factors –

Militating the demands and providing barriers to error causes, recovering errors

Challenging Context Characters
- What characteristics leads to error causes

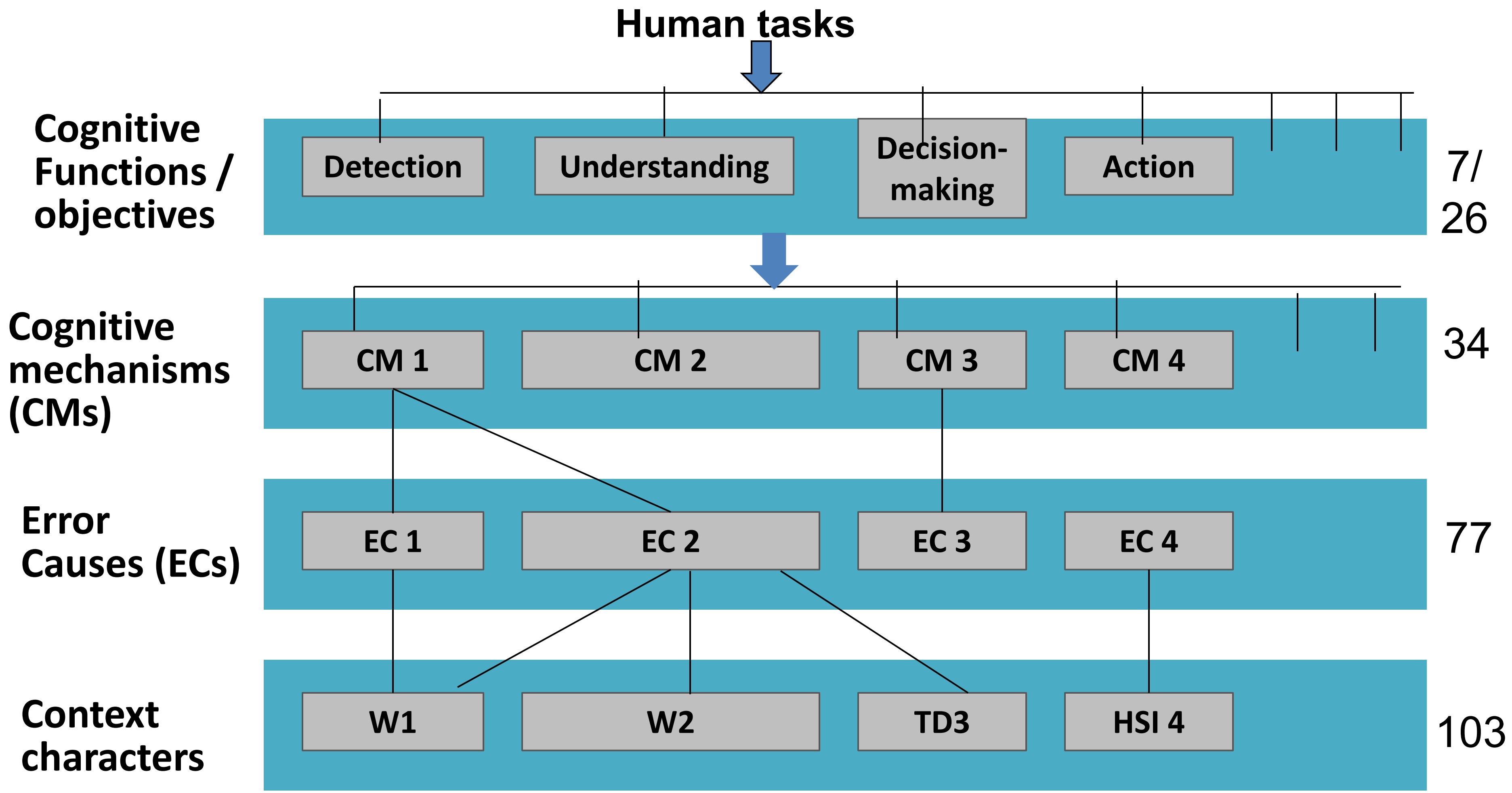
Context character list - For each context factor, we identified the context characters that challenge the cognitive mechanisms of every cognitive functions.

Example context characters for *Understanding*

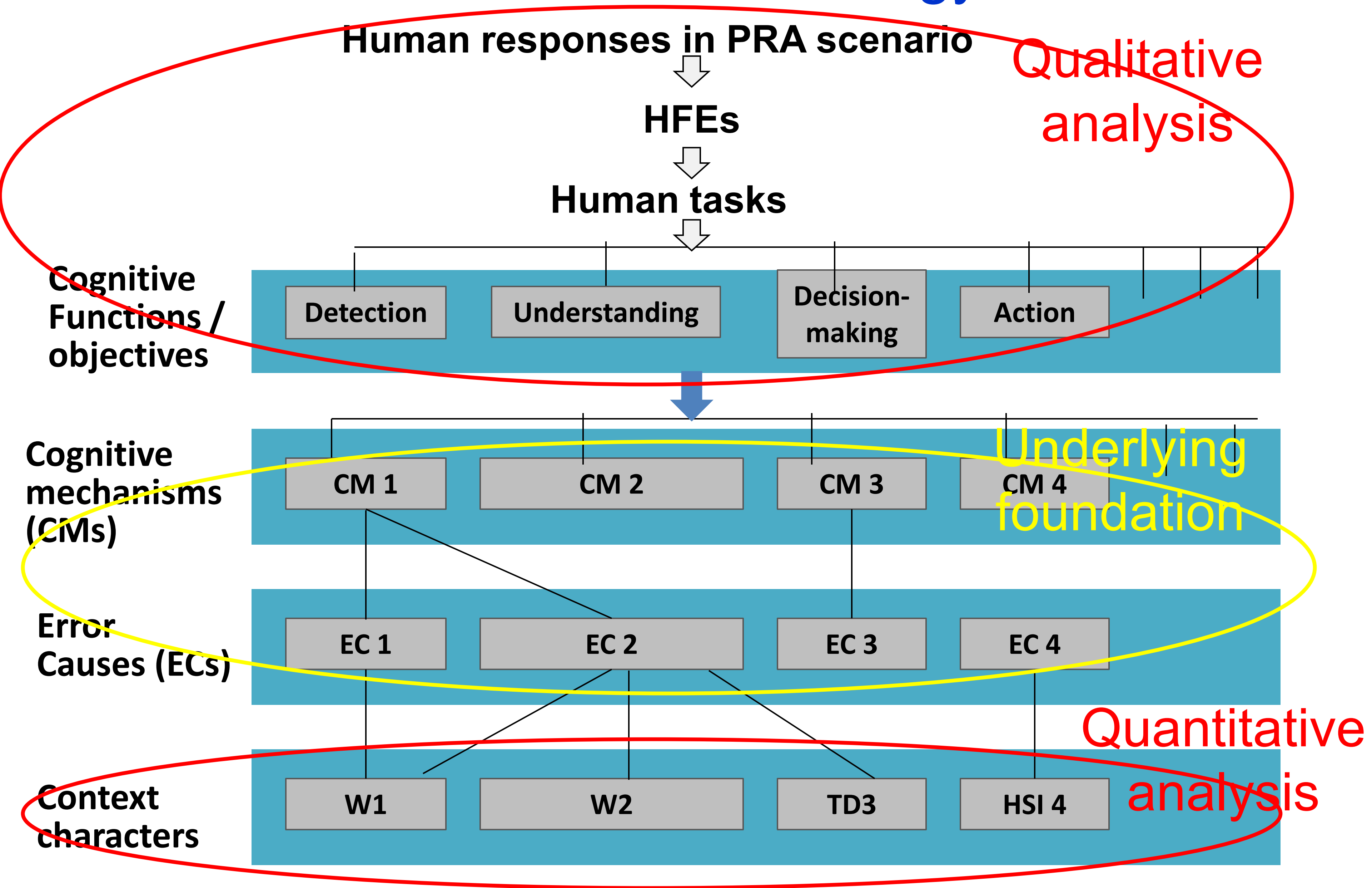
Context factor	Example challenging context character	Cognitive mechanism
Workload	Multitasking, Interruption	integration
Task demands	Unfamiliar scenario	Mental model
HSI	System behavior is not apparent or masked	Information selection
Procedure	Criteria are ambiguous	Integration
Training	Under-trained system failure modes	Mental model

Summary of the cognitive error-causal tree

Each cognitive function and its objectives are associated with cognitive mechanisms, error causes, and challenging context characters.



From cognitive error-causal tree to HRA – IDHEAS methodology



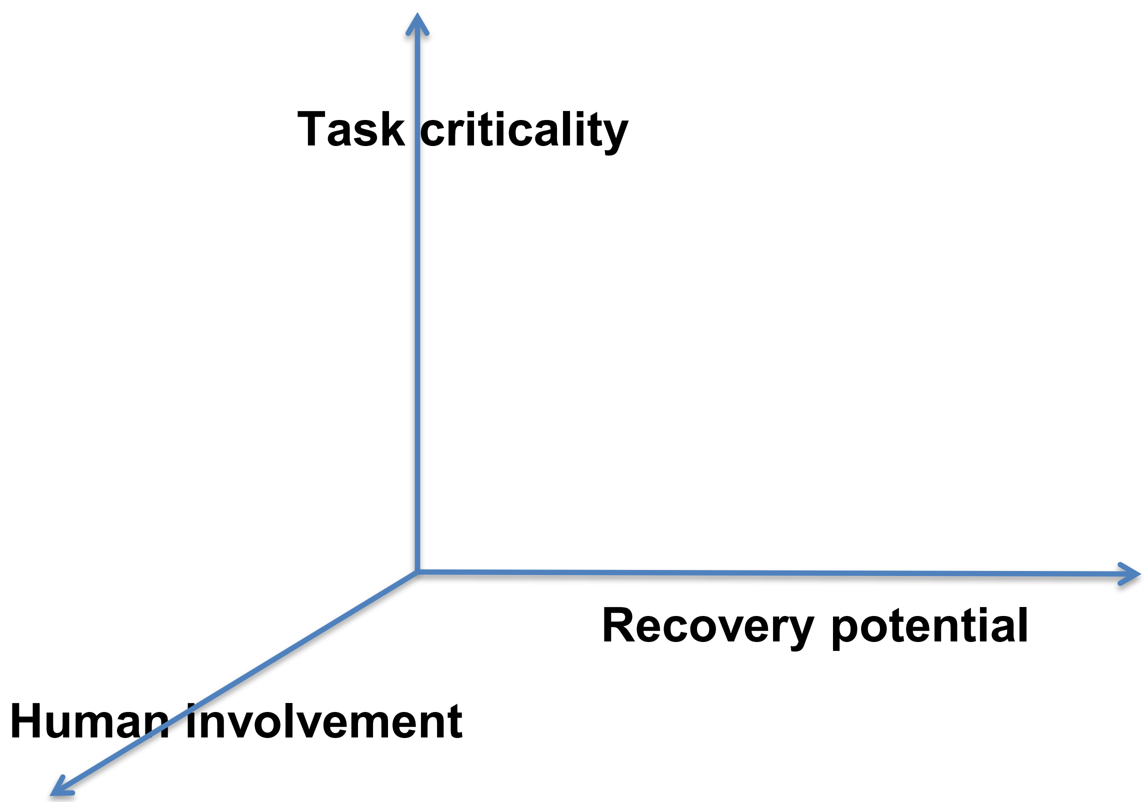
Qualitative Analysis Structure

- Guidance for HFE analysis
- Cognitive error-causal tree
- **Qualitative analysis structure**
  - Identify the tasks critical to the HFE success
  - Characterize cognitive aspects of each critical task
  - Perform timing and workload analysis
  - Develop the operational story
- Quantification process
- Integrative analysis

Qualitative analysis structure – Part 1: Identify the tasks critical to the HFE success

Objective –

Identify and represent safety-critical tasks for quantification; failing each critical task leads to failure of the HFE.



Qualitative analysis structure - Part 2: Characterize cognitive aspects of the critical tasks

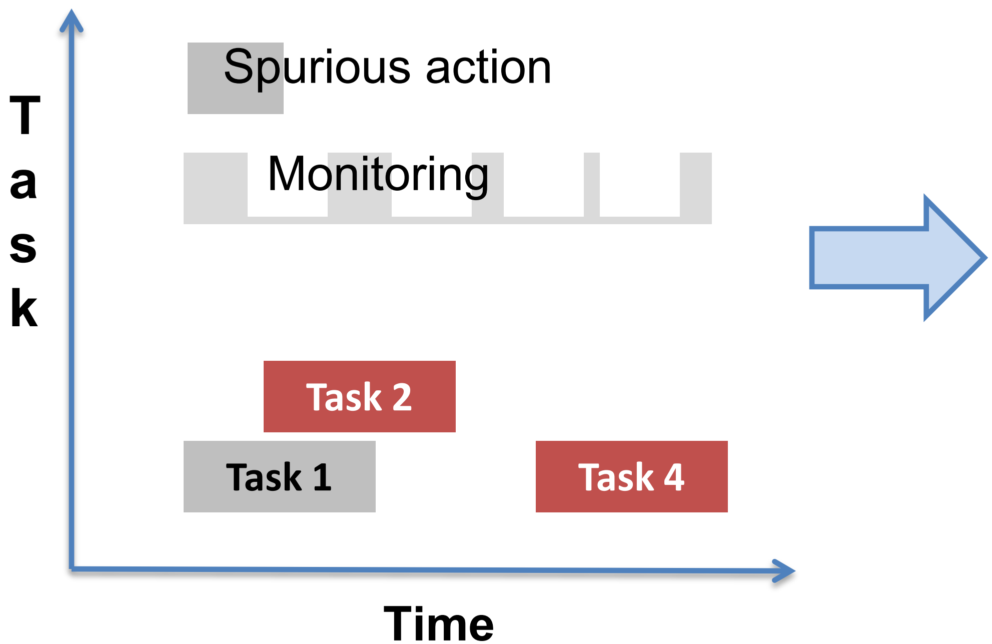
Objective - Identify cognitive characteristics of every critical task.

Cognitive features	**Description**
Task goal	The expected outcome of the task (e.g., reach hot shutdown within 3 hours) including the constraints of operation (e.g., cooldown RCS but not exceeding 100 ºF/hr)
Cognitive functions and objectives	Activities to achieve the goal and the desired outcome of the activities
Plant cues and supporting information	The information (i.e., cue) to initiate the task. A cue could be an alarm, an indication, a procedure instruction or others (e.g. onsite report). The supporting information is in addition to the cue and is needed to perform the task.
Procedures and operational guidance	Guidance used to perform the tasks.
Personnel	Personnel who performs the task or specific task objectives.

Qualitative analysis structure - Part 3: Perform timing and workload analysis

Objective – Assess workload

Timing



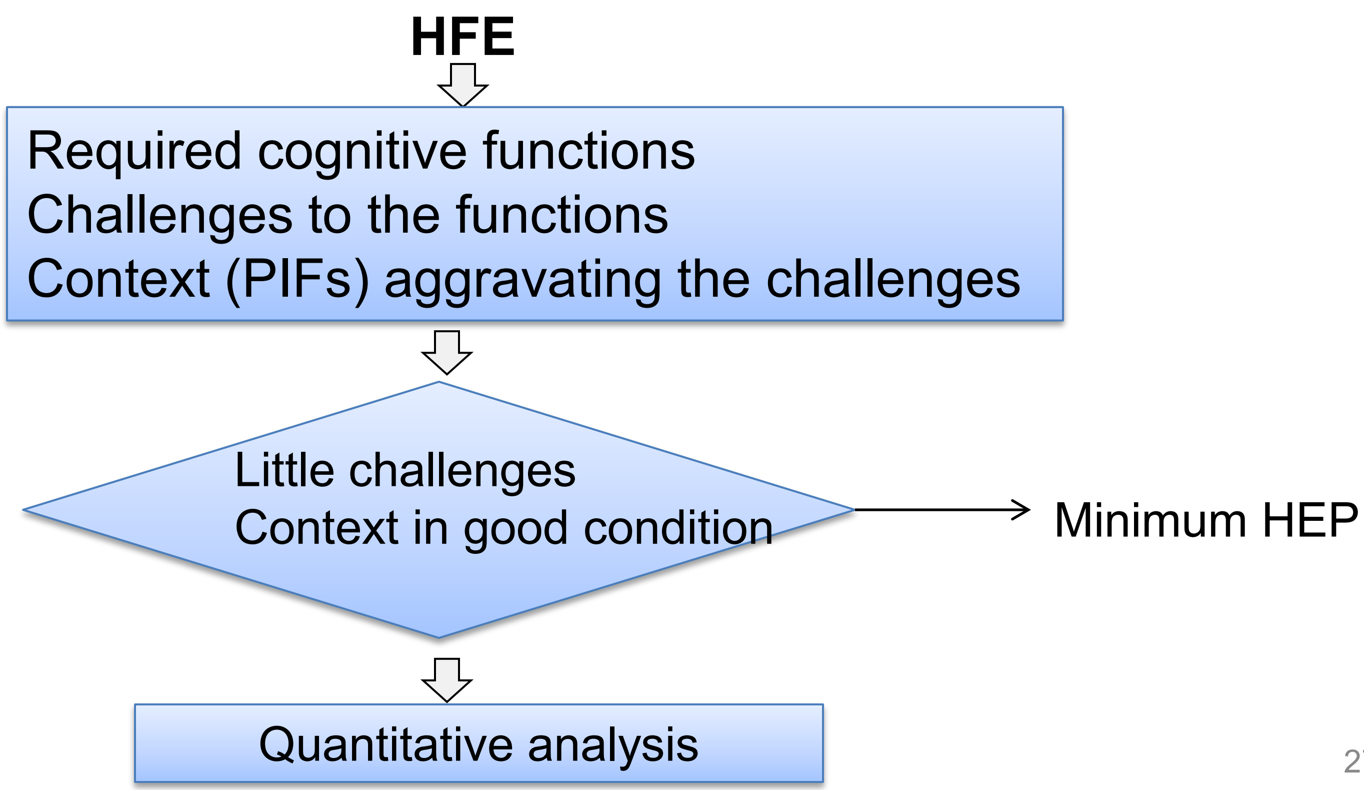
Workload characters

- W1 - Multitasking interference
- W2 - Interruption / distraction
- W3 - Complex, sustained cognitive demand
- W4 - Timing

Qualitative analysis structure
- Develop the operational story

Objective -

1) Develop the operational story to have a coherent understanding of the HFE
2) Screen out very low probability HFEs (little or no challenges).

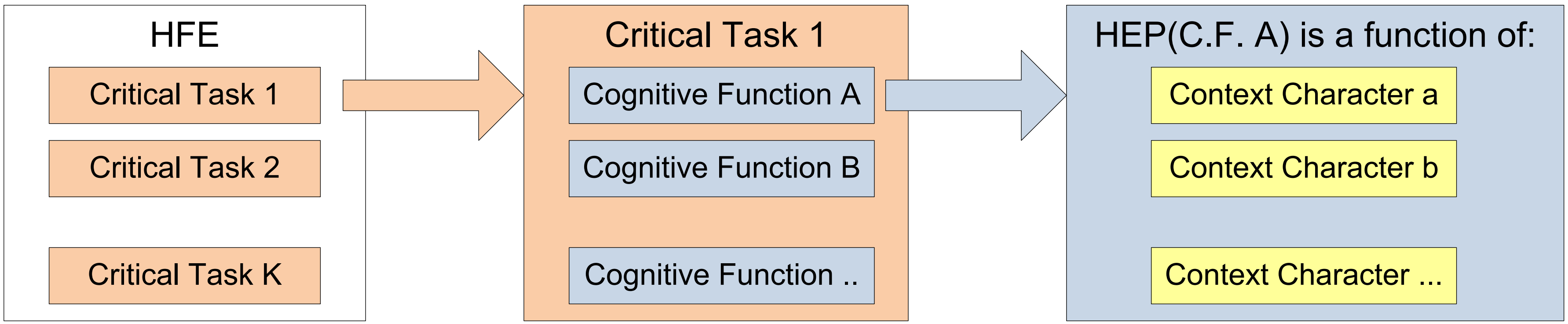


Quantification process

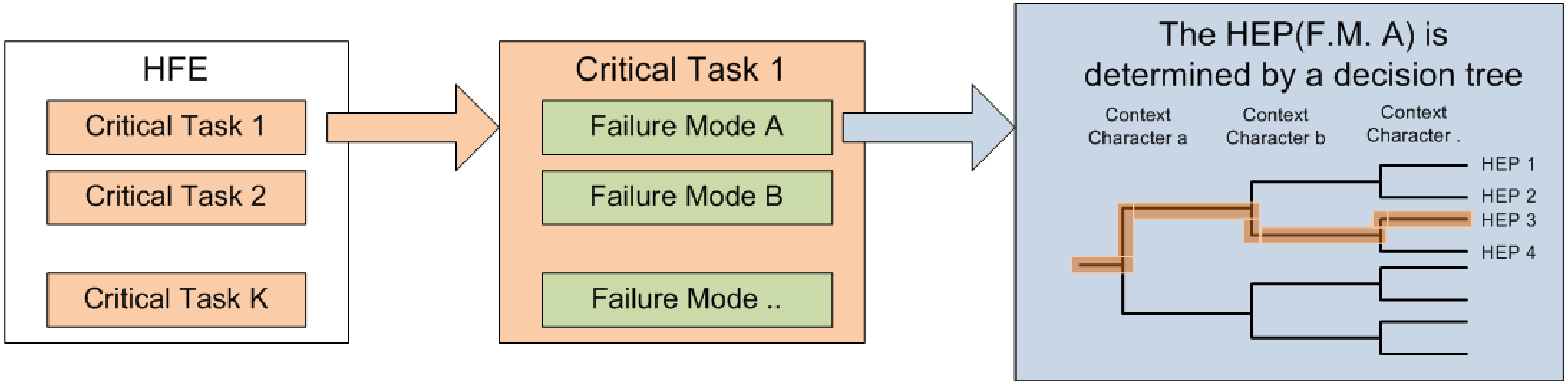
- Guidance for HFE analysis
- Cognitive error-causal tree
- Qualitative Analysis Structure
- **Quantification process**
  - Scoping analysis
  - Detailed failure-mode analysis
  - HEP estimation
- Integrative analysis

Two levels of quantitative analysis

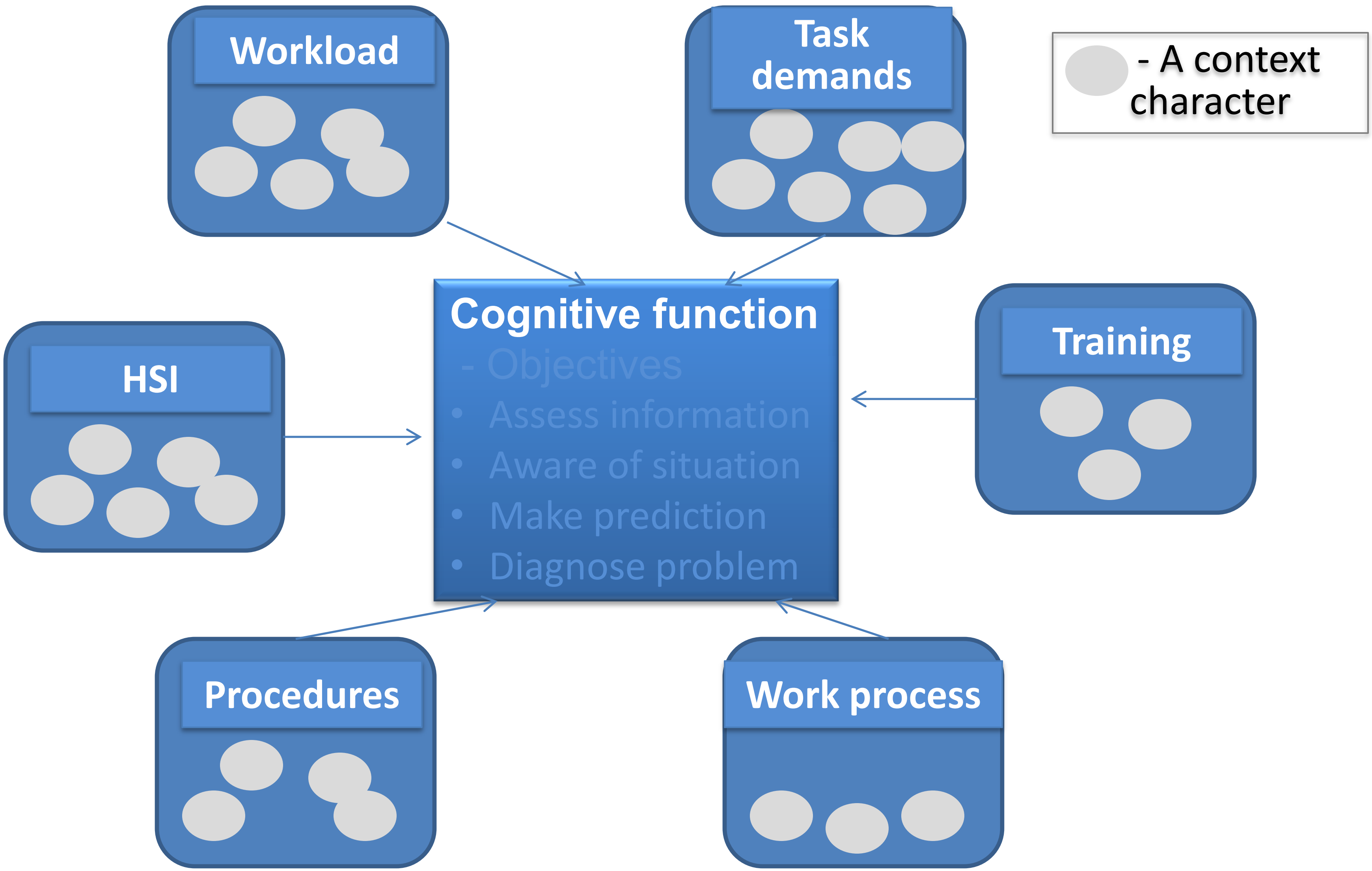
Scoping analysis – Determine the HEP range



Detailed failure-mode analysis – Estimate HEPs

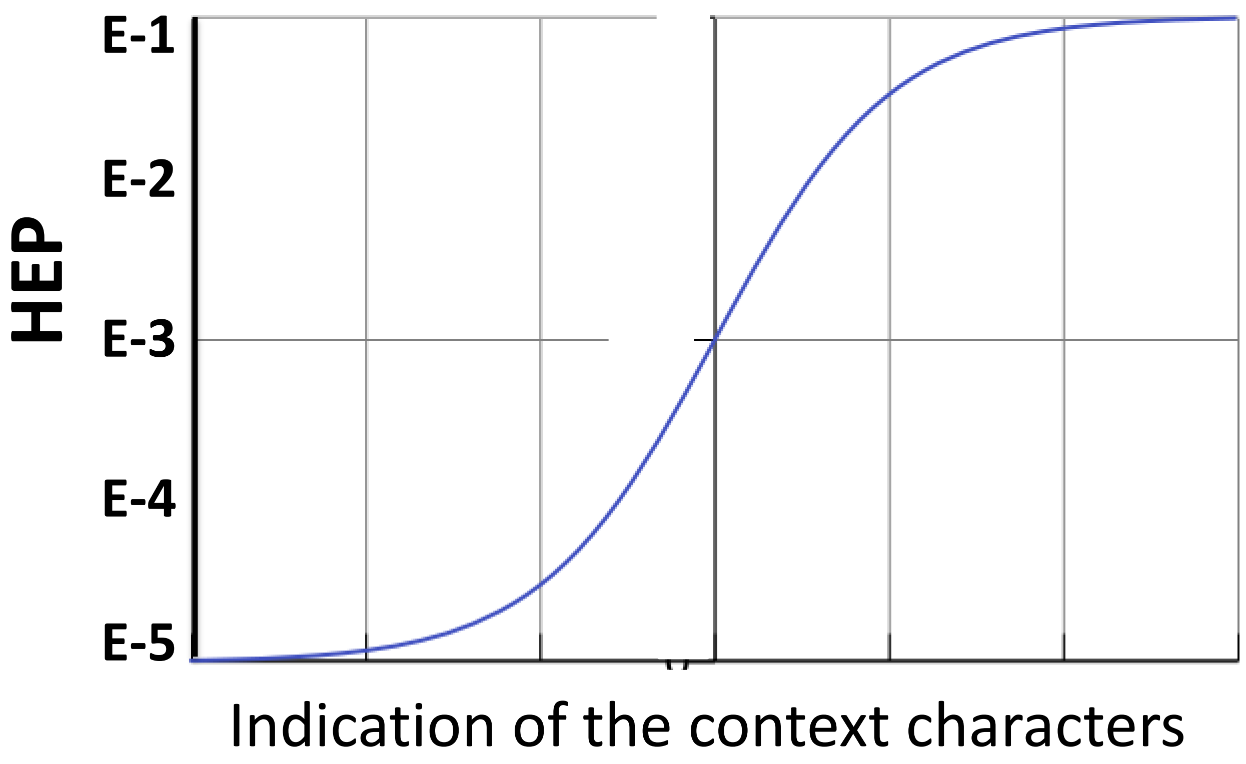


Scoping analysis

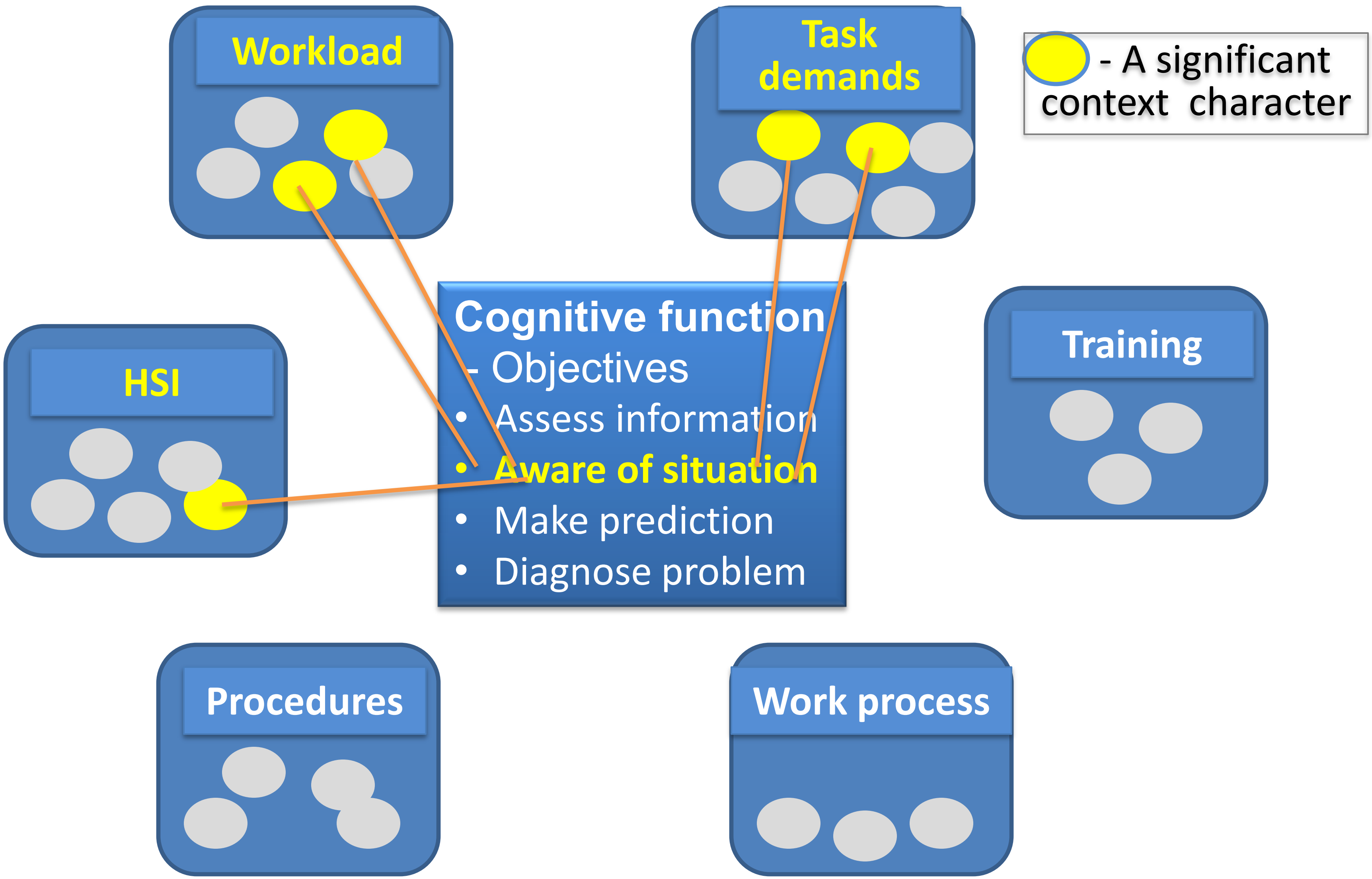


Scoping analysis

- Scoping analysis is to rank and group the failure probability of each cognitive function
- A HEP range is determined by the indication of context characters



Detailed failure-mode analysis



Generic task failure modes

Generic task failure modes represent possible types of failure of cognitive task objectives (i.e., subtasks).

Example task failure modes for the *Understanding* function

Objectives	Generic failure modes	CFMs for internal at-power events
Assess and verify information	Not assess / verify conflicting or ambiguous information	Critical data misperceived
Maintain situational awareness	Fail to maintain situation awareness	Critical data not checked with appropriate frequency
Diagnose problem s	Diagnose the wrong causes to the problems	
	Incomplete diagnosis	

Represent a failure mode in a Decision Tree (DT)

A DT consists of branches representing the context characters that are **most relevant** to the failure mode for the specific task domain.

The internal event IDHEAS method has DTs for the 14 CFMs.

To develop DTs for task domains other than internal at-power events:

- If a failure mode is an internal CFM, use the existing DTs in the internal event IDHEAS method and modify it as needed by

1) examining the character list to identify additional significant characters,
2) adjusting the DT branches.

- If a failure mode is not an internal CFM, develop the DT by

1) examining the context character list, and
2) selecting the characters that most significantly contribute to the failure mode.

HEP estimation

Objective – Obtain the HEPs in the scoping and detailed failure mode analysis

Short-term goals:

- Obtain HEP estimates through expert elicitation;
- Provide guidance for expert elicitation of HEPs;

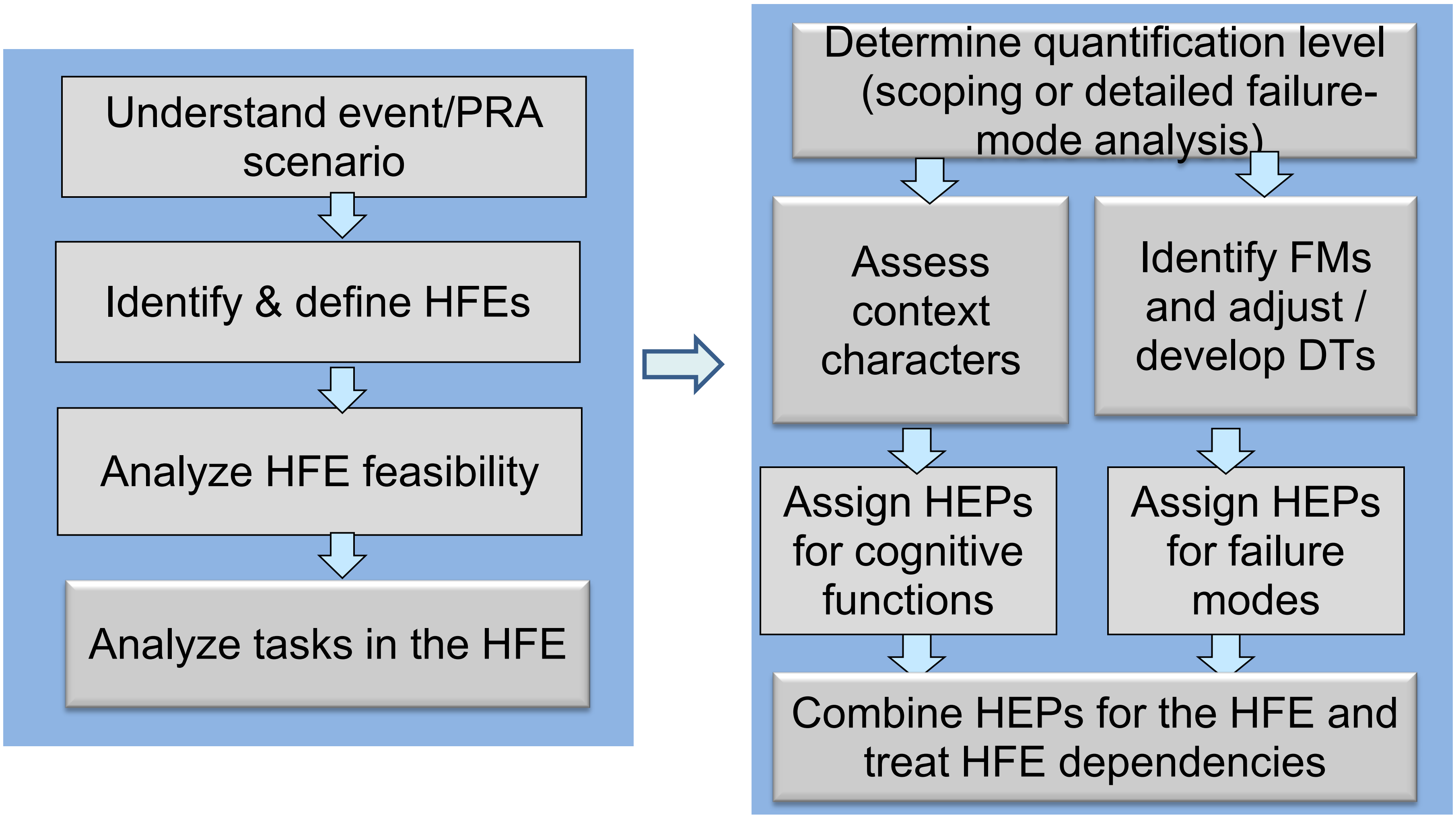
Long-term goals:

- Data-driven HEP estimation - Use the data from SACADA and other data sources to calibrate HEPs.

Integrative analysis
– Adapted from NUREG-1921

- Guidance for HFE analysis
- Cognitive error-causal tree
- Qualitative Analysis Structure
- Quantification process
- Integrative analysis
  - Dependency analysis
  - Sensitivity/uncertainty analysis
  - Recovery analysis

Summary of the IDHEAS process



Summary

- The cognitive error-causal tree enhances the state-of-the-art of HRA and human factors engineering.
- The generic IDHEAS methodology is intended to be applicable to all HRA domains in NPP operation.
- The generic methodology needs to be explored with its intended applications (e.g., LPSD, Level-3 PRA).
- Further development and refinement of the methodology will be made through exploration, piloting, and testing.

The IDHEAS HRA Method for internal at-power events

ACRS subcommittee on Risk and Reliability
January 16, 2013

Gareth Parry (ERIN)
John Forester (INL)

A Collaboration of U.S. NRC Office of Nuclear Regulatory Research (RES) & Electric Power Research Institute (EPRI)

IDHEAS products

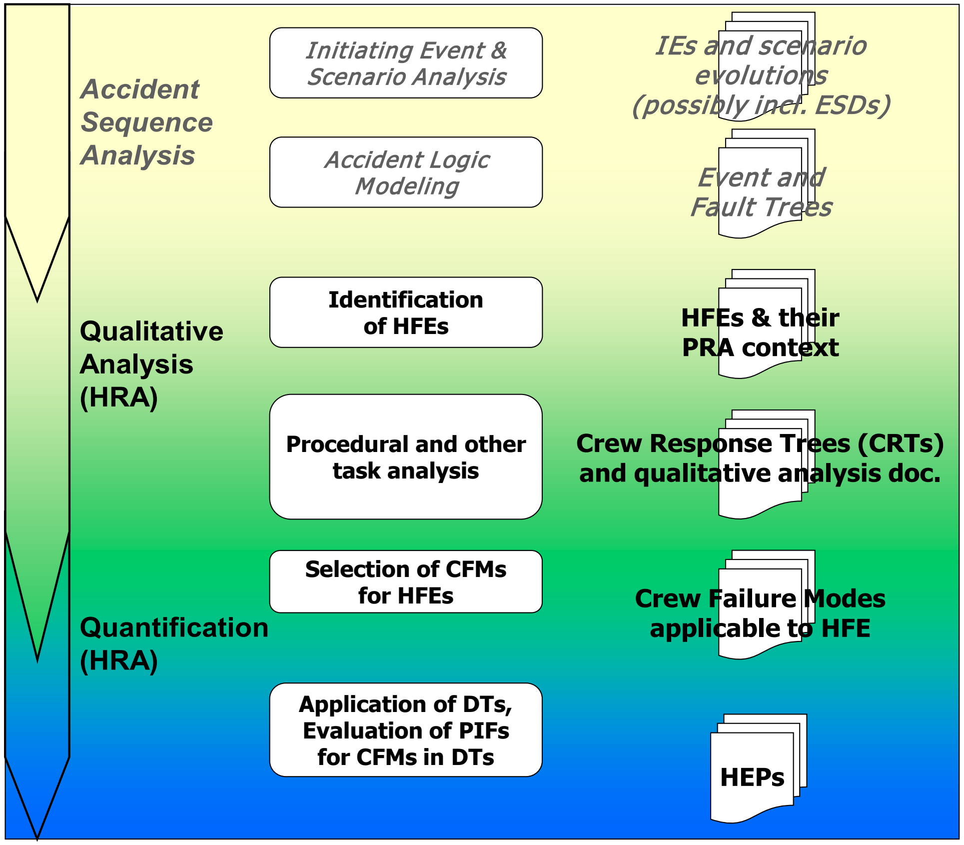
Product	Intended applications
Cognitive basis for human error analysis	• HRA • Human performance • Human factors engineering
Generic IDHEAS methodology for NPP applications	• HRA for all kinds of human events in NPP (Level-3 PRA, LPSD, external events, etc.)
An IDHEAS method for internal, at-power events	• **Internal, at-power event PRA (PRA models, SDP, ASP, etc.)**

Outline of Presentation

- Purpose is to provide an overview of the IDHEAS process for evaluation of HEPs of defined HFEs – draft report 12/17/2012
- Focus is on the method rather than the technical basis
- Elements of the Method:
  - Identification and definition of HFEs
  - Feasibility assessment
  - Task analysis and development of crew response tree (CRT)
    - Identification of critical tasks and opportunities for failure
  - Identification of applicable crew failure modes (CFM)
  - Use of decision trees to assess contextual impact on HEPs
  - Quantification
- Example Decision Tree
- Dependency

HRA Process

Overall HRA Process



Identification and Definition of Human Failure Events (HFEs)

Identification of HFEs

- ASME/ANS-RA-Sa-2009 definition: An HFE is a basic event that represents a failure or unavailability of a component, system, or function that is caused by human inaction, or an inappropriate action
- IDHEAS addresses those HFEs that represent failures to respond to a failure or unavailability of a system or function, or failures to manually initiate a required function following an initiating event
  - Post-initiating event HFEs
  - HFEs representing failures to respond that are included in fault tree models of initiating events
- Guidance for identification of HFEs is based on existing guidance
  - HRA Good Practices (NUREG-1792)
  - ATHEANA (NUREG-1624)
  - SHARP1 (EPRI-TR-101711)
  - Fire Human Reliability Analysis Guidelines (NUREG-1921)

Identification of HFEs (Cont'd)

- An HFE is included in a PRA model when an operator action is required to restore or maintain a function that is necessary to avoid core damage
- Identification based on the specific (procedural) guidance that specifies the required operator response (e.g., EOPs, AOPs, annunciator response procedures)
- Identification based on availability of cues that alert the operators to the need for response

Definition of HFE

- A typical HFE definition includes the following additional information:
  - Identification of additional cues or other information required to perform the response,
  - PRA scenario specific timing of cues and relevant information related to plant status,
  - The plant state or physical condition by which the operator action must be completed, and the corresponding time window (TW), and
  - The equipment (e.g., system or systems) the crew uses in order to achieve the functional goal and the way in which the equipment is to be used to achieve success (e.g., initiate injection using system X, perform depressurization using SRVs).

Assessment of Feasibility

Assessment of Feasibility

- An HFE is included in the model if the identified operator response is considered feasible in the context defined by the PRA scenario
- Feasibility of the identified response is performed based on:
  - Time available to complete the response
  - Availability of procedures and/or training
  - Availability of cues
  - Accessibility to areas where response is to be performed
  - Availability of resources
    - Personnel
    - Equipment
- Assessment of feasibility may be performed whenever the information required has been obtained
  - key information may not be available until the timeline for the IDHEAS task analysis has been completed

Task Analysis and Development of Crew Response Trees

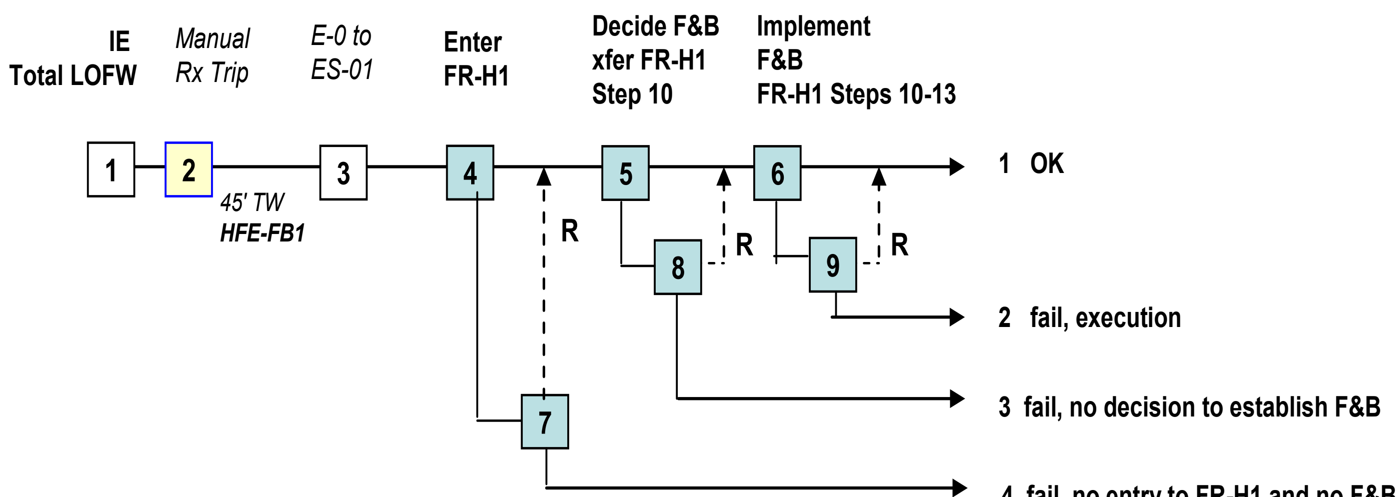
Task Analysis – Identification of Opportunities for Error

- HFE definition so far is at a fairly high functional level
- Success criteria are stated in forms such as "operators must do X by time T or before pressure exceeds Y"
- The next step is to understand in detail what activities are required to perform to "do X" and to construct the time line of items of significance (e.g., occurrence of cues, expected times to reach a certain step in the procedure, as well as the time window)
- Procedural task analysis:
  - Identify essential activities and their nature (e.g., collect data, interpret in light of criterion, decide to execute, perform execution, etc.)
  - Other steps (e.g., verification of status) add to time taken but their omission does not lead to failure
- Failure of these essential tasks results in the HFE

Overview of Task Analysis

- Stage 1. Characterization of the expected success path
  - Identify path through the procedures
  - Identify significant tasks (entry into procedure, transfer within or to a procedure, decision point, execution)
- Stage 2. Identification and definition of critical sub-tasks
  - Identification of critical sub-tasks associated with steps in the procedure (e.g., collecting data, comparison with decision criterion, specific execution tasks) and the requirements for success in those sub-tasks
- Stage 3. Identification of Recovery Potential
  - Identify opportunities to recover from failure of a critical sub-task as a result of new cues, other procedural steps, etc.

Representation as a Crew Response Tree (CRT)



Definition of Nodes of CRT

- Each node is specified by the subtasks required:
  - e.g., Node 4 requires an operator, using the CSFST to read the SG levels and AFW flow, compare values with a criterion, and when the criterion is satisfied, transfer to FR-H1.
- Development of time line is concurrent with the development of the CRT
  - Critical times include, time taken to reach the point at which entry into the CSFSTs is directed, time at which the cues are as stated for the criterion

Overview of Quantification Model

Overview of Quantification Model

- The HEPs are assessed on the basis of explanations of why the HFE might occur (e.g., due to various conditions the crew dismisses relevant information that results in their failure to achieve the required response).

- Explanations (crew failure scenarios) are grounded in an understanding of cognitive mechanisms and the related performance influencing factors (PIF).

- Crew failure scenarios are grouped by crew failure mode (CFM). Scenarios within the group are differentiated by the presence or absence of specific PIF characteristics that contribute to the likelihood of failure.

Crew Failure Modes

- The CFMs are determined from:
  - An identification of the ways in which an operating crew can fail to perform the functions of:
    - Plant status assessment
    - Response planning
    - Execution
  - Tailored to an understanding of the nature of the subtasks that need to be performed to achieve success in those functions
    - Responding to an alarm
    - Obtaining information
    - Evaluating information using procedures
    - Execution

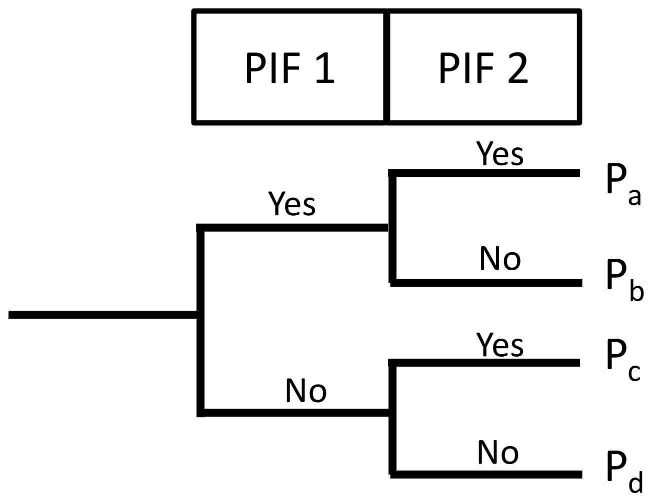
CFMs and Phase of Response

	Phase of Response		
	Plant Status Assessment	**Response Planning**	**Execution**
Crew Failure Mode	Key alarm not attended to	Delay implementation	Fail to initiate execution
	Data misleading or not available	Misinterpret procedure	Fail to execute response correctly
	Premature termination of critical data collection	Choose inappropriate strategy	
	Critical data misperceived		
	Wrong data source attended to		
	Critical data not checked with appropriate frequency		
	Critical data dismissed/discounted		
		Misread or skip step in procedure*	
	Critical data miscommunicated**		

*may occur in either ‘Response Planning’ or ‘Execution’ phases.
**may occur in any of the three phases.

Quantification Model - Decision Tree Approach

- Decision points relate to existence of those PIF categories that relate to the cognitive mechanism leading to the CFM
  - Determined from the literature review
- Decision tree paths represent different crew failure scenarios distinguished by the specific characteristics of the PIFs present
- A probability is assigned to each end point (path)



Quantification Equation

The quantification of the HEP takes the following form for a PRA scenario S:

$$HEP(HFE|S) = \sum_{CRT\ sequence} \sum_{CFM} Prob(CFM\ |CRT\ sequence, S)$$

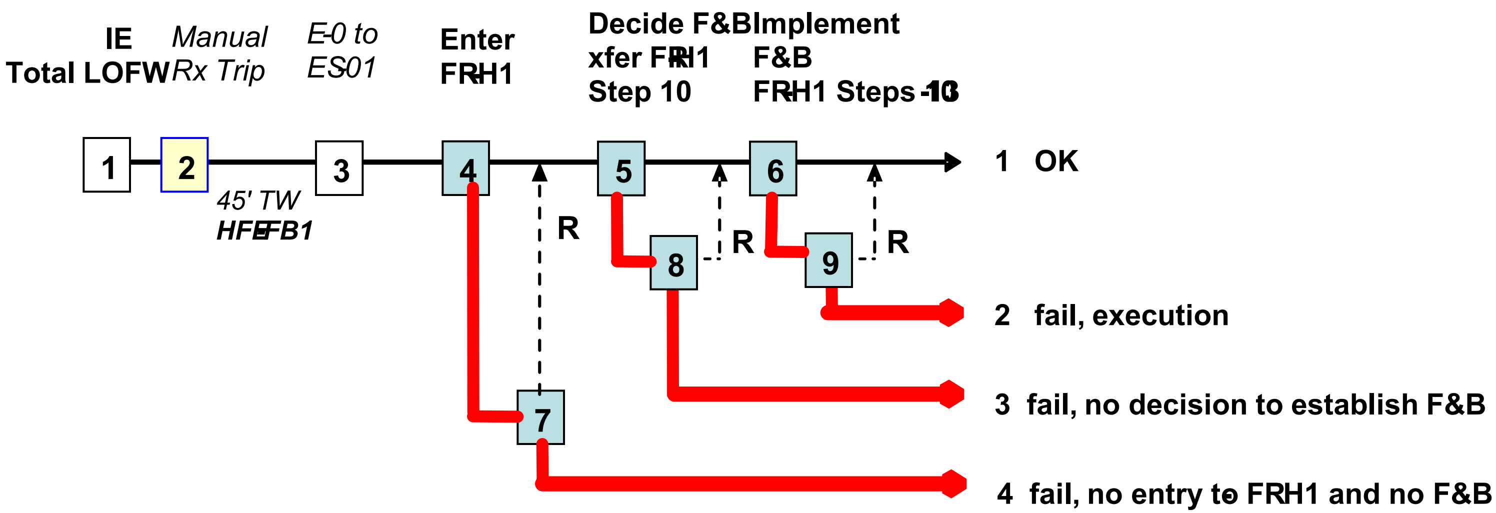
where the outer sum is over the CRT sequences that leads to the HFE, and the inner sum is over the CFMs that are relevant for the CRT sequence.

$$Prob(CFM\ |CRT\ sequence, S)$$

is the probability associated with the end point of the path through the DT for the specific CFM

The appropriate path is determined by the context (PIF characteristics) determined by the HFE boundary conditions.

Crew Response Tree (CRT) Sequences for Quantification



Steps in Quantification Process

Outcomes of qualitative analysis

- PRA scenario understanding;
- HFEs and definitions
- CRT and critical tasks
- Description of critical tasks
- Timeline and feasibility assessment

HFE quantification

Cognitive analysis of critical tasks

↓

Select crew failure modes (CFM)

↓

Select DT paths and assign HEPs

↓

Calculate combined HEP for the HFE

↓

Treat dependencies of HFEs.

Selection of CFMs

- For each CRT sequence, the initial node will be defined in terms of the critical sub-tasks (i.e., what do the crews have to do to respond correctly)

- Based on the nature of the subtasks, choose the appropriate CFMs using the following table (next slide)
  - As implied in the equation there may be multiple subtasks and multiple CFMs for each node.

Response phase	Nature of activities/sub-tasks included in CRT Node success	CFM
Plant status assessment	Does success require alarm response?	Key alarm not attended to
	Does success require data collection to assess plant status?	Data misleading or not available
		Decide to stop collecting critical data
		Critical data incorrectly processed / misperceived
		Wrong data source attended to
		Critical data dismissed/discounted
	Does success require monitoring for a critical plant parameter as a cue to initiate response?	Critical data not checked with appropriate frequency
Response planning	Does success require responding when a critical value is reached (given the value has been recognized)?	Delay implementation
	Does the success require a decision (e.g., transfer to another procedure, or initiate action) which if performed incorrectly would lead to an incorrect path through the procedures?	Misinterpret procedures
	Does the procedure allow a choice of strategies?	Choose inappropriate strategy
Execution	Does the branch point address execution?	Fail to initiate execution
		Fail to execute response correctly
All phases	Is written procedure being used?	Misread or skip steps in procedures
All phases	Is communication between crew members required?	Miscommunication

Example – Node 4

- Essential activities
  - Monitoring the NR Level in the SGs and the total AFW flow to SGs
  - Comparison with Criterion
    - Criterion 1 “NR Level in at least one SG GREATER THAN X%” – NO
    - Criterion 2 “Total AFW Flow to SGs GREATER THAN Y GPM” – NO
  - Making a decision

Example – Node 4 (Cont'd)

- All CFMs are potentially relevant except:
  - Key alarm not attended to
  - Choose inappropriate strategy (no alternate strategies)
  - Fail to initiate execution (not an execution step)
  - Fail to execute response correctly (not an execution step)
  - Miscommunication (dedicated operator)

Treatment of Recovery

- The recovery branches on the CRT are addressed integrally within the DTs

- In that way, recovery potential is assessed as a function of the CFM so that the cognitive dependence is handled appropriately

For each relevant CFM Select DT Path

- For each of the relevant CFMs for CRT sequence, assess the branches on each DT
- based on the HFE context and determine the path through the DT
- Pick end point probability
- Substitute HEPs into the equation (slide 20)

Integration into PRA Model

- What we have presented is a model for the estimation of an HEP for a defined HFE
- In implementing this or any other method that addresses HFEs one at a time, the issue of dependency needs to be addressed
- Because it is a cause-based approach, IDHEAS affords the opportunity to deal with dependency in a more complete way than the current "Commonality of PSFs" approaches

Elements of Treatment of Dependency

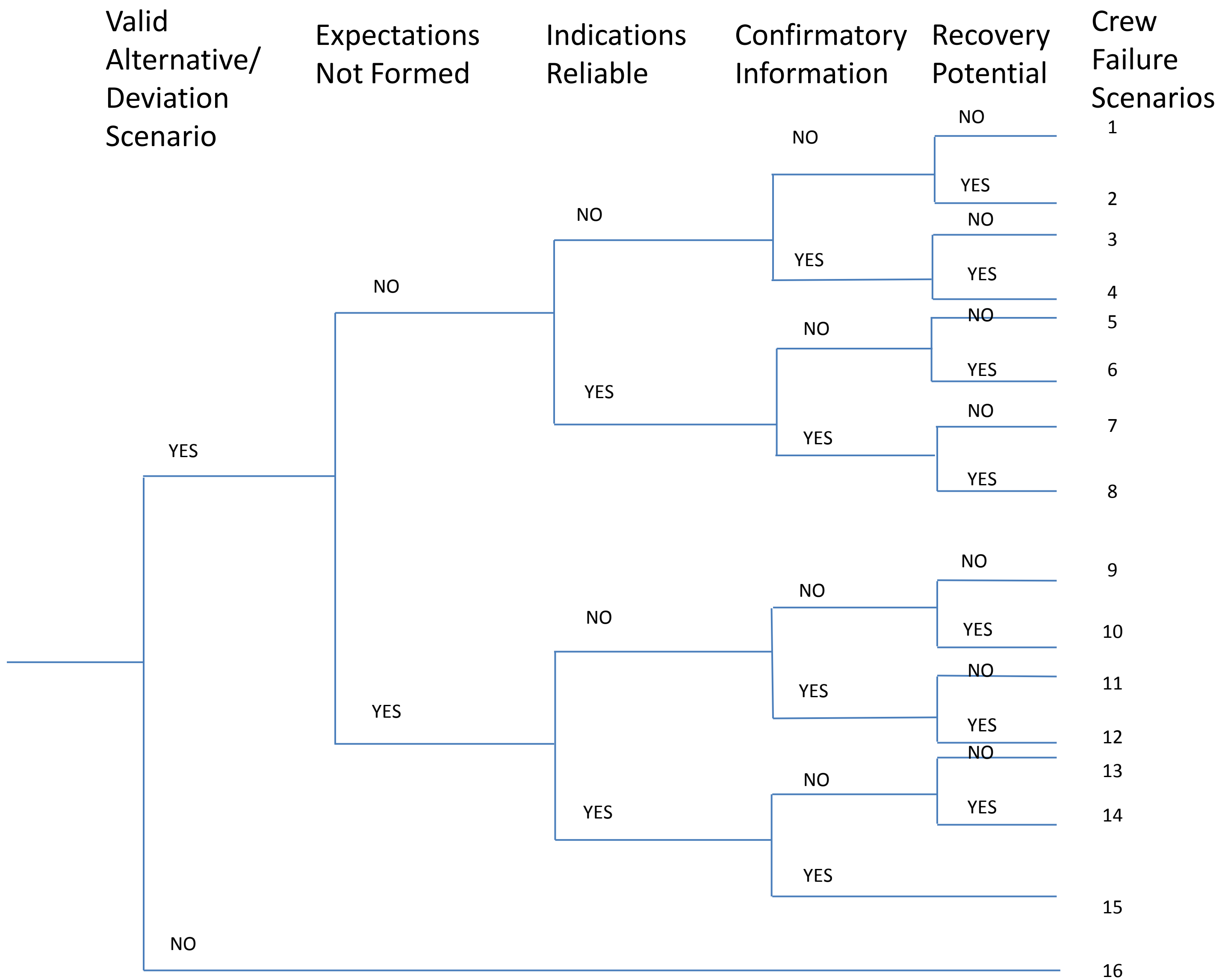
- Identify why and how the first HFE can occur in terms of the CFMs and PIF characteristics
- Develop a complete picture of the overall scenario to understand the responses subsequent to the first failure and identify the task demands, cognitive resources, and CFMs and PIF characteristics associated with the new context
- Understand the cognitive processes underlying the CFMs of the initial and subsequent HFEs and use the cognitive mechanisms to examine the dependency effects (e.g., is the potential for an incorrect mental model that would lead to failure of both responses to persist)

Example Decision Tree

Example of Decision Tree – Critical Data Dismissed/Discounted

- CFM definition: The crew is aware of and has obtained the correct information (e.g., the value of a key plant parameter, the status of a piece of equipment, information that has been communicated by another person, etc.), but has discounted it from the assessment of the plant status (and therefore represents an incorrect synthesis of the information they have).
- This CFM is applicable when the information being dismissed is an essential part of assessing the plant status for which there is one (or possibly more than one) successful response.

Critical Data Dismissed/Discounted



Explanation of branch points

- Branch Point 1: "Valid Alternative/Deviation Scenario" assesses whether, with the data dismissed, there is a plant status that is valid and should be within the knowledge base of the crew. If this is not the case, take the NO branch. Otherwise, take the YES branch

- Branch Point 2: "Expectations Not Formed/Irrelevant" addresses the question of bias from training and knowledge/experience/expertise with respect to the plant status. Importantly, are the training and experience sufficient to create a strong expectation that the critical data is irrelevant and can be dismissed?

Explanation of branch points (cont'd)

- Branch Point 3: "Indications Reliable". If the crew judges the plant indications (HSI output, procedural quality, etc.) as unreliable, they may be likely to dismiss the information that the indicators are providing. This does not apply when the known areas of unreliability are well understood by the crew or when a warning of the potential unreliability is given in the procedure

- Branch Point 4: "Confirmatory Information" questions whether if it is the case that, before dismissing any piece of information, the crew searches for some confirmatory information as this may mitigate the discounting of information. There may be specific procedural steps that the operators engage in to confirm the information and/or the operators may perform confirmatory checks as a matter of good practice

Explanation of branch points (cont'd)

- Branch Point 5: "Recovery Potential" addresses the possibility that, even if the crew/operator makes the wrong decision initially, there is a means of timely self-recovery. For instance, the operator (given the incorrect plant status assessment) might be expecting a particular plant response. If this response does not occur or is different than what is expected, the operator may re-analyze the plant status which may result in correcting the previously inaccurate assessment. In addition, future procedural steps may lead the operators to make the appropriate decisions to get back on track for that function. If the crew has opportunities to reassess the plant status, this could serve as a recovery potential

Consideration of Feasibility Assessment in IDHEAS HEP Estimation

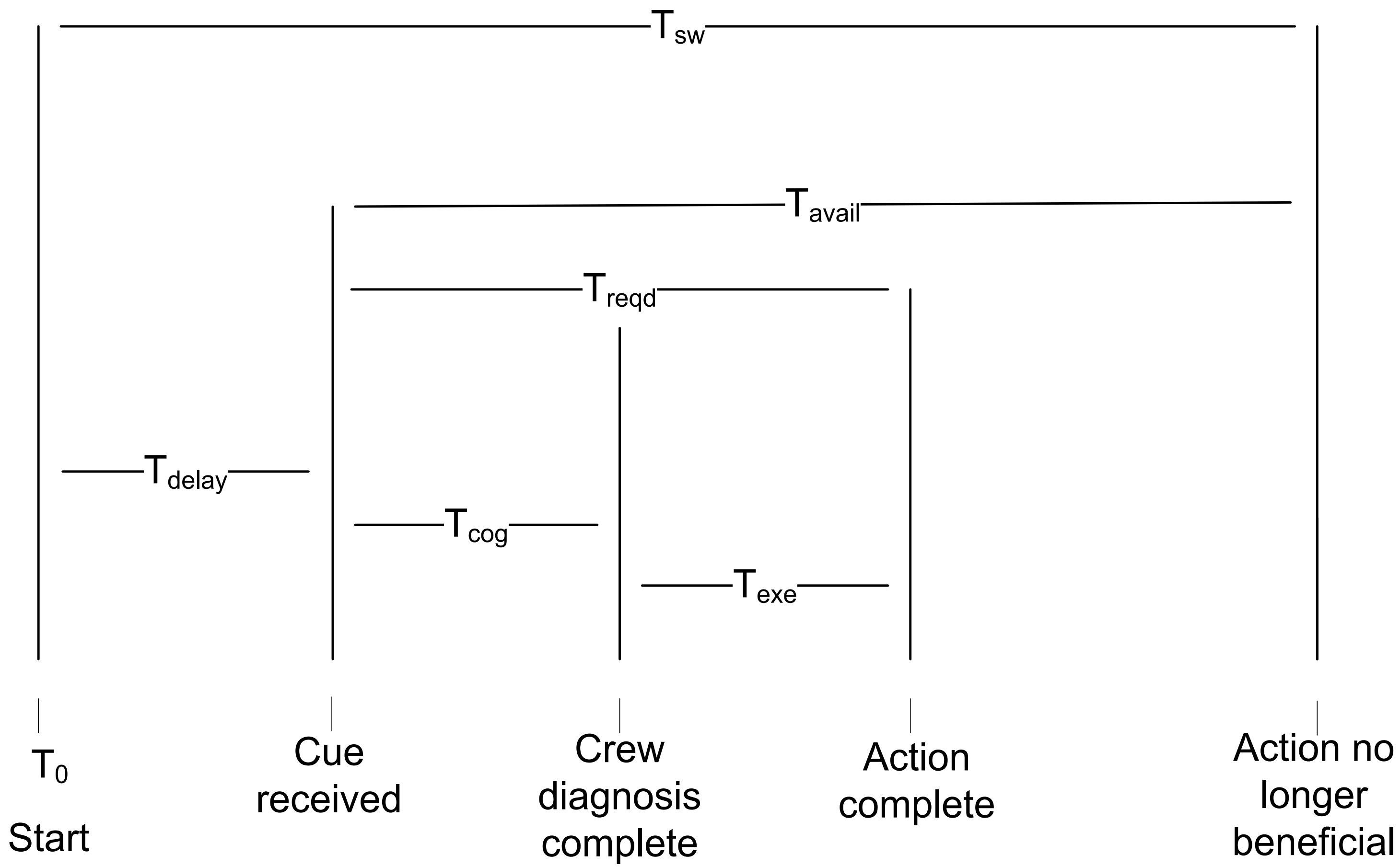
Assessment of Feasibility

- An HFE is included in the model if the identified operator response is considered feasible in the context defined by the PRA scenario
- Feasibility of the identified response is performed based on:
  - Time available to complete the response
    - <u>Must be adequate time available to diagnose and complete the response</u>
  - Availability of procedures and/or training
  - Availability of cues
  - Accessibility to areas where response is to be performed
  - Availability of resources: personnel, equipment
- If not feasible, then HFE not included in the model or the HEP set to 1.0
- If feasible, still another consideration for assessing <u>reliability</u> with the IDHEAS DTs

Time and Use of IDHEAS DTs

- An underlying assumption of the DTs is that the actions are feasible from a timing perspective.
  - Adequate time is available for the operating crew to diagnose the need for and complete the actions for a particular HFE.
  - So time availability is not a driving issue on performance
- There can be variability in the time required by different operating crews to complete the actions
- Uncertainty associated with estimating the time required for the operator actions associated with an HFE
- How do we have some confidence that the DTs can be applied without worrying about time limitations?
- Want to keep the assessment as simple as possible
  - Currently a couple options

Timeline Diagram



Two Options

- IDHEAS provides detailed guidance for obtaining realistic estimates of the time required – Strength of the approach
- Option 1 – Use the maximum time it would be expected for all of the crews to complete the actions required (T_{cog} + T_{exe}), under the conditions present in the scenario. If less than the time available, OK to use DTs
- Option 2- Use average crew response time with a time margin (extra time for the action) of 100%.
- Note that procedure based actions (e.g., those in EOPs, alarm and abnormal plant procedures) have been vetted in terms of whether there should generally be enough time available for the actions

United States Nuclear Regulatory Commission

Protecting People and the Environment

International HRA Empirical Study Overall Lessons Learned

Erasmia Lois, PhD
Human Factors and Reliability Branch
Division of Risk Analysis

Presentation to:

Advisory Committee on Reactor Safeguards
Reliability and PRA Subcommittee
January 16 2013

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

Briefing Objectives

- Inform the ACRS on the overall results and lessons learned from this international collaborative effort
- Obtain feedback on the draft NUREG-2127, Summary of Insights and Lessons Learned

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

Study Objectives

- Assess HRA methods and practices in light of NPP control room simulator data
  - Characterize the methods
    - Identify strengths and weaknesses
  - Provide the technical basis for improving the methods method and method implementation
  - Improve HRA practices in general
  - Support addressing ACRS and Commission direction on HRA

Motivation

- Differences in the underlying frameworks, data, and quantification algorithms of HRA methods yield different human error probabilities and different insights regarding the potential drivers of error/failure
- Models are based on formal and informal human performance theories but have not been tested with empirical data
- Improving the robustness of PRA/HRA has been the NRC's focus
- The ACRS and Commission have recommended the need to address variability in HRA results

Status

- Simulation runs performed at the Halden HAMMLAB 11-12/2006
- Phase 1—Pilot: *Description of Overall Approach and Pilot Results from Comparing HRA Methods Predictions to Simulator Performance Data*, NUREG/IA-0216, Vol.1 (HWR-844), 11/2009
- Phase 2: *Comparisons of HRA Predictions to Simulator Data from SGTR Scenarios,* NUREG/IA-0216, Vol. 2 (HWR-915), 8/2011
- Phase 3: *Comparisons of HRA Predictions to Empirical Simulator Data from LOFW Scenarios*, NUREG/IA-0216, Vol. 3 (HWR-951), to be published, February 2013
- NUREG-2127 (HRP-313), *The International Empirical study, Lessons Learned from Comparing HRA Methods Predictions to HAMMLAB Simulator Data*, to be published, March 2013

Authors

NUREGIA-0216 & NUREG-2127

John Forester[1,6], Andreas Bye[2], Vinh N. Dang[3], Erasmia Lois[4], Jeff Julius[5], Salvatore Massaiu[2], Helena Broberg[2], Per Øivind Braarud[2], Ronald Boring[6], Ilkka Männistö[7], Huafei Liao[1], Gareth Parry[4,9], Pamela Nelson[8]

1 Sandia National Laboratories, USA

2 Institute for energy technology, OECD Halden Reactor Project, Norway

3 Paul Scherrer Institute, Switzerland,

4 U.S. Nuclear Regulatory Commission, USA

5 Scientech, USA

6 Idaho National Laboratory, USA

7 VTT, Finland

8 Universidad Nacional Autónoma de México, Mexico

9 ERIN Engineering

HRA Methods &Teams

- ASEP/THERP-- NRC staff and consultants, USA
- ATHEANA--NRC staff and consultants, USA
- CBDT EPRI--(Scientech), USA
- CESA--PSI, Switzerland
- CREAM--NRI/Czech Rep
- Decision Trees + ASEP NRI--Czech Rep
- HEART--Vattenfall & Ringhals NPP, Sweden
- KHRA--KAERI, Korea
- MERMOS ---EDF, France
- PANAME--IRSN, France
- SPAR-H--Idaho National Laboratory, USA
- SPAR-H--Idaho National Laboratory, USA
- THERP with Bayesian Enhancement/--VTT

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

International HRA Empirical Study Overview

Scenario Definitions
OECD Halden, Assessment Group

crews – one power plant

HRA Analyses
Analysis teams

Simulator Sessions: Observation, Data Collection
OECD Halden

Summary of Predictions
Assessment Group

Data Analysis
OECD Halden

Comparison: HRA Predictions vs. Crew Data
Assessment Group

Development of Insights to Improve HRA Methods and Practices

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

HRA Method Assessment

- Method assessment was based on a *comparison of method results to empirical evidence* —we call it the "*predictive power*" of
  - qualitative analysis
  - quantitative analysis
- The *traceability* of qualitative/quantitative analysis
- The usefulness of qualitative and quantitative results to *human error reduction*
- The *adequacy of the guidance* provided by each method for qualitative and quantitative analysis

HRA Method Assessment (cont)

- Qualitative predictive power evaluations based on
  – PSF assessments - how well the method applications predicted the specific performance issues and drivers observed in the reference data
  – Operational expression assessments - how well the method applications predicted the ways crews could fail and the operational situations that could contribute to the failure paths

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

HRA Method Assessment (cont)

- Quantitative predictive power evaluations based on
  - Potential optimism of the most difficult HFEs
  - Consistency of the ranking of the HFEs (on the basis of estimated HEPs) with the difficulty rankings based on the empirical evidence
  - Quantitative differentiation of the HFEs by HEP
  - Predicted HEPs relative to the confidence and uncertainty bounds of the reference data

Empirical Results

- Designing easy (base case) and complex scenario variants produced variability in crew performance
  - provided a basis to go beyond failure counting and examine a broader spectrum of performance issues as well as to rank order the HFEs
- Developing “operational descriptions” allowed comparisons of empirical evidence with method predictions
  - such descriptions explain how tasks were performed, why performed in this way, and consequences
  - exemplified how HRA concepts such as diagnosis can be observed in actual crew responses
  - pointed out that HRA practices in which cognitive demands on operators are frequently not well examined can lead to missing important impacts on performance

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

Empirical Results (cont)

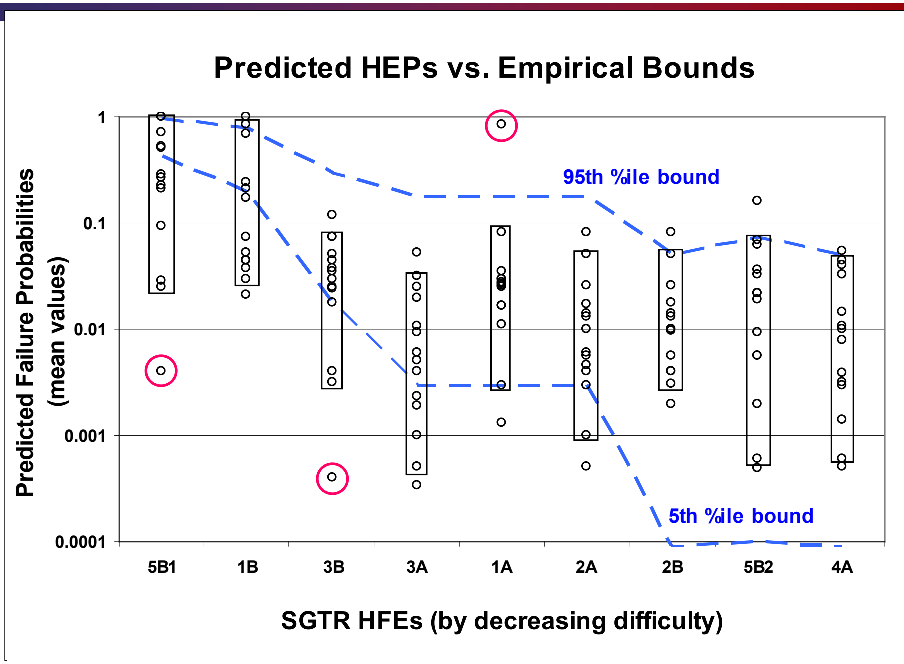
- Variability in crew behavior was observed
  - The frequently made analysts' assumption that crews will behave generally the same in a given scenario is not supported by the empirical evidence
  - taking into consideration what the impact might be if a crew takes a different path in executing the procedures is not being addressed in most methods
- The PSF analysis produced evidence of the presence and strengths of various PSFs in crew responses
  - the usefulness of the PSF definitions for explaining crew behavior was to some extent validated and provided a means for evaluating the HRA results
- Established clarity to fundamental HRA concepts
  - a major achievement of this study

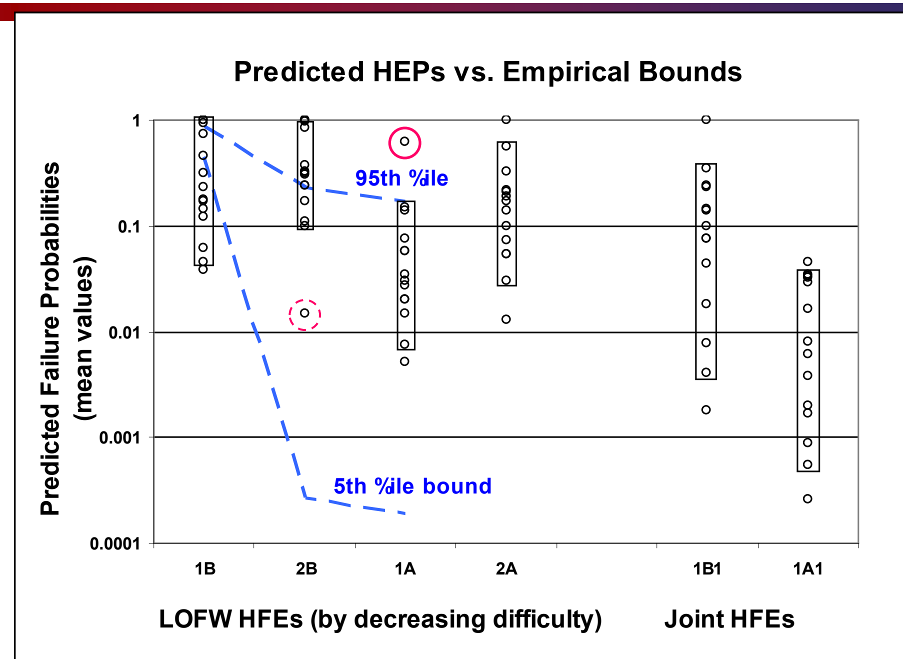
Overall Findings for Quantification

- *Optimistic HEPs for the most difficult HEFs:* Evidence of producing optimistic HEPs for the most difficult HEPs
- *Ranking of HEPs*: In many cases, HEPs do not reflect the relative difficulty levels of the HFEs observed in the evidence
- *Range and differentiation of HEPs*: The analyses did not always adequately discriminate among the difficulty levels, even in cases where they produced appropriate ranking
- *Conservative or realistic* HEPs: None of the methods consistently produced high (or low) HEPs for the set of HFEs
- *Comparisons of HEPs against confidence or uncertainty bounds of the reference data.* The uncertainty bounds of the reference data were broad for the easier HFEs and relatively narrow for the more difficult HFEs, reflecting the relative strength of the evidence for small HEPs vs. large HEPs

Predicted HEPs vs. empirical HEPs

(Bayesian results)



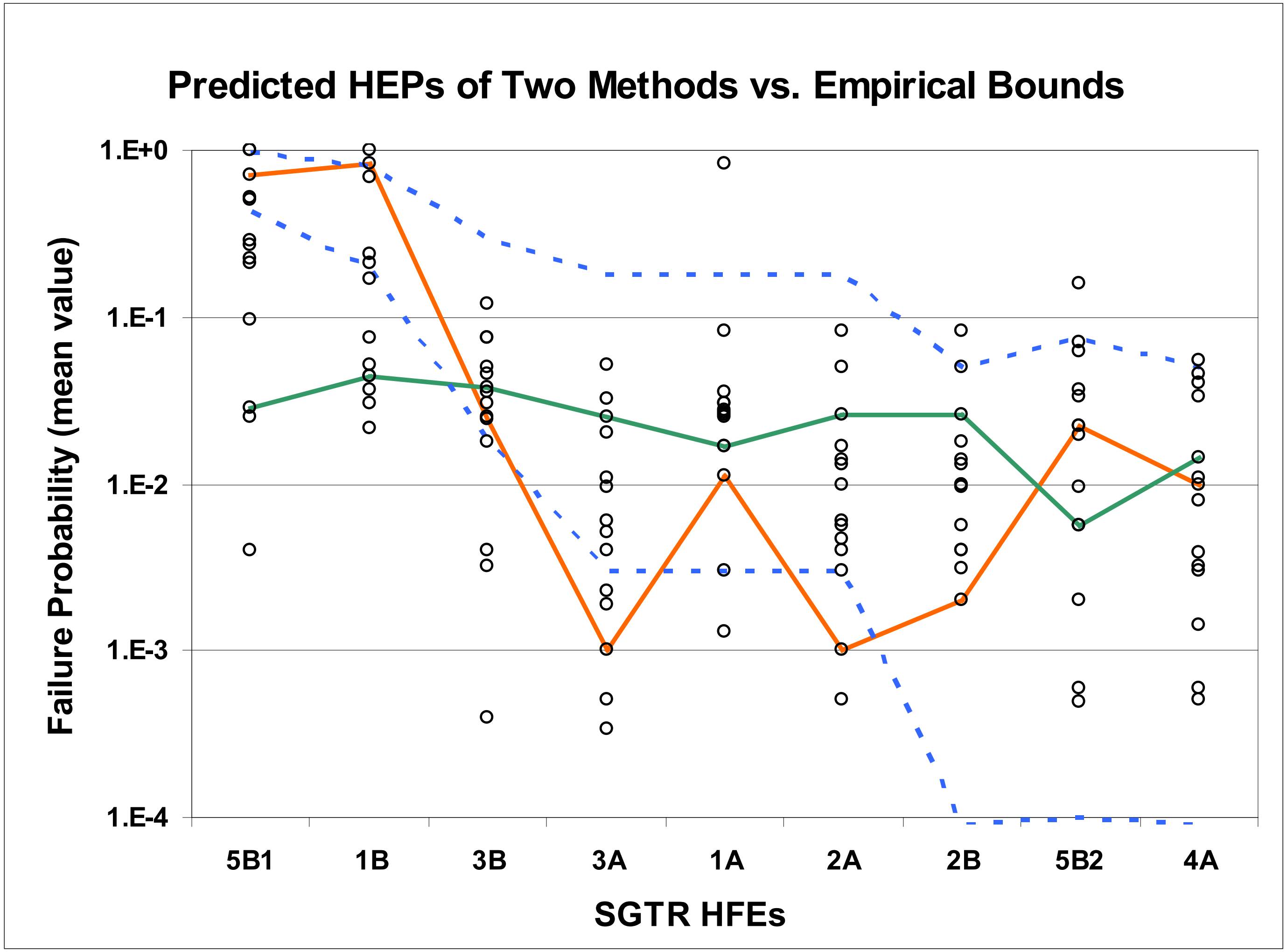


* Breadth of Bayesian confidence bounds are due to small data set.

- This show limits of comparisons based only on empirical (Bayesian) HEPs.

- After exclusion, most ranges span < 2 orders of magnitude
- Many methods underestimated HEPs of difficult HFEs in SGTR scenarios (4 at left)
- Rest of HFEs: nearly all predictions (mean values) fall within bounds, but very broad bounds
- Consistency of predicted ranks
– (by individual method)
– separate, important criterion for HRA methods

Quantitative Comparisons - Examples



from NUREG/IA-0216 Vol. 2, HWR-915

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

Overall Findings for Quantification

- The findings of this study w/r to quantification are important to be considered when HRA results are used for regulatory decision making.
- Issues such as the ranking of HFEs, the optimism of HFEs for even apparently difficult actions, the lack of adequate discrimination among HFEs, and the fact that no method can inherently produce “conservative” bounding values are important insights needed to considered in PRA/HRA applications

Overall Findings For Qualitative analysis

- *Handling of crew cognition tasks*: Failure to adequately consider cognitive activities can lead to a failure to identify important influencing factors and result in underestimations of HEPs
  - Some methods allow only addressing response execution in some situations
  - Others simply do not provide adequate guidance/treatment of normal cognitive activities in accident scenarios
- *Incorporation of failure mechanism and contextual factors: S*ubstantial evidence that methods that focus on identifying failure mechanisms and associated contextual factors produce richer content and frequently predicted actual crew performance
  - Evidence that HRA *does have the capability* to predict what could or would occur in responding to the scenario.
    - Methods using causal analysis (e.g., ATHEANA, CBDT, CESA, MERMOS) appear to have better capability to actually identify observed failure paths and causes

Overall Findings For Qualitative analysis

- *PSF Treatment:* selection of an appropriate PSF and judging the degree of influence of the PSF on performance is an important factor and contributed to both over- and under-estimation of HEPs.
- The range of PSFs provided by a method is an issue—analysts miss identifying failure contributors because the method is not covering them thru PSFs

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

Traceability of qualitative and quantitative analyses

- Two aspects of traceability
  - in qualitative analysis how judgments are made, e.g., basis for choices of PSFs and their weights
  - In quantification, given the choices made in the analysis
- PSF-based methods are good w/r to traceability in quantification but not as good w/r to qualitative analysis
- Context-based methods are good w/r to qualitative traceability
  - develop strong operational stories in which judgments made about the conditions facing the operators are easily understood and traceable
- Context-based methods are not as good w/r to traceability in quantification
  - Lack an easily traceable way of translating scenario stories into HEPs
  - No guarantee of reproducibility even when the analysts agree on the assumptions and aspects of the scenario descriptions

Insights for error reduction

- Most methods do not offer specific guidance for error reduction.
- Capability of error reduction using the more traditional PSF-based methods depends on the rigor of the underlying analysis and judgments made
- The newer, narrative-based methods have better capability to identify error reduction
  – describe how elements of the scenario, task, human-machine interface, and operator aids may contribute to the HFE.
  – Failure scenarios can be directly understood by plant and directly use them for error reduction.

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

Insights for improving guidance and methods

- Improved guidance in selection and treatment of PSFs for methods with limited range of PSFs
  – Identifying a comprehensive set of factors
  – How the identified factors can be assigned to the method’s PSFs, and
  – How to select an appropriate strength of the PSF
  – Better link between qualitative and quantitative analysis
- Improved guidance on HFE decomposition into sub-tasks and determining the level of decomposition
- Improved guidance for
  – considering not only the primary diagnosis/situation assessment but also for cognitive activities throughout the execution

Insights for improving guidance and methods (cont)

- Improved guidance
  - for performing a reasonableness check regardless of the method used to perform an HRA—This is one of the most fundamental insights of the study
    - the Empirical Study suggest that a reasonableness check was not performed in a number of analyses
  - For performing a thorough assessment of potential failure mechanisms in connection possible operational contexts compatible with the PRA scenario
- Evidence that such qualitative analysis leads to more comprehensive basis as an input to HFE quantification
  - Good qualitative analysis is needed along systematic way to tie to quantification
- However, given the limitations of methods, it is questionable if the focus should be on improving individual methods

Conclusions

- Comparison of HRA outcomes to crew performance outcomes was a major achievement of the study
  - Document crew performance from an HRA perspective required an in-depth analysis well beyond envisioned
  - Allowed an understanding of how methods are applied (vs their intended application)
  - Identified methods strengths and weaknesses of individual methods
  - Allowed a comparative analysis of methods to some extent
  - Identified strengths and weaknesses of the HRA field as a whole
- Major conclusion: no method meets all desirable attributes
  - A hybrid could be the solution
- More studies could further improve the robustness of HRA

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

- Back up slides

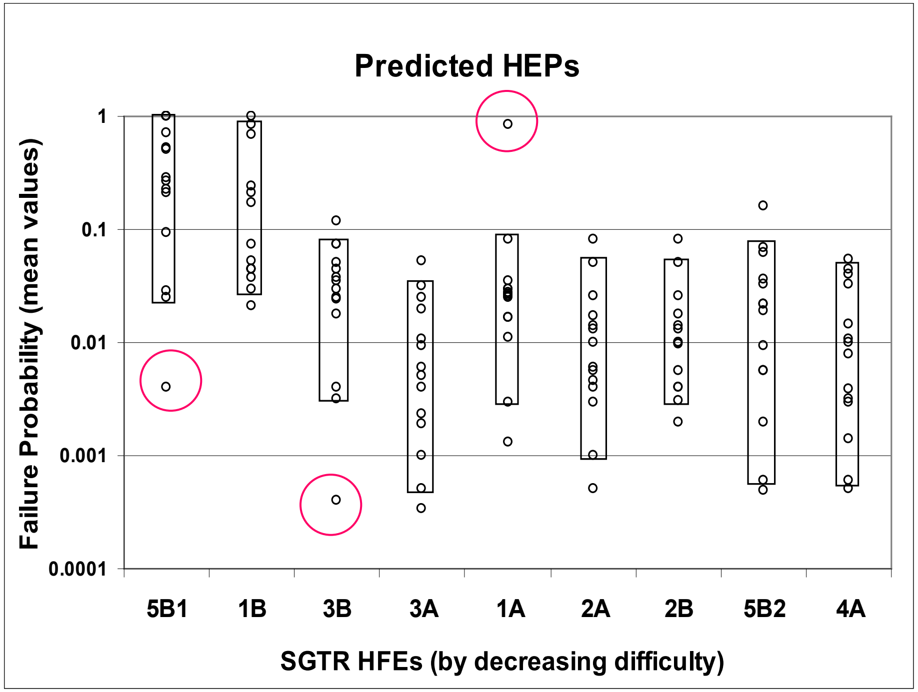
U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

Empirical Data Development

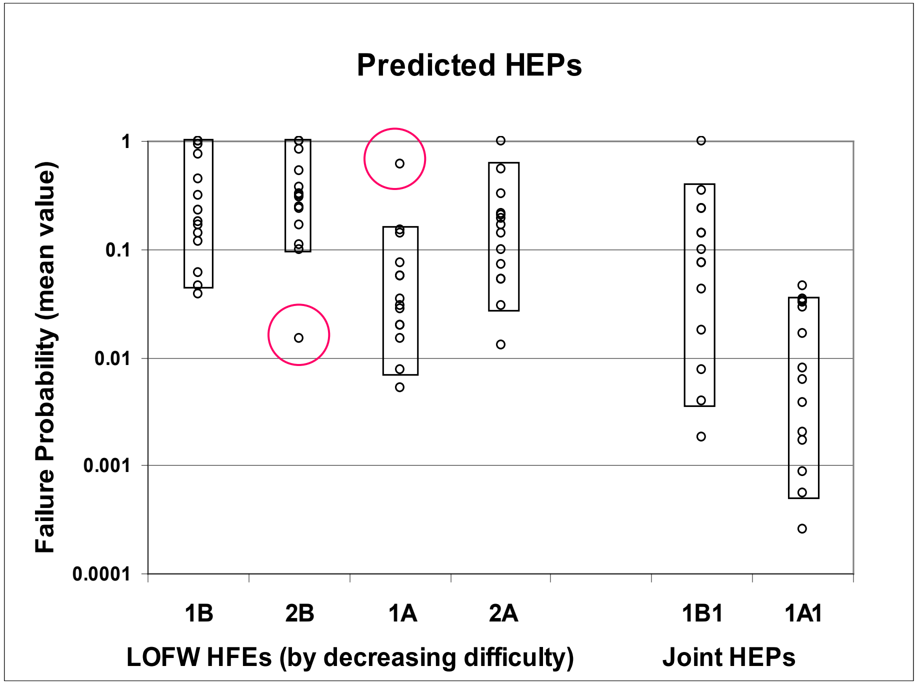
- Collection of raw data in operator logs, audio/video recordings, and crew interviews
- Crew-level data analysis to determine to what degree crews accomplished the tasks
- Determination of crew failures associated with various HFEs
- Development of operational descriptions
- Identification of performance shaping factors (PSFs) in crew performance and PSF ratings
- Ranking the difficulty of the HFEs on the basis of the empirical evidence

Range of predicted mean HEPs

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 (circled) are highlighted. Many due to faulty analysts' assumptions



The US HRA Empirical Study:

Assessment of HRA Method Predictions against Operating Crew Performance on a US Nuclear Power Plant Simulator

Presented by: Huafei (Harry) Liao, PhD

Sandia National Laboratories

Presentation to: Advisory Committee on Reactor Safeguards

Reliability and PRA Subcommittee

January 16, 2013, Rockville, MD

Briefing Objectives

- To inform the ACRS on the findings from the US HRA Empirical Study
- To obtain feedback from the ACRS to incorporate in the documentation of the study

Motivation

- Follow-on limited scope of the International HRA Study to address limitations
  - Use of Halden facilities and European crews
  - Analysts visit to reference plant
  - Mainly method-to-data comparison--no analyst-to-analyst effects
  - Commission interest to ensure the applicability of Halden human performance studies in US applications
- SRM-M090204B – February 2009
  - Pursue testing U.S. nuclear plant operating crews' performance in a variety of situations
  - Keep the Commission informed on benchmarking and HRA database projects

Contributors/Authors

John Forester[1,6], Huafei Liao[1], Vinh N. Dang[2], Andreas Bye[3], Mary Presley[4], Julie Marble[5], Helena Broberg[3], Michael Hildebrandt[3], Erasmia Lois[5], Bruce Hallbert[6], Tommy Morgan[6]

[1] Sandia National Laboratories, Albuquerque, NM, USA
[2] Paul Scherrer Institute, Villigen PSI, Switzerland
[3] OECD Halden Reactor Project, Institute for Energy Technology, IFE, Halden, Norway
[4] Electric Power Research Institute (EPRI), Palo Alto, CA, USA
[5] U.S. Nuclear Regulatory Commission (NRC), Washington, DC, USA
[6] Idaho National Laboratory, Idaho Falls, ID, USA

Contributors/HRA Teams

- Sandia National Labs
- Idaho National Lab
- NRC staff/contractors
- SAIC
- EPRI
- University of Mexico
- Czech Republic

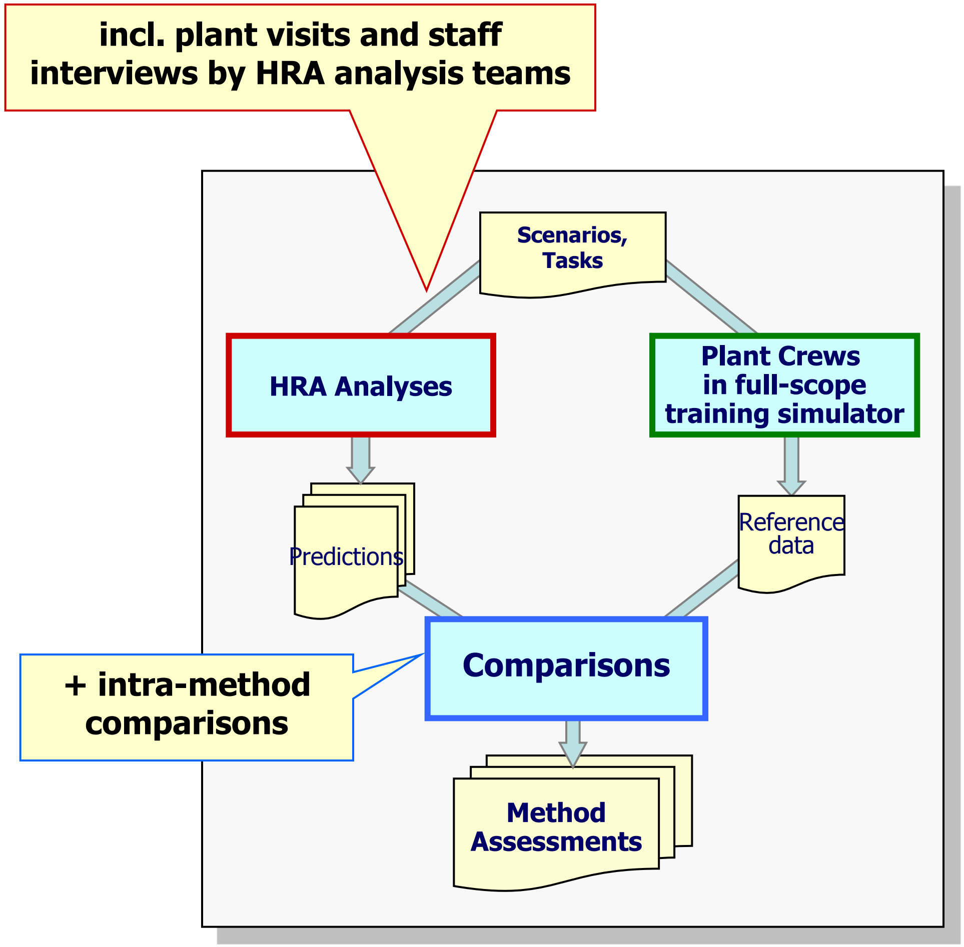
Study Objectives

- To improve insights developed from the International HRA Empirical Study
  - To be able to separate analyst effects from method effects
  - To control for the limitations in the International Study where the HRA teams were not able to visit the plant, interview plant personnel, or observe simulator runs
- To obtain insights on generalizability of Halden human performance studies to US applications

Study Status

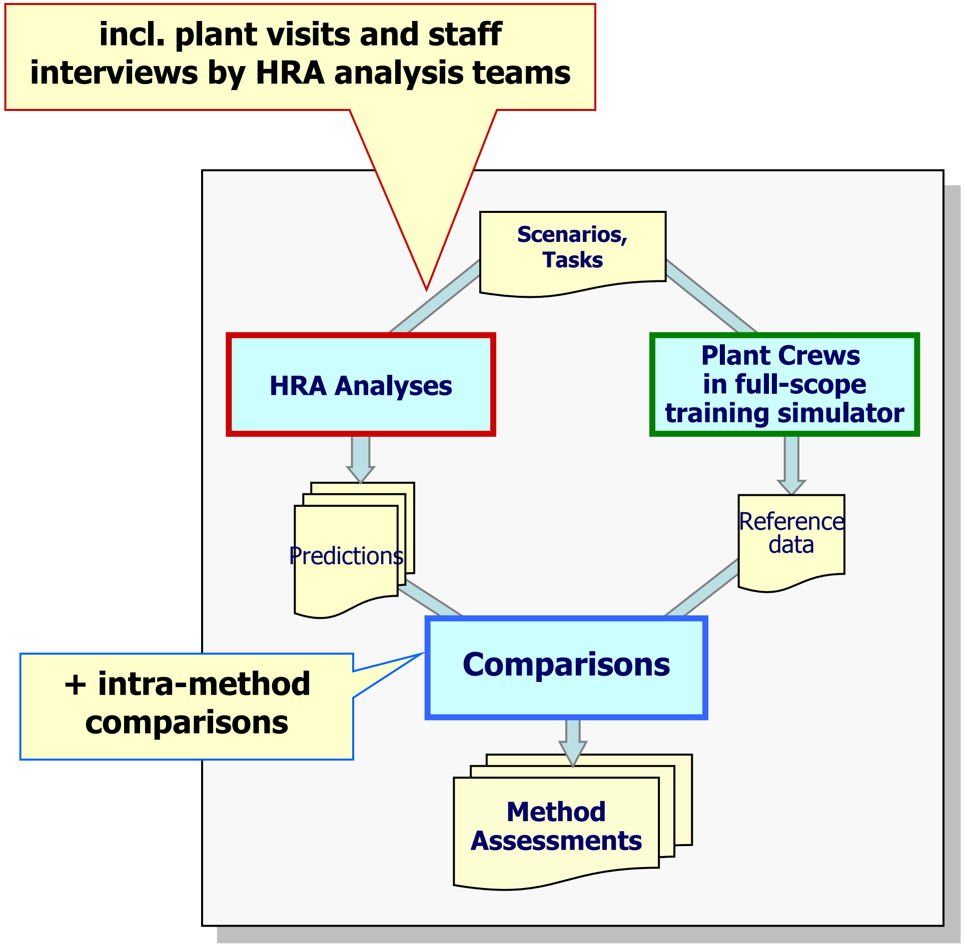
Task #	Task Description	Status
1	Experimental design	Completed, December 2009
2	Development of accident scenarios and human failure events (HFEs)	Completed, December 2009
3	Form and support HRA teams	Completed, April 2010
4	Collect and evaluate the HRA submittals	Completed, November 2010
5	Crew data analysis	Completed, March 2011
6	Comparison of HRA predictions to reference data, initial comparison of analyst teams using the same method, and workshop	Completed, June 2011
7	Perform intra-method comparison and submit initial study findings to PSAM 11	Completed, January 2012
8	Final NUREG/CR report documenting study design, methodology, and results	To be completed, 2013

Study Methodology Overview (1/3)



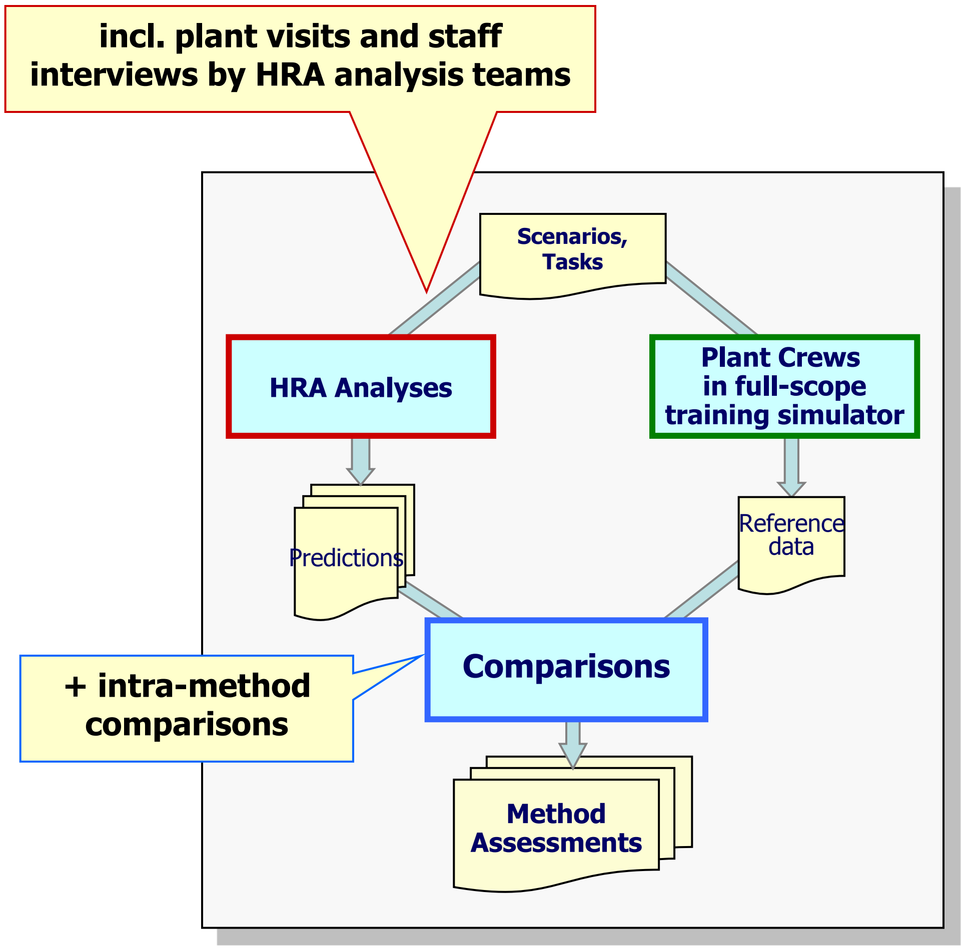
- Scenarios (# of HFEs)
  - LOFW followed by SGTR (3)
  - Loss of CCW & RCP seal water (1)
  - Textbook SGTR (1)
- Nine HRA teams
  - ASEP: 2
  - HCR/ORE & CBDT: 3
  - SPAR-H: 2
  - ATHEANA: 2
- Four US crews & full-scope simulator

Study Methodology Overview (2/3)



- HRA predictions vs. reference data
  - Start with *quantitative* results
  - *Qualitative* predictions are weighted more strongly

Study Methodology Overview (3/3)



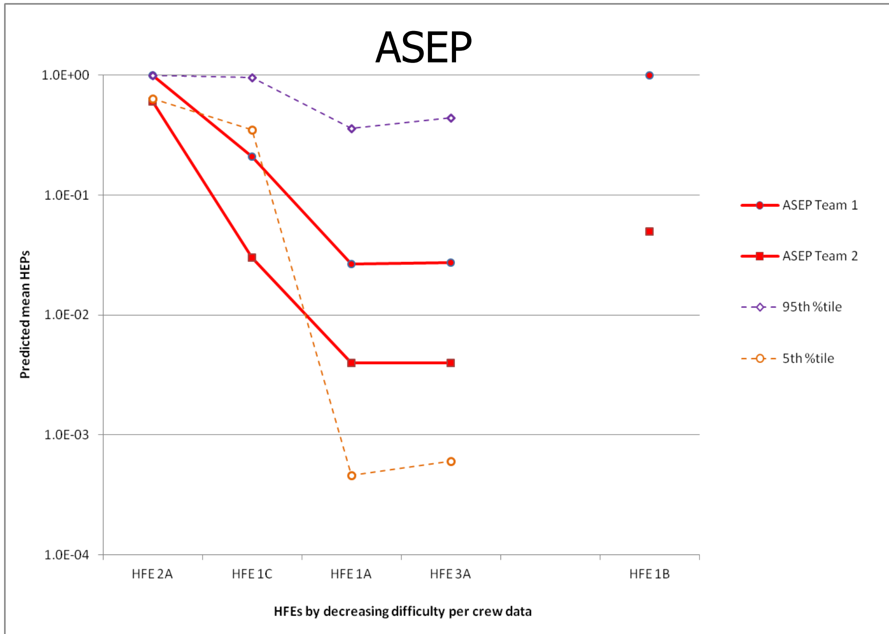
- Intra-method comparison (comparing analyst teams)
  - Differences in qualitative predictions
  - Differences in quantification and ranking of HFEs
  - Differences in the analyses and assumptions
  - Potential contributors to these differences

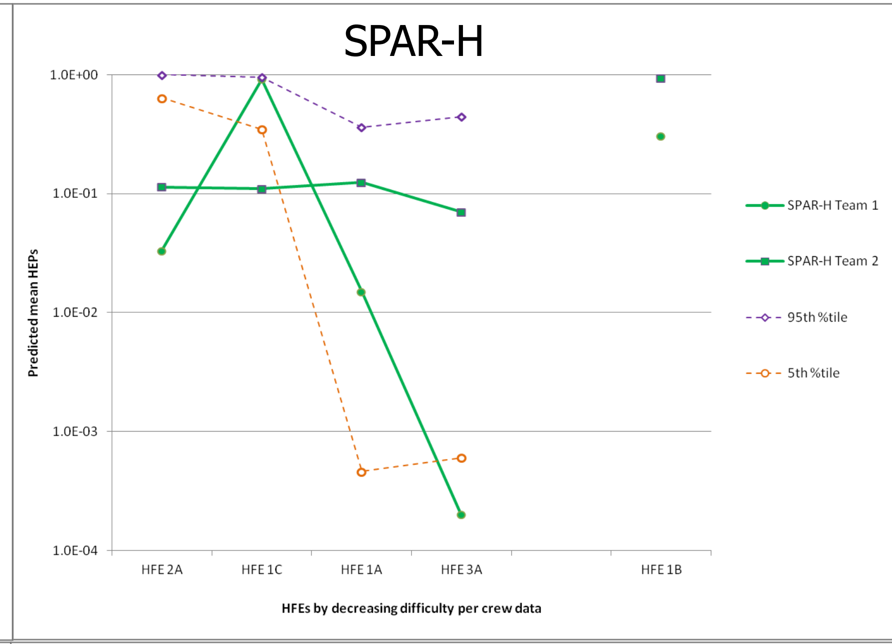
HFE Difficulty Ranking

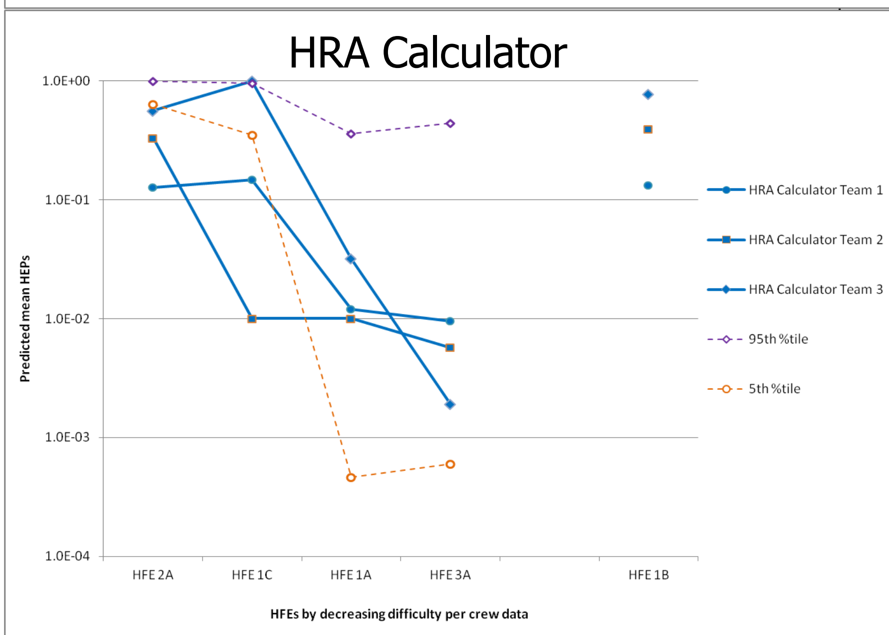
HFE	Task	Difficulty
2A	Stop RCPs and start PDP in Scenario 2	Very difficult
1C	Identify and isolate ruptured steam generator in Scenario 1	Difficult
1A	Start bleed and feed in Scenario 1	Fairly difficult to difficult
3A	Identify and isolate ruptured steam generator in Scenario 3	Easy

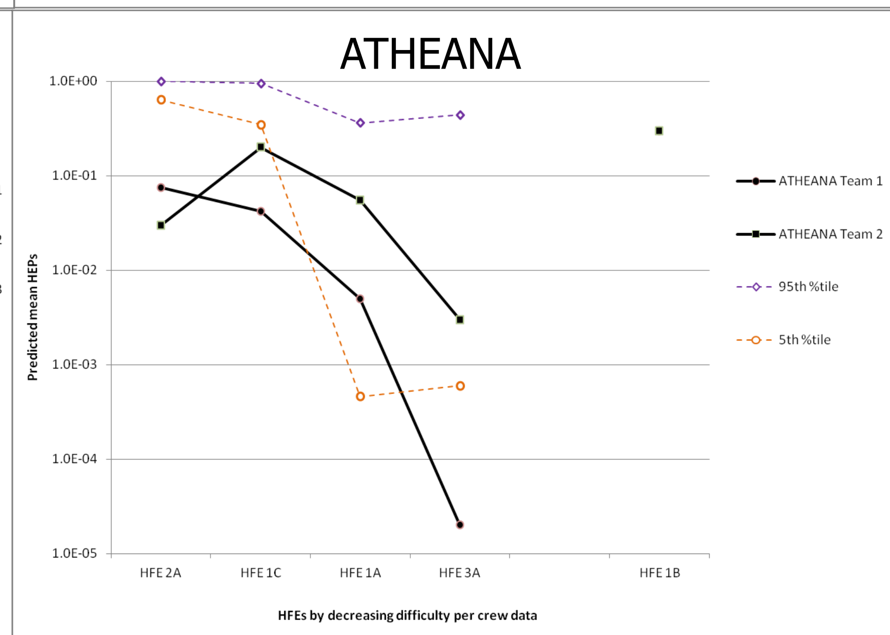
Note: No crew data was available for HFE 1B.

Predicted HEPs with Empirical Bounds









Overview of Quantitative Findings

- Ranking of HFEs moderately improved for most methods compared to International study
- For most HFEs, one order of magnitude difference among predictions from a given method
- Some methods seem more consistent than others
- Many teams underestimated HFE 2A
- All teams ranked HFE 3A easiest, but significant variability across teams, within method

Example Intra-Method Findings (1/4)

- ASEP
  - For Team 1, a detailed review of procedure paths led to better estimate of 'required time' as well as identification of potential issues for HFE success.
    - Limited guidance for reviewing procedure paths and estimating time required. ASEP focuses on evaluation of TRC for diagnosis, which can lead to missing consideration of diagnosis difficulties.
  - Method poorly equipped to quantify diagnosis difficulties. Although analysts may compensate based on experience, it may lead to quantitative differences in results.
  - Teams obtained different HEPs for post-diagnosis actions.
    - Limited guidance on what to include in analyzing post-diagnosis actions
    - ASEP vs. THERP execution modeling (option in ASEP)

Example Intra-Method Findings (2/4)

- HCR/ORE & CBDT
  - Analysts made different judgments in addressing complex scenarios, leading to variability.
    - In situations where procedures do not cope well with complexity and available time is short, analysts have difficulties treating HFE-specific aspects of procedural guidance and diagnosis complexity.
    - Method's basic assumption – operators are able to follow procedures and procedural cues are available – is strongly challenged.
  - Analysts obtained different timing estimates for diagnosis and execution.
    - Limited guidance on what to include in timing analysis.
  - Team 3 decomposed HFEs into subtasks but other teams did not
    - Limited guidance on HFE decomposition.

Example Intra-Method Findings (3/4)

- SPAR-H
  - Team 1 treated HFEs as single tasks while Team 2 decomposed in detail. (Both approaches "consistent" with method). Transitions (decision points) in procedures not treated by Team 1.
    - Limited guidance for qualitative analysis.
  - Detailed decomposition contributed to relatively better qualitative predictions for Team 2, but the good qualitative analysis did not yield better quantitative results.
    - Limited guidance for translating qualitative analysis to quantitative results
    - Analysts' experience with the method.
  - Teams accounted for some factors under different PSFs (e.g., lack of plant cues treated under "complexity" vs. "HMI"), leading to different quantitative impact on HEP.
    - Limited guidance on PSF selection

Example Intra-Method Findings (4/4)

- ATHEANA
  - Substantial differences in how teams performed qualitative analysis and quantification. Team 1's HEPs consistently lower for many HFEs, but too many differences to identify main causes.
  - Team 1 quantified Monte Carlo simulation of a "scenario map" (set of potential plant-crew responses) with expert-elicited probabilities and task duration distributions. Led to clear and traceable link between qualitative analysis and resultant HEPs.
  - Team 2 was more holistic in identifying operational challenges (less emphasis on time and more on potential difficulties) and quite effective even if less detailed. Used point estimates for durations and did not quantify Unsafe Actions (UAs) separately.
    - The implementation of ATHEANA is resource intensive, which can lead to differences in implementation of the method.

Conclusions and Recommendations (1/4)

- Overall, the findings on method strengths and weaknesses from the International Study were confirmed. However, the US Study produced more findings.
- <u>Method effect 1:</u> Variability due to scope and depth of the qualitative analysis (PSF characterization and driver identification)
  - More emphasis on structured, consistent qualitative analysis is needed
  - Coupling to quantification method needs to be assured

Conclusions and Recommendations (2/4)

- Method effect 2: Variability due to inadequate method guidance (lack of clarity, specificity, examples on aspects of the method)
  - Extend guidance on method implementation
  - Test guidance for analyst-to-analyst interpretation
- Analyst effect 1: Analysts make different judgments in compensating for method limitations, leading to variability (e.g., lack of coverage of a given factor or type of performance issue)
  - Method needs modifications
  - Not a guidance issue per se

Conclusions and Recommendations (3/4)

- Analyst effect 2: Variability in information from plant visits and staff interviews
  - Allowing plant visit and staff interviews may have been a contributor to somewhat better HRA predictions, but also caused some problems contributing to variability.
  - Differences in interview skills
  - Teams using the same method made different assumptions and interpretations
  - Guidance is needed for performing interviews and collecting information
- US crews vs. European crews
  - Similar variability in crew performance in both studies
  - Somewhat better HRA predictions in the US Study, but no evidence this was due to a crew effect

Conclusions and Recommendations (4/4)

- Results provide clear evidence of method limitations and indicate specific ways to improve individual methods.
- Essentially all methods have limitations in qualitative analysis and the corresponding link with the quantitative models.
- Identified limitations in HRA practices
  - Estimation of time required
  - Reasonableness checks
  - How to conduct plant interviews
- Creation of a hybrid method combining effective elements and features of existing methods should significantly improve HRA and HRA practices.

Thanks!

Questions and Comments?

Scenarios and HFEs (1/2)

- LOFW followed by SGTR
  - Mis-positioned recirc valve with no indication in the control room
  - Indicated flow from AFW pump on the HSIs masked the fact that no water at all was going to the steam generators
  - HFE 1A: Failure to establish bleed and feed (B&F) within 45 minutes of the reactor trip, given that the crew initiates a manual reactor trip before an automatic reactor trip.
  - HFE 1B: Failure to establish B&F within 13 minutes of the reactor trip, given that the crew does not manually trip the reactor before an automatic reactor trip occurs.
  - HFE 1C: Failure to isolate the ruptured SG and control pressure below the SG PORV setpoint to avoid SG PORV opening. The time window to perform the required actions is estimated to be approximately 40 minutes.

Scenarios and HFEs (2/2)

- Loss of Component Cooling Water (CCW) and Reactor Coolant Pump (RCP) Sealwater
  - Failing distribution panel increased the complexity and masked the status indications
  - Very short time windows
  - HFE 2A: Failure to trip the RCPs and start the Positive Displacement Pump (PDP) to prevent RCP seal LOCA.
- SGTR
  - HFE 3A: Failure to isolate the ruptured SG and control pressure below the SG PORV setpoint before SG PORV opening. The time window to perform the required actions is estimated to be 2 to 3 hours.

Assessment Criteria

Qualitative predictive power – driving factors

- Prediction of observed driving factors and performance issues
- Predictions not supported by the data

Qualitative predictive power – operational expressions

- Prediction of observed failure mechanisms

Quantitative predictive power

- Optimism wrto most difficult HFEs
- Consistency of the ranking of the HFEs (by predicted HEP) with reference difficulty ranking
- Predicted HEPs relative to the confidence/uncertainty bounds of the reference data
- Quantitative differentiation of the HFEs by HEP

Guidance and traceability

Insights for error reduction

Intra-Method Comparison (1/4)

ASEP Team 1	ASEP Team 2
Detailed qualitative analysis beyond method guidance.	Tended to stick relatively close to method guidance.
Detailed analysis of procedure paths. Considered role of procedures in diagnosis.	Not as detailed as Team 1. Only considered whether post-diagnosis actions were covered in procedures.
Included relatively more procedural steps in estimating time required for post-diagnosis actions.	Included relatively less procedural steps in estimating time required for post-diagnosis actions.
Made different decision on what to include in estimating time required for post-diagnosis actions.	Made different decision on what to include in estimating time required for post-diagnosis actions.
Used THERP to quantify post-diagnosis actions per ASEP instructions.	Used ASEP to quantify post-diagnosis actions.
Made different assumptions or interpretations of information from interviews with operators.	Made different assumptions or interpretations of information from interviews with operators.

Intra-Method Comparison (2/4)

Cal. Team 1	Cal. Team 2	Cal. Team 3
HRA methodology implemented with HRA Calculator	HRA methodology implemented with HRA Calculator	CBDT + THERP + ASEP (Did not use actual software). Cognitive contribution is the sum of identification & diagnosis/delay
Did not decompose HFEs.	Did not decompose HFEs.	Decomposed HFEs.
	Misunderstood HEE 1C	
Did not consider recovery.	Did not consider recovery.	Considered recovery.
Quality of documentation varied across teams. Greatest variation was seen in cases where multiple procedural or knowledge-based success paths existed.		
The area of least traceability is the operations story, cues and timing analysis.		
Cognitive vs. execution contributions to final HEPs.		

Intra-Method Comparison (3/4)

Sandia National Laboratories

SPAR-H Team 1	SPAR-H Team 2
Treated HFEs as single tasks. Did not consider transitions in procedures.	Decomposed HFEs to basis events based on break points in procedures.
Qualitative analysis seemed to be built on scenario insight.	Operating experience contributed to qualitative analysis.
Accounted for lack of plant cues under complexity.	Accounted for lack of plant cues under HMI (greater multiplier).
Information from interviews with operators caused optimism in HFE 2A.	Lack of differentiation between HFEs. Did not perform sanity/reasonableness check.
Difficulties in mapping qualitative analysis to PSFs and assigning PSF levels. Traceability depends on documentation.	

Intra-Method Comparison (4/4)

ATHEANA Team 1	ATHEANA Team 2
Detailed qualitative analysis with Monte Carlo Simulation of a detailed scenario map presenting alternative paths. (250 man-hours)	Holistic and streamlined approach to qualitative analysis. Did not consider alternative paths. But effective. (90 man-hours)
Probabilities and task duration distributions elicited from experts. Focused on whether time was available. Duration distributions include potential effects of PSFs.	Quantified HFEs as one task. Point estimates for timing estimates. Less rigor in accounting for delays.
Performed sanity check	Did not perform sanity check
Good traceability due to detailed scenario map and documentation.	Less transparent in translating qualitative information into quantitative estimates.
	Biased by experience to underestimate HFE 2A

SACADA Database for Human Reliability and Human Performance

Y. James Chang, Ph.D.

Human Factors and Reliability Branch
Division of Risk Analysis
Office of Nuclear Regulatory Research

Presented to ACRS Subcommittee
January 16, 2013

HRA Data Program Goal

Bridging the human performance data and the human reliability analysis (HRA) applications

SACADA

- SACADA: Scenario Authoring, Characterization, and Debriefing Application
- Aims at informing human error probabilities (HEPs)
  - Collect licensed operator simulator training data
  - Provide HEPs as well as what's and why's of human errors
- Developed under a memorandum of agreement between NRC and STPNOC signed in March 2011
  - Collected data will not be used for regulatory actions
  - Data are proprietary to the data providers (i.e., plant)
- A web-based database
  - Database located at and maintained by the INL
  - In transition from developmental phase to production phase

Objectives

Objective 1: produce anchor Human Error Probabilities (HEPs)

- Produce anchor HEPs with sufficient contextual information
- Focused data: licensed operator simulator training data
- Tool: SACADA database

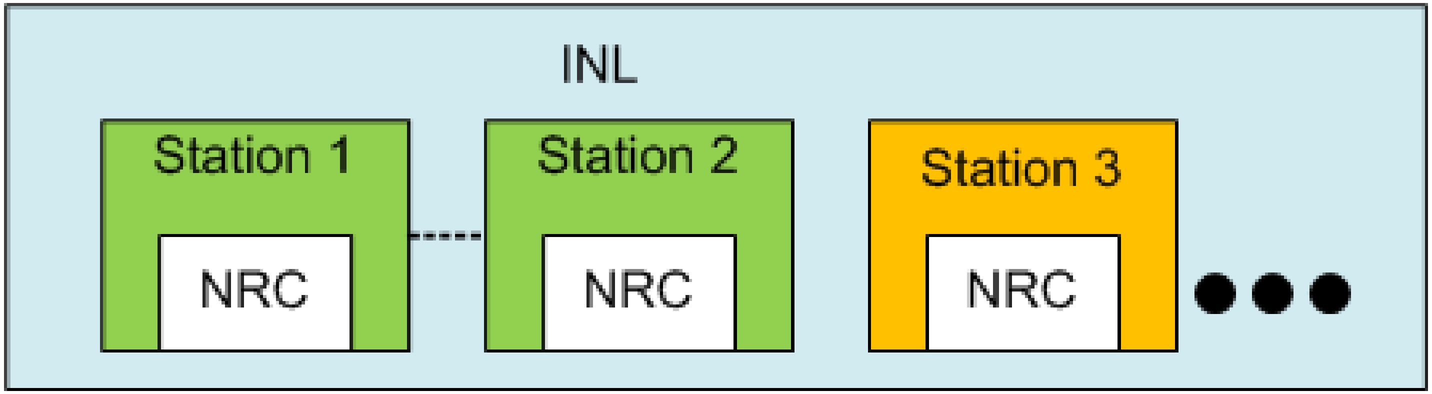
Objective 2: Improve understanding of operators' behavior during accidents or incidents

- Analyze and document operators' responses to complicated events
- Focused means: event investigations and research oriented simulator exercises
- Tool: Improved event timeline

Operation Model
- In Production Phase

- A long term sustainable data collection program
  - All data are entered by plant staff
    - Training department: operator trainers
    - Operations department: operating and staff crews
  - NRC & contractors spot check data quality
- Engage users by
  - Providing data for improving operator performance
  - Streamlining simulator training process
  - Reducing redundant data entry efforts

Accessibility Control



Role-based accessibility control

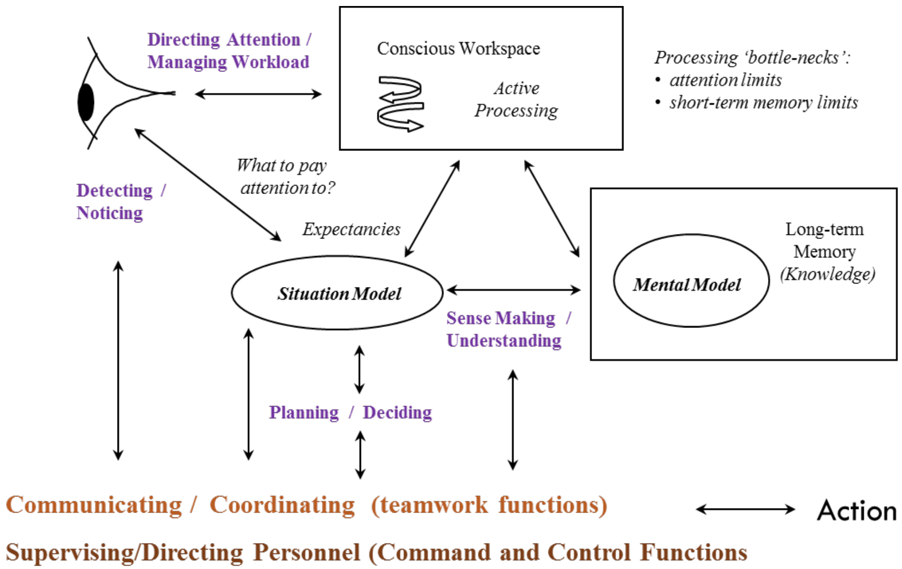
- Roles and accessibilities of plant staff are based on STP's input
- A database administrator for each station

Current Status

- Collected 5 training cycles debriefing data
  - 10 scenarios for 14 crews (Missing some crew-scenarios)
- Outreach:
  - Presented 5 papers at PSAM11/ESREL 2012 conference in June 2012
  - Hosted 5 GoToMeetings in the past 30 days
    - Domestic: INL, SNL, CurtissWright (ScienTech), UMD, STP, NRC and consultants
    - International: Halden, PSI, INER, KAERI, NRI, and Tsing-Hua univ. (China), and Scandpower
  - Presented at the EPRI HRA User Group Meeting 1/2013

Theoretical Foundation

- Macrocognitive function model
  - – Detecting
  - – Understanding
  - – Deciding
  - – Executing actions
  - – Teamwork
  - – Communication
  - – Supervision

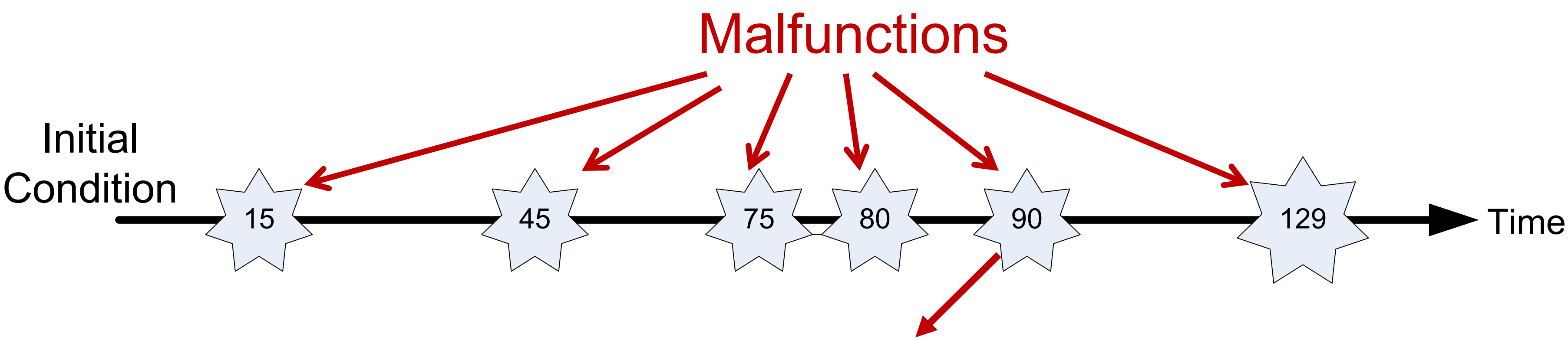


- Same basis as the IDHEAS method

Consider Common Components for Estimating HEPs

- Task analysis (Task decomposition)
- Generic tasks
- Error modes
- Performance shaping factors
- Error recovery
- Task dependency

Simulation Scenarios



Elements

Loss of ECW 1A	
POSITION	**EXPECTED RESPONSE**
CREW	Recognize loss of EW flow to A train.
Crew	Secure ECW pump 1A
SM	Manually trip Diesel Generator prior to any of Diesel Generator trips
Crew	Ensure CCP 1A is in service
Crew	Verifies Natural Circulation
SM	Determines need to cooldown
SM	Declare an Alert HA1/EAL2 due to damage to EW structure or notify ED that escalation is appropriate.

Data Point - Element

A Data Point

Context

Crew Responding to the context

Responding Results

- SAT, SAT+, SATΔ, UNSAT…
- If any, Performance problem classifications and specifics

Characterization

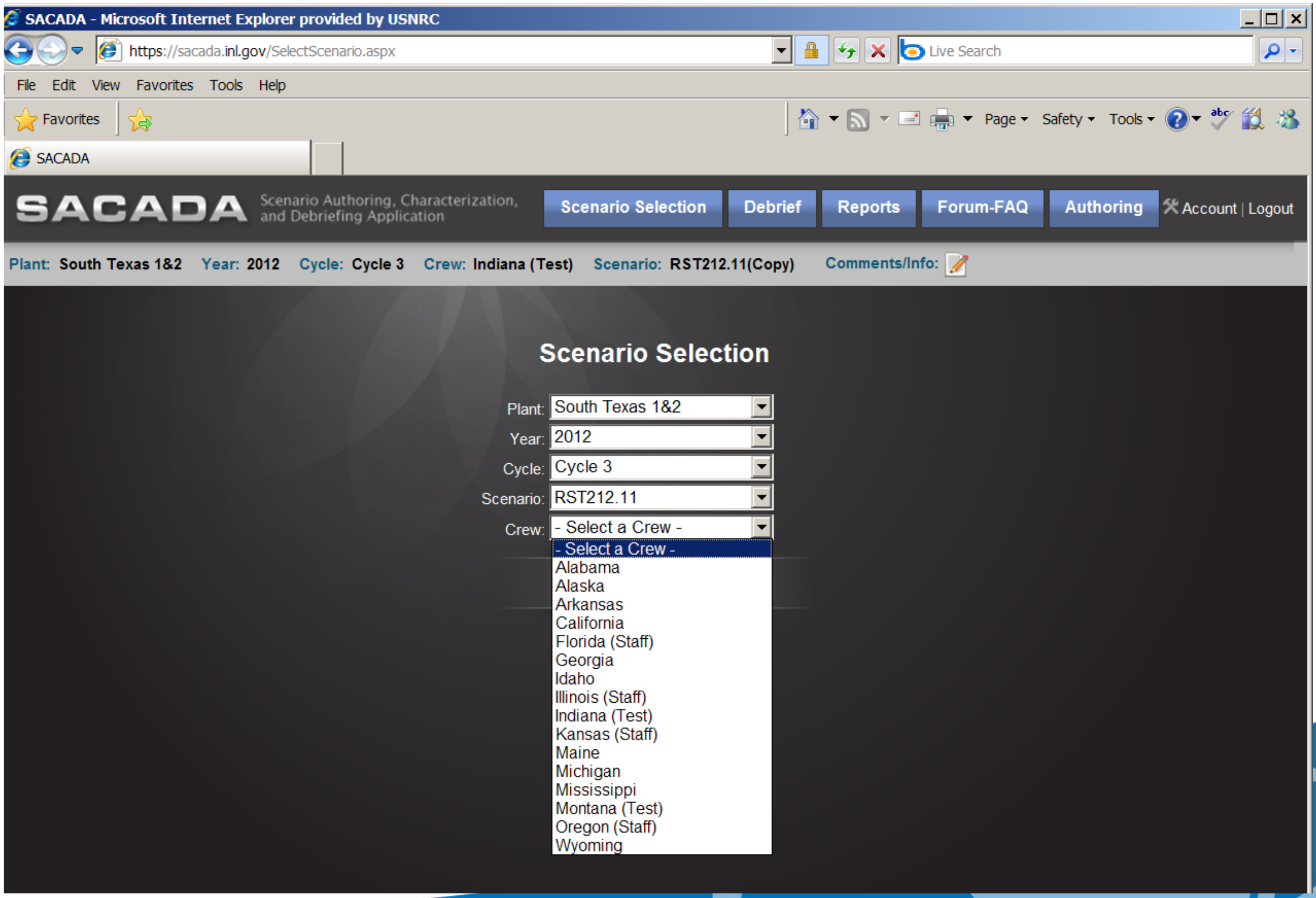
Debriefing

Four Main Functions and Statuses

Function	1. Function Description 2. Performer	Status
Authoring	1. Design scenarios 2. Scenario designers/Operator trainers	STP testing
Characterization	1. Characterize context/challenges 2. Scenario designers/Operator trainers	STP testing
Debriefing	1. Document performance results 2. Crew	Implemented
Data Outputting	1. Data analysis, generate reports, output data 2. Plant staff, NRC staff, and contractors	Need more work

SACADA Login Screen

- For NRC Users



Authoring – Before

- In MS Word

United States Nuclear Regulatory Commission

Protecting People and the Environment

RST 211.07 Rev. 0
Page 1 of 2

POP04-ZO-02 Addendum 4 actions		
POSITION	EXPECTED RESPONSE	
US	START and LOAD Standby Diesel Generator 13 based on Winds In Excess Of 73 MPH within 2 hours.	
CREW	START And LOAD The Remaining Standby Diesel Generators based on Unstable Offsite Electrical Supply Developing AND is Predicted by TDSP	
SM	Notifies NRC Operations Center within one hour and to NRC Region Office as soon as practical when Temporary suspension of Safeguards Security Measures to protect personnel has been approved.	

A loss of Standby Bus 1G		
POSITION	EXPECTED RESPONSE	
US	Enter 0POP04-AE-0001, First Response To Loss Of Any Or All 13.8 KV Or 4.16 KV Bus	
Crew	Ensure SDG 12 & EW 1B has started	
Crew	Ensure B train sequencer is loading the SDG	
Crew	VERFY RCP Seal Cooling: CCW or Seal Injection	
Crew	Ensure CCP has cooling established	
Crew		
US	Evaluate Tech. Spec. 3.8.1.1.e	

LOSS of GRID		
POSITION	EXPECTED RESPONSE	
US	Reenters from beginning 0POP04-AE-0001, First Response To Loss Of Any Or All 13.8 KV Or 4.16 KV Bus	
CREW	ESTABLISH CCW flow to the RCFC's within 30 minutes of the initiation of the event.	

RST 211.07 Rev. 0
Page 2 of 2

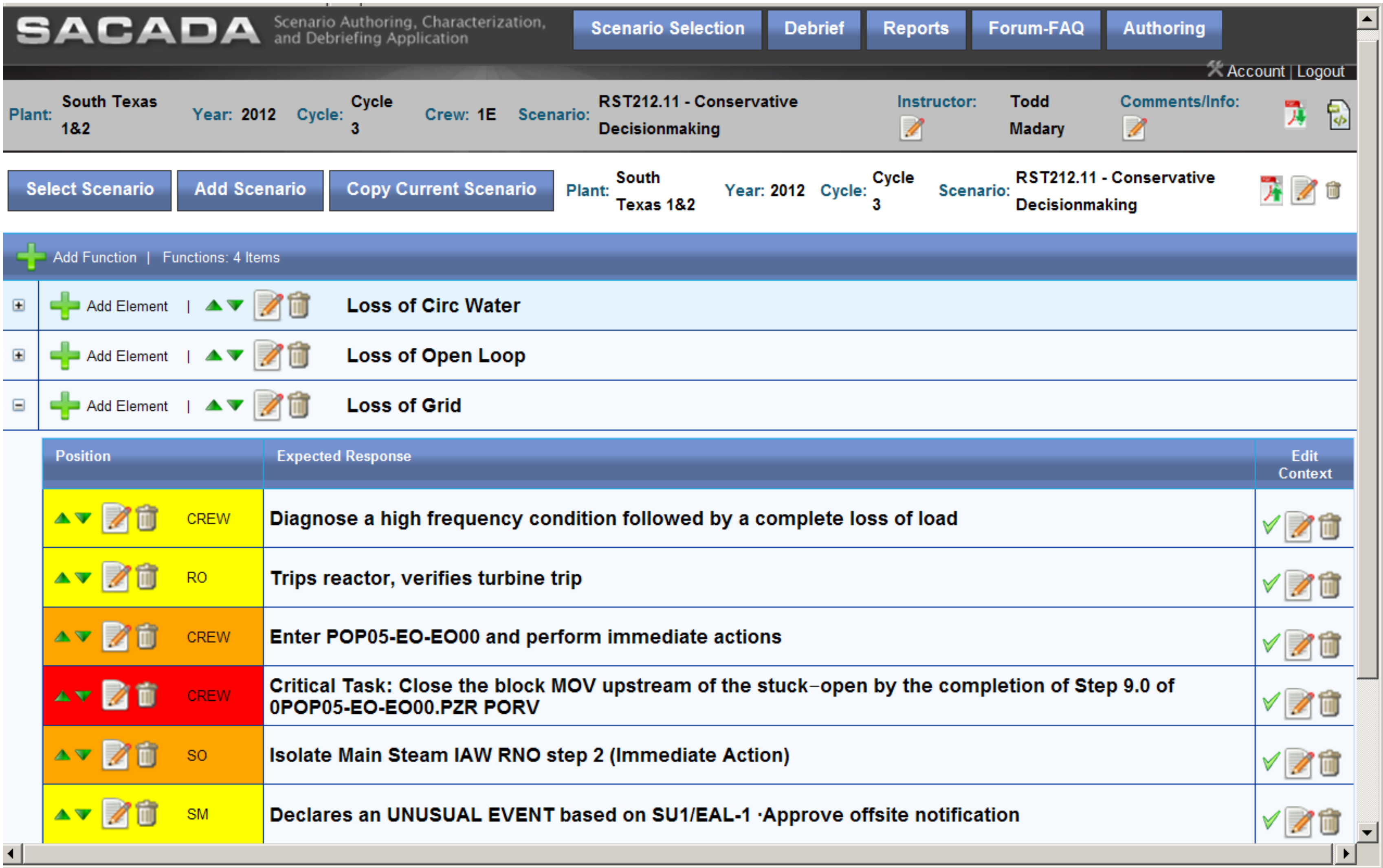
SM	Notifies Emergency Director of conditions.	
Crew	ENSURE one of the following FHB Relief Dampers is OPEN: IAW POP04-AE-0001 CIP "RELIEF SPLY DMPR FV-9500" "RELIEF SPLY DMPR FV-9500A"	
Crew	ENSURE ALL Dilution Flow Paths – ISOLATED	

Loss of ECW 1A		
POSITION	EXPECTED RESPONSE	
CREW	Recognize loss of EW flow to A train.	
Crew	Secure ECW pump 1A	
CT	Manually trip Diesel Generator prior to any of the following occurring: Diesel Generator tripping	
	SAFETY SIGNIFICANCE -- Failure to manually start the ECW pump in an operating safeguards train represents a "demonstrated inability by the crew to: Recognize a failure/incorrect auto actuation of an ESF system or component Effectively direct/manipulate ESF controls" Additionally, under the postulated plant conditions, failure to manually start at least the minimum required number of ECW pumps (when it is possible to do so) is a "violation of the facility license condition."	
Crew	Ensure CCP 1A is in service	
Crew	Verifies Natural Circulation	
SM	Determines need to cooldown	
FYI The ED would declare	Declare an Alert HA1/EAL2 due to damage to EW structure or notify ED that escalation is appropriate.	

Loss of TSC Diesel		
POSITION	EXPECTED RESPONSE	
SM	Confers with the Emergency Director on response to loss of TSC.	

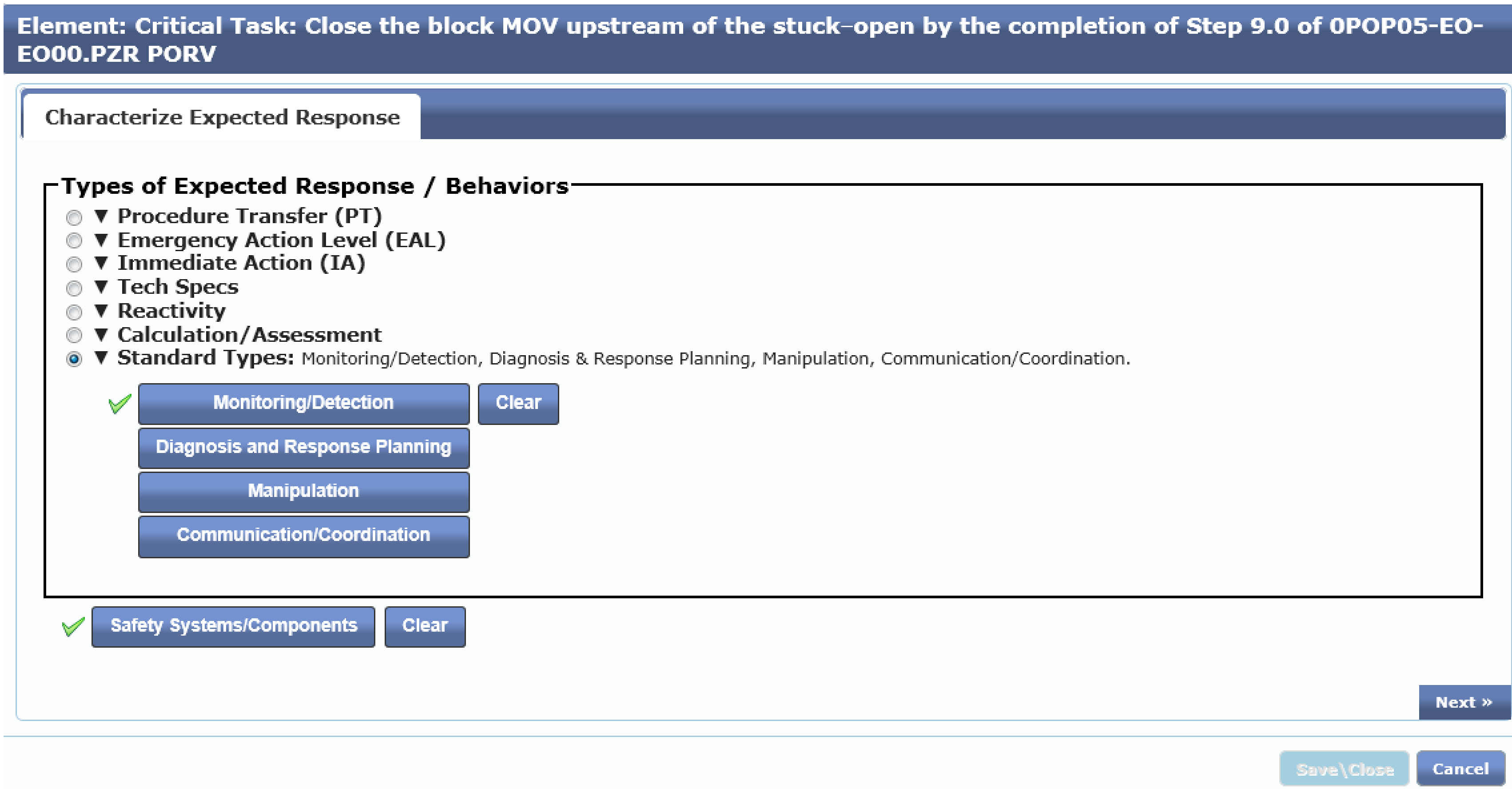
Authoring - Now

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment



Characterization
- Entry Screen

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment



Characterization

- Monitoring/Detecting - Indicators

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

Element: Critical Task: Close the block MOV upstream of the stuck-open by the completion of Step 9.0 of 0POP05-EO-EO00.PZR PORV

Monitoring/Detection: Select Type of Input to be Detected

Monitoring/Detection

- ▼ **Alarms**
- ▼ **Indicators**

Simulator Tag (or Component ID):

Detection Mode

- **Procedure Directed Check:** Procedure directs crew to check a specific indicator or parameter.
- **Procedure Directed Monitoring:** Procedure instructs the crew to monitor a parameter.
- **Knowledge Driven Monitoring:** Crew decides to monitor an indication or parameter based on knowledge of the situation or expectation of change in the parameter.
- **Awareness/Inspection:** Non-procedurally directed monitoring or awareness of plant parameters.

Individual Indicator

- **Slight:** Requires some effort to detect the change.
- **Distinct Change:** Prominent and readily detected if looked at.

Mimics/Display etc.

Expectation on the indication change.

- **No Mimics:** Requires operator to rely on memory.
- **Small Indications:** Can be read only from a close distance.
- **Similar Displays:** Multiple identical displays in the same bank of control panel.

Other Challenges

Any Additional Challenges Not Covered Above:

Overarching Issues

Workload

- **Normal:** All crew members have peer check and backup.
- **Concurrent Demands:** One crew member has own task with no backup; all others have normal peer check and backup.
- **Multiple Concurrent Demands:** Overloaded, no peer check. Everyone has their own task with no backup.

Time Criticality

- **Expansive Time Available**
- **Nominal Time Available**
- **Barely Adequate Time Available:** e.g., high-tempo, time-pressured tasks.

Other Overarching Issues

- **Coordination:** Requires close coordination with on-site personnel.
- **Multiple Demands:** Multiple competing demands on attention/distractions.
- **Memory:** Demands on memory.

Characterization

- Diagnosis & Response Planning

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

Element: Enter POP05-EO-EO00 and perform immediate actions

Diagnosis and Response Planning: Contextual Factors

Diagnosis Basis
- (selected) **Procedure:** The diagnosis is driven by procedures or other guidance.
- **Skill/STAR:** Skill-driven diagnosis; without procedure, operator can perform diagnosis from memory.
- **Knowledge Based:** Knowledge-based diagnosis (no procedure applicable; crew relies on engineering or technical knowledge and operating experience).

Expectation of Alarm/Indication Change
- (selected) **Expected:** Given their understanding of current plant status (including systems out for maintenance or testing), the alarm or indication change is expected.
- **Not Expected:** Operators do not anticipate this alarm or change in indications.
- **Not Applicable**

Information Specificity
- (selected) **Specific:** Alarm/alarm pattern/indication(s) point the the specific system problem.
- **Not Specific:** The alarm(s) and/or indication(s) do not directly point to the specific system problem, which requires operator cognitive effort to integrate the information and identify the specific system problem (e.g., multiple cues needed for diagnosis).
- **Not Applicable**

Information Quality
- [ ] **Missing Information:** Includes masked information.
- [ ] **Misleading Information:** Information points to an incorrect diagnosis.
- [ ] **Conflicting Information:** Information points to more than one possible diagnosis, or conflicts with other alarms or indications.

Information Integration
- [ ] **Timing of Information:** Includes slow information feed or delayed information.
- [ ] **Ambiguous Information:** Information provided by system is vague, unclear, or does not point to the nature of the problem.
- [ ] **Integration Required:** Integration of multiple pieces of information is required for diagnosis.

Familiarity
- (selected) **Standard:** Crew has previously trained on this challenge.
- **Novel:** This involves a change to the way the challenge is addressed, such as a new procedure, scenario, or role.
- **Anomaly:** Standard training must be adapted to fit an anomalous situation (e.g., the procedures do not cover the circumstances).

Other Challenges

Any Additional Challenges Not Covered Above:

Overarching Issues

Workload
- **Normal:** All crew members have peer check and backup.
- (selected) **Concurrent Demands:** One crew member has own task with no backup; all others have normal peer check and backup.
- **Multiple Concurrent Demands:** Overloaded, no peer check. Everyone has their own task with no backup.

Time Criticality
- **Expansive Time Available**

Characterization
- Safety Components/Systems

Element: Critical Task: Close the block MOV upstream of the stuck-open by the completion of Step 9.0 of 0POP05-EO-EO00.PZR PORV

Safety Systems: Impacted

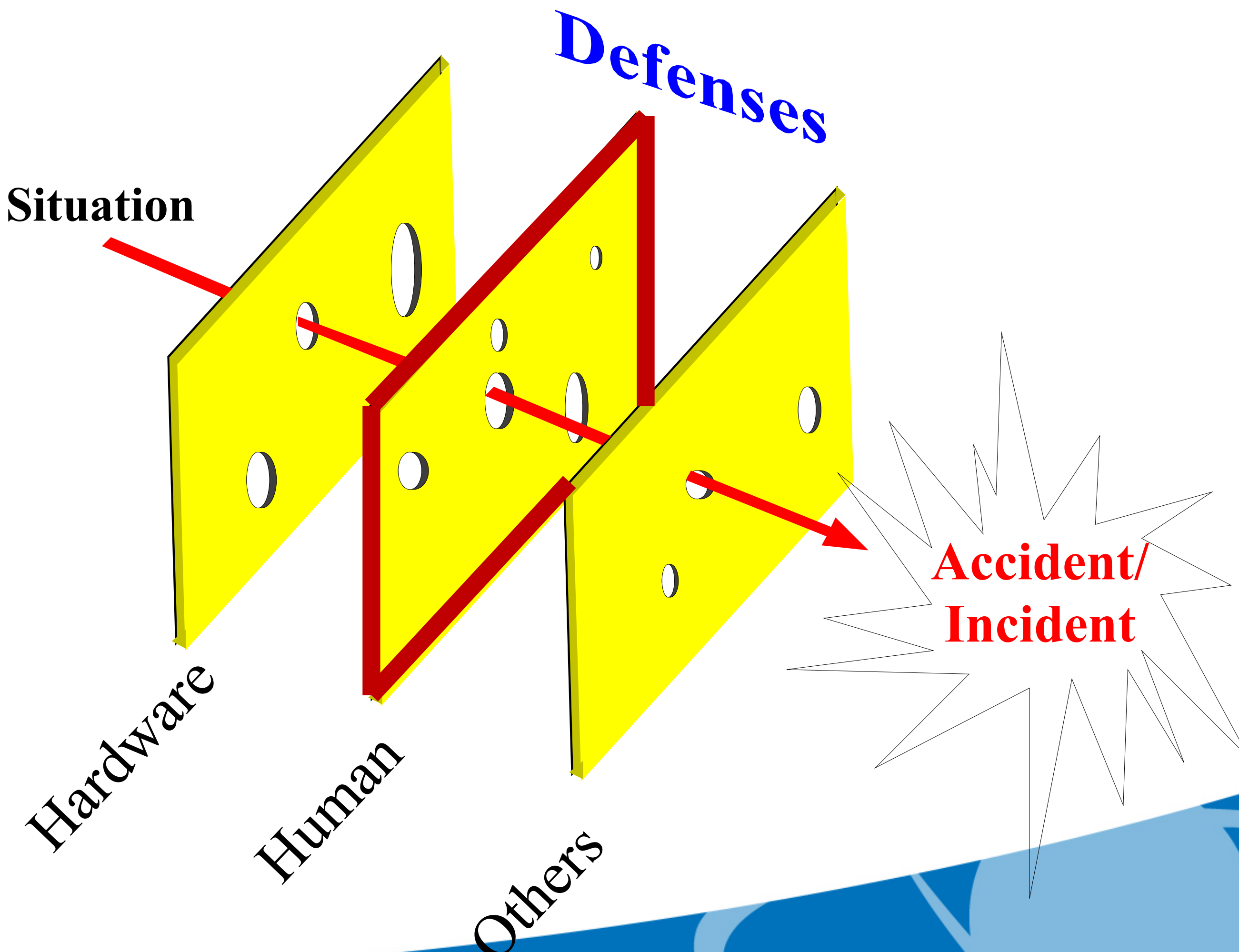
- [ ] ▼ Reactivity Control
- [ ] ▼ Reactor Coolant System Inventory Control
- [x] ▼ Reactor Pressure Control
  - [ ] Emergency Core Cooling System
  - [x] Pressurizer Pressure Control System
- [ ] ▼ Heat Removal From Reactor Core
- [ ] ▼ Containment Integrity
- [ ] ▼ Electrical
- [ ] ▼ Instrumentation
- [ ] ▼ Plant Service System
- [ ] ▼ Radioactivity Release

Next »

Cancel

Concept of Debriefing Screens

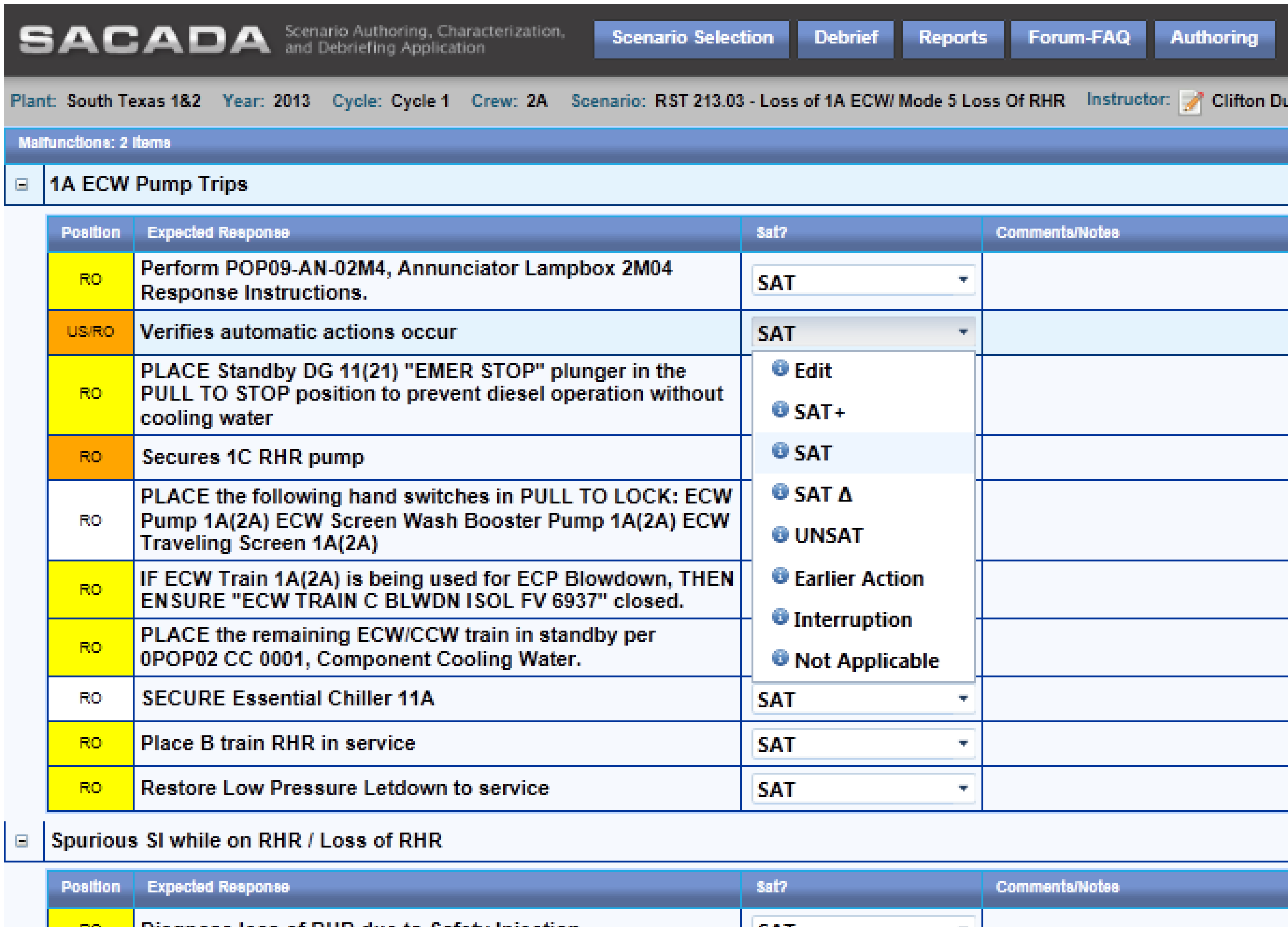
- The Swiss Cheese Concept



Debriefing

- Entry Screen

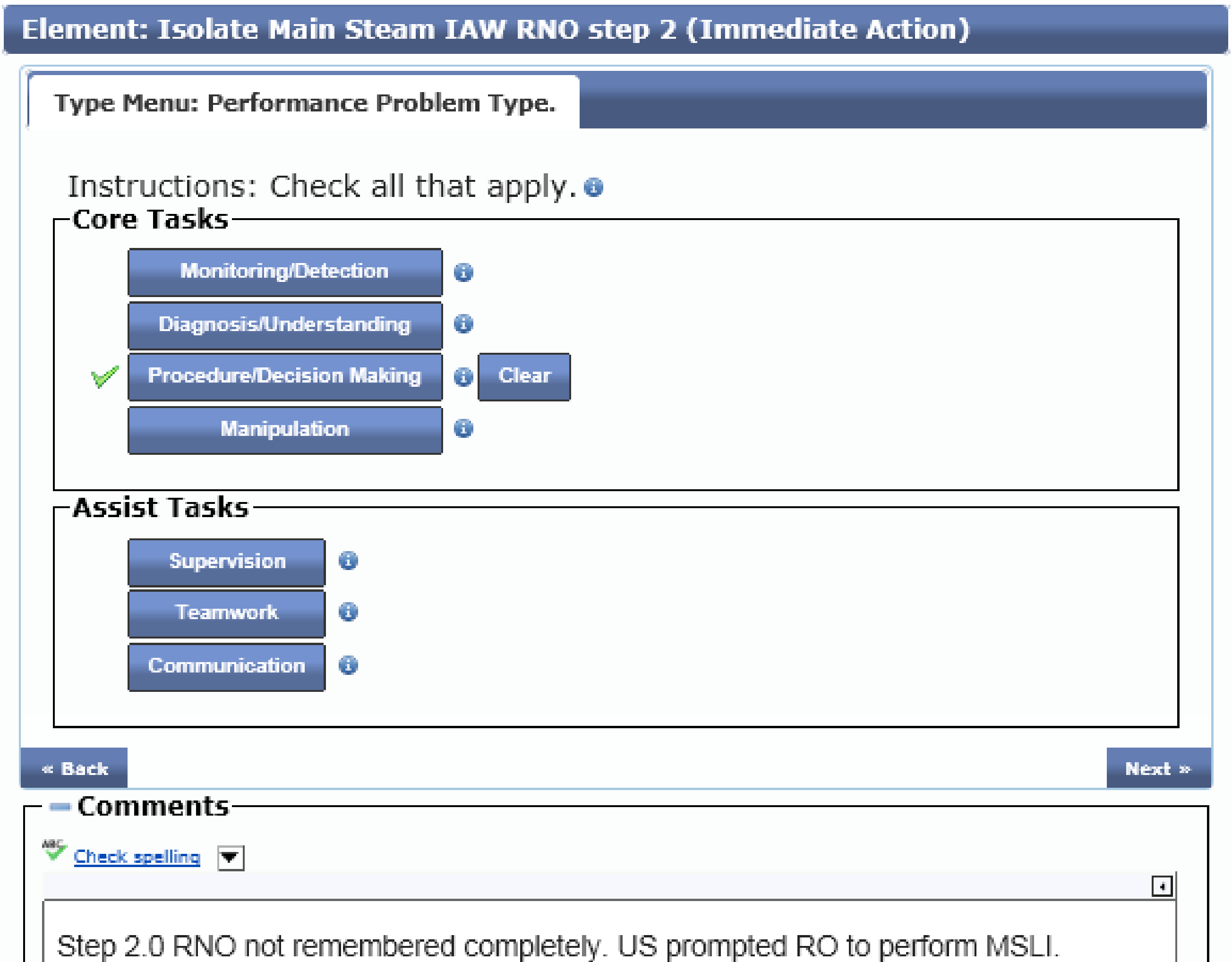
U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment



Debriefing

- Types of Performance Deficiency

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment



Debriefing

- Specifics of Performance Deficiency

United States Nuclear Regulatory Commission
Protecting People and the Environment

Element: Isolate Main Steam IAW RNO step 2 (Immediate Action)

Procedure/Decision Making: Performance Problem Sub-Type.

Procedure/Guidance Availability

- (●) **Relevant Procedural Guidance Available**
  - ( ) **Not Consulted:** Failed to consult available procedure.
  - (●) ▼ **Following Problem:** Trouble following/using procedure.
    - ( ) **Wrong:** Used or transferred to a wrong procedure.
    - (●) **Misinterpreted:** Misinterpreted procedure instruction.
    - ( ) **Deviated:** Incorrectly decided to deviate from procedure.
    - ( ) **Specific/focused Error:** Misinterpreted, omitted or incorrectly performed one or more substeps of a single step.
    - ( ) ▼ **Usage Rules:** Violated general usage rules.
    - ( ) **Other:** Explain.
  - ( ) ▼ **Not Adapted:** Failed to adapt to the situation.
- ( ) **Relevant Procedure/Guidance Not Available**

« Back | Next »

Comments

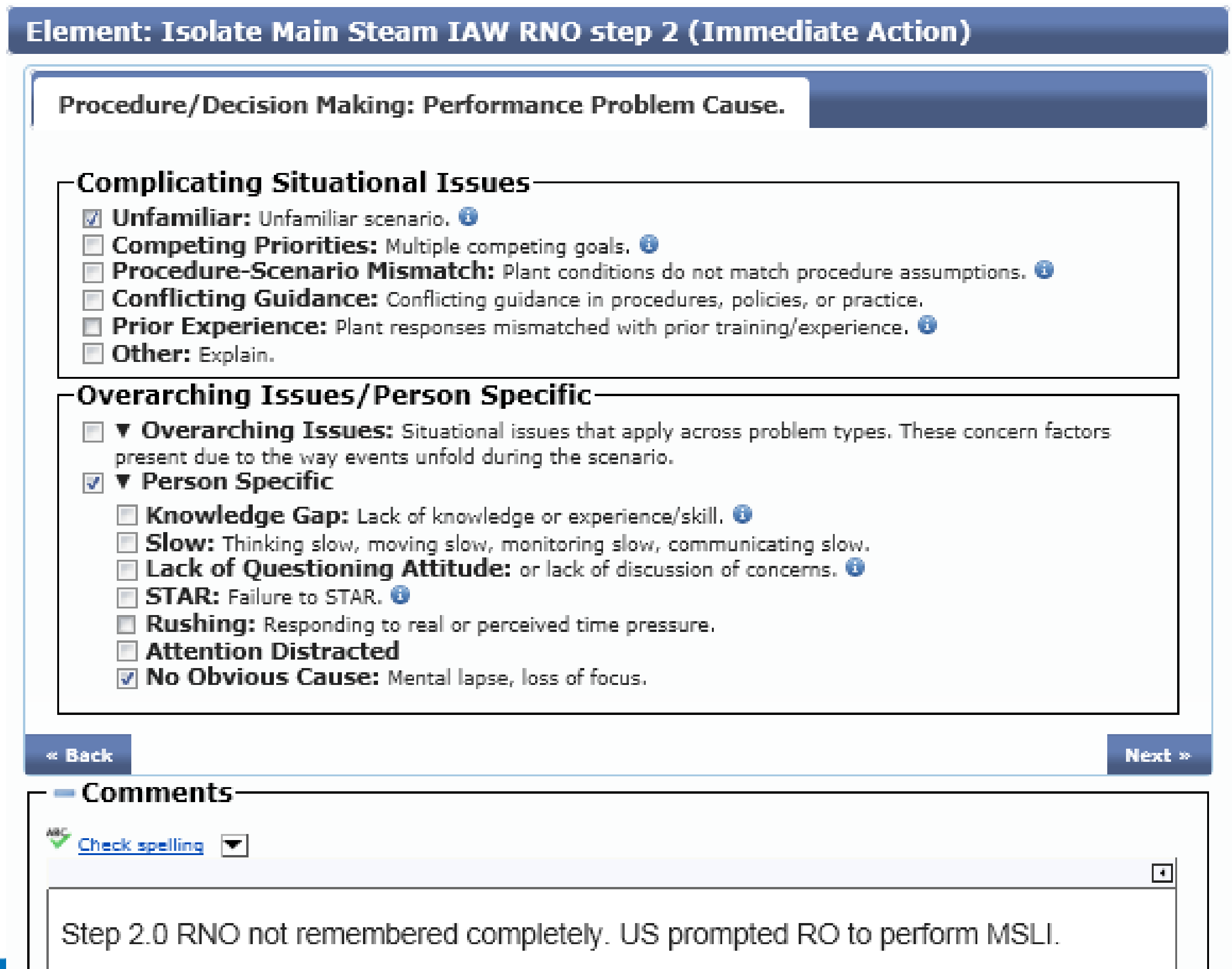
Check spelling

Step 2.0 RNO not remembered completely. US prompted RO to perform MSLI.

Debriefing

- Causes of Performance Deficiency

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment



Element: Isolate Main Steam IAW RNO step 2 (Immediate Action)

Procedure/Decision Making: Performance Problem Cause.

Complicating Situational Issues

- [x] **Unfamiliar:** Unfamiliar scenario.
- [ ] **Competing Priorities:** Multiple competing goals.
- [ ] **Procedure-Scenario Mismatch:** Plant conditions do not match procedure assumptions.
- [ ] **Conflicting Guidance:** Conflicting guidance in procedures, policies, or practice.
- [ ] **Prior Experience:** Plant responses mismatched with prior training/experience.
- [ ] **Other:** Explain.

Overarching Issues/Person Specific

- [ ] ▼ **Overarching Issues:** Situational issues that apply across problem types. These concern factors present due to the way events unfold during the scenario.
- [x] ▼ **Person Specific**
  - [ ] **Knowledge Gap:** Lack of knowledge or experience/skill.
  - [ ] **Slow:** Thinking slow, moving slow, monitoring slow, communicating slow.
  - [ ] **Lack of Questioning Attitude:** or lack of discussion of concerns.
  - [ ] **STAR:** Failure to STAR.
  - [ ] **Rushing:** Responding to real or perceived time pressure.
  - [ ] **Attention Distracted**
  - [x] **No Obvious Cause:** Mental lapse, loss of focus.

« Back

Next »

Comments

Check spelling

Step 2.0 RNO not remembered completely. US prompted RO to perform MSLI.

Debriefing

- Error Recovery

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

Element: Isolate Main Steam IAW RNO step 2 (Immediate Action)

Recovery

Recovery Status

- (selected) ▼ **Immediate Recovery:** Error was recovered with minimal impact on crew performance.

  How Recovered
  - **Self Recovery:** Error was caught and recovered by individual who made the error.
  - **Peer Check:** Intervention by peer checker prevented error.
  - (selected) **Team Recovery:** Questioning attitude by team member realigned crew.
  - **Supervision:** Intervention by supervisor identified or prevented error.
  - **STAR:** Review of system response uncovered error.
- ▼ **Delayed Recovery:** Scenario objectives met but mistakes resulted in delay or confusion with noteworthy impact on crew performance.
- **Unrecovered:** Error was never recovered.

« Back | Next »

Comments

Check spelling

Step 2.0 RNO not remembered completely. US prompted RO to perform MSLI.

Debriefing

- End Results/Effects on Scenario

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

Element: Isolate Main Steam IAW RNO step 2 (Immediate Action)

End Result

End Result From Perspective of What Plant Needs

- (●) **None:** Error may have affected crew performance but did not impact control of the plant.
- ( ) **Required Action Not Taken:** Failed to take required action (did not attempt action).
- ( ) ▼ **Incorrect Timing**
- ( ) ▼ **Process Control Complication**
- ( ) ▼ **Executed undesired action:** Incorrect action from perspective of what plant needs or requires.

« Back | Next »

Comments

Check spelling

Step 2.0 RNO not remembered completely. US prompted RO to perform MSLI.

Debriefing

- Remediation/Follow Through

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

Element: Isolate Main Steam IAW RNO step 2 (Immediate Action)

Remediation and/or Follow Through Items

Refresh Preview

Type	Sub Type	Cause
Procedure/Decision Making	Proc Guidance Available: Following Problem: Misinterpreted	Unfamiliar Person Specific: No Obvious Cause

Recovery	End Result	Remediation
Immediately Recovered: Questioning Attitude	No impact on plant control	

Hide Preview

Export to Crew Notebook

- [ ] **Export:** Add this element to the list of items to export to the crew notebook.

Remediation (Data not shared with NRC)

- ○ **No Follow through/Remediation Required.**
- ◉ ▼ **Follow through/Remediation Required.**
  - [ ] **Performance gap:** Individual/Team performance gap that should be addressed.
  - [ ] ▼ **Simulator Deficiency:** Fix needed to simulator equipment or code.
  - [ ] ▼ **Condition Report**
  - [ ] **Other:** Explain.

« Back

Debriefing

- Performance Results

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

SACADA Scenario Authoring, Characterization, and Debriefing Application

Scenario Selection | Debrief | Reports | Forum-FAQ | Authoring

Account | Logout

Plant: South Texas 1&2 Year: 2012 Cycle: Cycle 3 Crew: 1E Scenario: RST212.11 Instructor: Todd Madary Comments/Info:

Malfunctions: 4 Items	Malfunction Comments	Edit
Loss of Circ Water		
Loss of Open Loop		
Loss of Grid		

Position	Expected Response	Sat?	Comments/Notes	Edit
CREW	Diagnose a high frequency condition followed by a complete loss of load	Not Applicable		
RO	Trips reactor, verifies turbine trip	SAT		
CREW	Enter POP05-EO-EO00 and perform immediate actions	SAT	US Command and Control. US should order reactor trip when appropriate.	
CREW	Critical Task: Close the block MOV upstream of the stuck–open by the completion of Step 9.0 of 0POP05-EO-EO00.PZR PORV	SAT	US should attempt to shut PORV before the block valve. This step was recognized and performed following IAs and before read through.	
SO	Isolate Main Steam IAW RNO step 2 (Immediate Action)	SAT Δ	(see below)	

Type	Sub Type	Cause
Procedure/Decision Making	Proc Guidance Available: Following Problem: Misinterpreted	Unfamiliar Person Specific: No Obvious Cause

Recovery	End Result	Remediation
Immediately Recovered: Questioning Attitude	No impact on plant control	Technical Knowledge: Formal Remediation: Responsible Individual: Steve Kaspar

Hide Preview

Debriefing Summary

- Performance results
  - SAT, SATΔ, UNSAT, and Earlier Actions, etc.
- If SATΔ or UNSAT
  - Types of performance deficiency
  - Specifics of performance deficiency
  - Causes of performance deficiency
  - Recovery
  - Effects on scenario
  - Remediation
- If Earlier Actions: task dependency

Debriefing

- Tasks Dependency

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

Loss of Open Loop — Effect of fouling not considered for Open Loop System until alarms were received.

Position	Expected Response	Sat?	Comments/Notes
RO	Check Open Loop Load temperatures	Not Applicable	
US	Enter POP04-OC-0001	SAT	
RO	Direct the TGB watch to isolate RCB & MAB Chillers	Earlier Action	Earlier error: Enter POP04-OC-0001 Reactor manually tripped when lowering Open Loop pressure was observed. Loss of offsite power coincident with Reactor trip.
US	Dispatch Plant Operator verify discharge strainer DP, attempt a strainer swap	SAT	
US	Direct securing of standby Open Loop pump	Earlier Action	Earlier error: Enter POP04-OC-0001 Reactor manually tripped when lowering Open Loop pressure was observed. Loss of offsite power coincident with Reactor trip.
US	Perform a Main Turbine load reduction	Earlier Action	Earlier error: Enter POP04-OC-0001 Reactor manually tripped when lowering Open Loop pressure was observed. Loss of offsite power coincident with Reactor trip.
SM	Notify Co-Owners	Earlier Action	Earlier error: Enter POP04-OC-0001 Reactor manually tripped when lowering Open Loop pressure was observed. Loss of offsite power coincident with Reactor trip.
RO	Maintain <300 MVARS	Earlier Action	Earlier error: Enter POP04-OC-0001 Reactor manually tripped when lowering Open Loop pressure was observed. Loss of offsite power coincident with Reactor trip.

Prospective Data for HRA

- Statistics of
  - Critical tasks and PRA tasks
  - Actions on safety component/system
  - Time information
- Context Similarity based HEPs
  - Use context similarity to generate HEPs
  - Increase data usability
- Others, e.g.,
  - Research oriented analysis, e.g., relation between context, error, and causes
  - Task dependency

Generate Context Similarity Based HEPs (1/3)

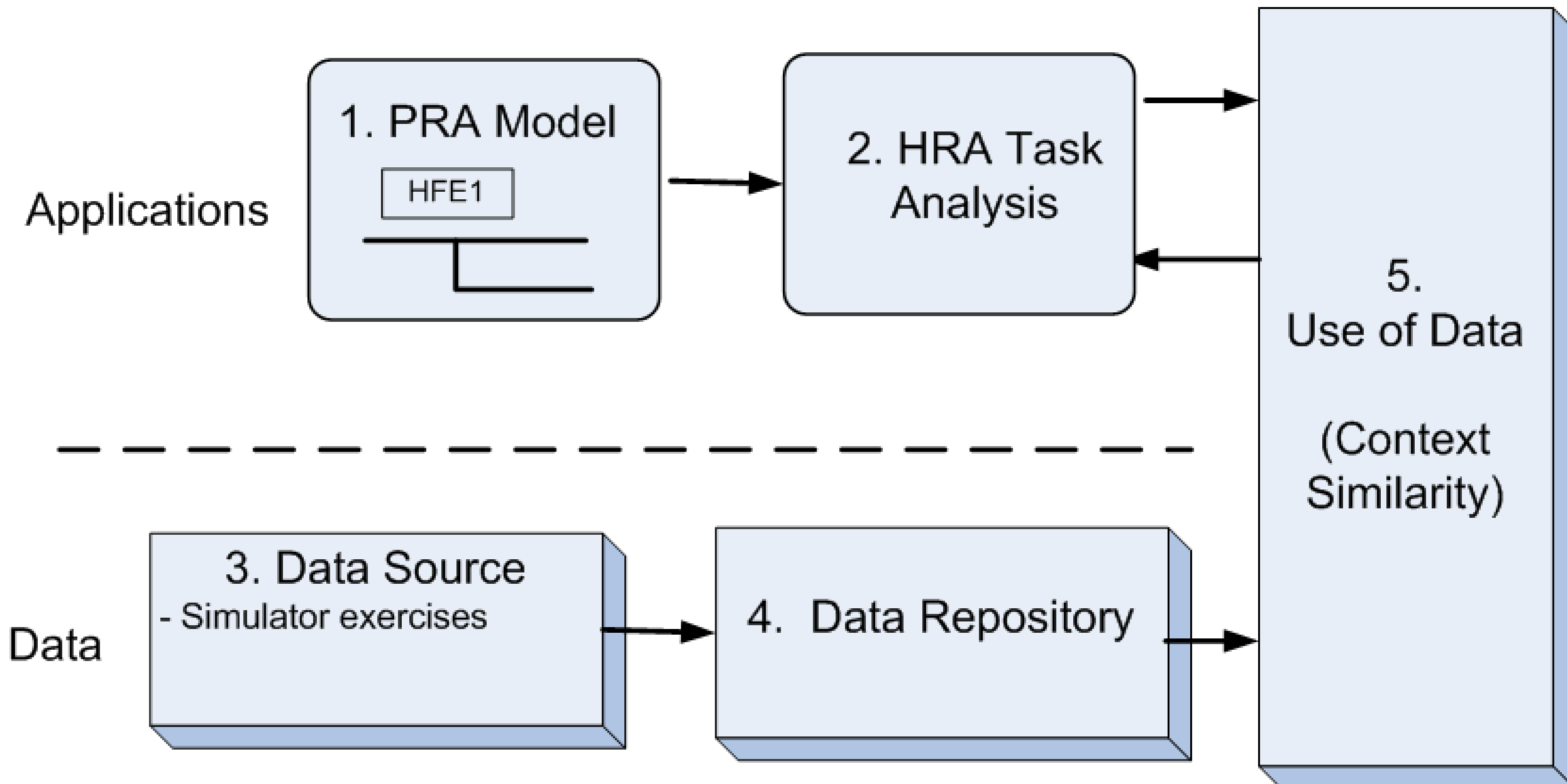
- Hypothesis
  - Tasks with similar context have similar HEPs
- Context of
  - Detecting cues
  - Diagnosis and response planning
  - Manipulation
  - Communication and coordination

Generate Context Similarity Based HEPs (2/3)

$$\text{HEP}(\text{Context}) = \frac{\#\text{ of Failures}(\text{Context})}{\#\text{ of Response Opportunities}(\text{Context})}$$

Generate Context Similarity Based HEPs (3/3)

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment



Data Quantity (Rough Estimates)

#	Unit
104	Unit
7	Crews/Unit (Operating crew & staff crew)
5	Training weeks/crew-year
4	Scenarios/Training week
4	Malfunctions/scenario
5	Elements/Malfunction

291,200 Elements (Data Points)/Year

Simulator Data Based HEPs for HRA

- Gaps to be addressed, e.g.,
  - Success criteria
  - Operator mindset (simulated environment)
  - Crew composition
- Mixed effects on HEP values
- Avoid blind use
  - Need case-by-case determination on appropriateness
  - e.g., Feed-and-bleed – main HEP driver may not exist in simulation context

Summary

- For human reliability and human performance
- Piloting with STP
  - Gained STP crews' acceptance
  - Gained STP managers and trainers' support
  - Improve trust to include simulator exam data
- Address gap in using simulator data for HRA
- Outreach to more plants/organizations to increase data generation rate

CycleYear	2012				
CycleCount	(All)				
Scenario_ID	(All)				

Count of ScenarioType_ID	**Column Labels**				
	Not Staff Crew				
Row Labels	**1_UnSat**	**2_Delta**	**3_Sat**	**4_Plus**	**5_NA_EarlrErr_Unrchd_Intrpt**
RST 212.18	**2**	**6**	**176**	**4**	**4**
RST212.08	**1**	**3**	**232**	**4**	
RST212.11	**2**	**15**	**138**	**4**	**41**
RST212.12	**1**	**6**	**190**	**1**	**10**
RST212.13	**3**	**5**	**287**	**2**	**3**
RST212.14	**2**	**3**	**225**	**1**	**9**
RST212.15	**2**	**6**	**307**	**4**	**11**
RST212.16	**1**		**210**	**2**	**27**
RST212.20			**35**		**1**
Grand Total	**14**	**44**	**1800**	**22**	**106**

Not Staff Crew Total	**Staff Crew**				
	1_UnSat	**2_Delta**	**3_Sat**	**4_Plus**	**5_NA_EarlrErr_Unrchd_Intrpt**
192	**2**	**3**	**66**	**1**	
240		**1**	**69**	**2**	
200	**2**	**9**	**51**		**18**
208	**2**	**1**	**75**		
300	**1**	**2**	**57**		
240		**2**	**64**		**6**
330	**1**	**4**	**89**	**3**	**2**
240		**2**	**61**		**9**
36			**24**	**1**	**2**
1986	**8**	**24**	**556**	**7**	**37**

Staff Crew Total	Grand Total
72	264
72	312
80	280
78	286
60	360
72	312
99	429
72	312
27	63
632	2618