Official Transcript of Proceedings

NUCLEAR REGULATORY COMMISSION

Title:	Advisory Committee on Reactor Safeguards Human Factors and Reliability & Probabilistic Risk Assessment Subcommittees Meeting
Docket Number:	(not applicable)
Location:	Rockville, Maryland
Date:	Friday, December 16, 2005
Work Order No.:	NRC-775 Pages 1-111

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)

HUMAN FACTORS AND RELIABILITY & PROBABILISTIC RISK ASSESSMENT SUBCOMMITTEE MEETING

+ + + + +

FRIDAY,

DECEMBER 16, 2005

+ + + + +

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

MEMBERS PRESENT:

GEORGE E. APOSTOLAKIS ACRS Member

MARIO V. BONACA ACRS Member

THOMAS S. KRESS ACRS Member

ACRS STAFF PRESENT:

ERIC A. THORNSBURY ACRS Staff

ALSO PRESENT:

ZOUHAIR ELAWAR

BRUCE HALLBERT

ALI MOSLEH

ERASMIA LOIS

A-G-E-N-D-A

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

8:28 a.m.

CHAIRMAN APOSTOLAKIS: This is the second day of the Advisory Committee on Reactor Safeguards, Subcommittees on Human Factors and Reliability and Probabilistic Risk Assessment. I'm George Apostolakis, Chairman of the Reliability and Probabilistic Risk Assessment Subcommittee.

Members in attendance are Mario Bonaca, Chairman of the Human Factors Subcommittee, and Tom Kress.

The purpose of this meeting is to review the status of the Agency's current research and human reliability analysis. The subcommittee will gather information, analyze relevant issues and facts, and formulate proposed positions and actions as appropriate for deliberation by the full committee.

Eric Thornsbury is the Designated Federal Official for this meeting.

The rules for participation in today's meeting have been announced as part of the notice of this meeting previously published in the Federal Register on November 28, 2005.

A transcript of the meeting is being kept and will be made available as stated in the Federal

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

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

We will now proceed with the meeting, and I call upon Mr. Bruce Hallbert to begin the presentations.

DR. LOIS: I would like to start out the meeting.

CHAIRMAN APOSTOLAKIS: Okay, Dr. Lois.

DR. LOIS: I just want to note that Dr. Mosleh is here representing himself. He has a long-time interest in the area of human reliability and the use of Bayesian frameworks for improving the human reliability technology, and he offered to address the committee. However, he's not in any capacity -- contracted capacity for the NRC.

And, with that, I will ask Dr. Hallbert to start the discussion on HERA database development, and then pursue the discussion on Bayesian frameworks.

DR. HALLBERT: Thank you, Dr. Lois.

We are happy to be here this morning to talk about two activities that we're conducting for the Nuclear Regulatory Commission.

One is the development of the Human Event Repository and Analysis Database, HERA database, and I'm going to provide you with some information about the project status, what's going on, and sort of refamiliarize members of the ACRS and the other people who are here today with what HERA is, and what we are hoping to do with HERA to support human reliability analysis.

And secondly, I'm going to provide a short introduction to, and a summary of, a workshop that we had related to the use of Bayesian methods for employing information, such as information in HERA, to support human reliability analysis.

As always, our purpose for being here is not only to tell you about what we are doing, but also to get feedback from you on where you think areas of our research are practical within the field of PRA and what areas you'd like to make recommendations on.

The background for the HERA project is, to some extent, similar to the current state of the art in human reliability. HRA methods use structured processes to identify the kinds of situations that are likely to produce errors, and the ways in which errors occur.

Most of the methods direct analysts to try to account for variously termed elements of the environment, whether you call it context, PIFs, PSFs, whatever. They are a way to say that we know that human performance is causally connected to variables that are internal to the human and external within the environment and in the work setting.

Identifying these things is somewhat challenging, because from person to person, or context to context, or accident scenario to scenario, the influence of these things varies, and even the factors that contribute to performance, they themselves vary.

So, accounting for these things is important, yet is somewhat difficult. As a result, there's a lot of analyst judgment that's applied in the process.

The concern is that differences employed in the judgment process can materially affect the risk methods that are produced by the HRA process.

The objective of the HERA project is to collect data and information about human performance in PRA relevant settings. The approach has been, and we briefed the ACRS on this a couple of years ago in April, and the approach is still the same, to identify information sources that could be used and that are

PRA relevant, develop a formal process for taking information from those relevant settings and putting it into a database, and in parallel, the new thing that we've been doing over the course of the last year or so is developing approaches to use the information that we're collecting in a quantitative framework.

The approach to extracting information is based upon this layered model that I'm showing here. At the bottom here, we talk about very objective information, and this model is especially relevant to the activities that we are conducting where we are extracting human performance information from operating events, LERs and things like that.

We look at the information source and we identify, you know, specifically, what happened, when it happened, where it happened, what the consequence of this occurrence was, and in looking into it, where human performance comes into play, we identify instances where human performance was successful, and instances where it was not successful, in terms of its intended consequences.

CHAIRMAN APOSTOLAKIS: So, you are including successes?

DR. HALLBERT: Yes, absolutely.

CHAIRMAN APOSTOLAKIS: How do you get

those? Operators do things all the time.

DR. HALLBERT: Yeah.

CHAIRMAN APOSTOLAKIS: They don't submit LERs for those things, so unless there's something, you know, going on, and they are successful, is that the space?

DR. HALLBERT: Well, the space, the space of our analysis is in LER, at least for this portion that I'm presenting to you now. There's other sources of information that we are looking at as well, too, where it will be slightly different. But, the source -- but the frame of reference for this is in LER.

Within every LER, there are things that were done right and there are some things that were not done as well, or things that went awry, and we capture all those things.

DR. LOIS: This is Erasmia.

Therefore, we are to capturing the success in terms of number of opportunities versus number of failures. We are not saying that so many times the operators or the maintenance groups are out there and basing this specific involvement on this specific valve. It's not that the concept is to -- when we have an event, we analyze what went wrong, and then how they recovered from that event and did not evolve into a more serious event.

CHAIRMAN APOSTOLAKIS: So, you already have something going on if you operate your site correctly, like what they did at Brown's Ferry with the -- source of water that they were not supposed to use.

DR. HALLBERT: That's right.

CHAIRMAN APOSTOLAKIS: But, not routine operation.

DR. HALLBERT: No.

The setting for this analysis is really in the context of an LER report, so it implies that something didn't go right.

CHAIRMAN APOSTOLAKIS: That's very interesting, because usually we focus on failures, and looking at some successes is probably a good idea.

DR. HALLBERT: Well, we have to, because, you know, -- well, at least we believe we have to, because, you know, what differentiates success and failure sometimes in certain environments is a very small margin, and sometimes, you know, the context is the same for success as well as failure. And, we are trying to collect evidence of both kinds.

CHAIRMAN APOSTOLAKIS: Good. Good.

DR. HALLBERT: Yes.

CHAIRMAN APOSTOLAKIS: Now, why is the

triangle going like that? I mean, what does it mean, the dependencies of the -- there?

DR. HALLBERT: Well, what we are trying to show is that, you know, at the beginning, and this is sort of a time flow as well too, at the beginning of the analysis there's a lot of information to be extracted, and that as we go through sequentially and extract more information about human performance, we eventually get to the point where we think we've extracted -- we've extracted all the information that's available to us about the factors that influence behavior in that context. That's all the pyramid really represents.

CHAIRMAN APOSTOLAKIS: Remind us what error mechanisms and error types are.

DR. HALLBERT: Okay.

Well, we started at the lower level here on the diagram, and we are dealing with objective, almost demographic information, what happened, where it happened, and I'll show you an example of this in a few minutes.

The next thing we look at is, we try to identify what went right and what went wrong, and that starts to get into judgmental processes. Okay, somebody didn't do something right. You know, we have to have some criterion for making that decision.

Beyond that, we try to understand what influenced their performance, so that either they did something well or they didn't do it well.

In the case of an error, you know, we would call those error mechanisms. So, if somebody was out, and I'll give you an illustration in a few minutes, if somebody was out applying, you know, a piece of equipment to another piece of equipment, and set it up incorrectly, and detected that, okay, you've made a mistake, now can you recover from that?

If they stop and they formally analyze it and figure out the best way to approach, then they might do it correctly, but if they don't, if they just reflexively try to take it off the device, it might cause the device to trip. And then you have an error.

CHAIRMAN APOSTOLAKIS: As you know in literature, the psychology literature, human failure, the words mechanism and types means something specific. That's not what you mean. Error types, you know --

DR. HALLBERT: Phenotypes, or whatever --

CHAIRMAN APOSTOLAKIS: You don't mean that.

DR. HALLBERT: No, I don't.

CHAIRMAN APOSTOLAKIS: And, mechanisms you don't mean that either, because mechanisms I think really go into the minds of people.

DR. HALLBERT: Sort of the causes.

CHAIRMAN APOSTOLAKIS: Yeah, it's just the causes.

DR. HALLBERT: What triggered behavior in that particular context.

CHAIRMAN APOSTOLAKIS: And, maybe you can put error/success.

DR. HALLBERT: Yes.

CHAIRMAN APOSTOLAKIS: Since you are looking at successes too.

DR. HALLBERT: Yes. Okay.

I'm going to give you sort of a short description here of the information that's in HERA, and by the way I think, and correct me if I'm wrong here, Erismia, but I believe that you have a copy of the HERA draft NUREG?

DR. LOIS: Yes.

DR. HALLBERT: Okay.

CHAIRMAN APOSTOLAKIS: The NUREG? We have a bunch of slides.

DR. LOIS: You should have received a draft report on HERA.

CHAIRMAN APOSTOLAKIS: Let me see, Bayesian

Methods Workshop.

DR. HALLBERT: The one before that.

CHAIRMAN APOSTOLAKIS: The one before that, yes.

DR. HALLBERT: Yeah, I believe in Appendix B of that report there is a detailed summary and description of the data fields in HERA. So, what I'm going to do here is just describe some of these things.

There's two main sections of a HERA -- of an event that's coded in HERA. The front part is what we refer to as the event main profile, and it includes some of that, you know, factual demographic information that we pull off of the LER, the name of the plant, the LER, the event type, which involves whether it was an initiating event or whether the event resulted in either an active or a latent condition, and I'll describe those in just -- the next slide, I believe, and plant mode, power levels, losses, and system unavailabilities as a consequence of the event.

Initiating events, we are using the same terminology for initiating events as are used in PRA and elsewhere within the Agency, conditions that lead to shutting down, new products processes, and they include the same types of initiating events that are commonly used in PRA.

Conditions, not every event results in an initiating event, or every LER results in an initiating event. Some produce conditions, and these are, basically, undesired outcomes. They don't require shutting down the reactor or removing decay heat, but they result in some sort of an off-normal condition, and there's two kinds that we look for. One is an active condition, and that is where somebody does something and the effects of their performance are immediately observed. They do something, something trips, that's an example, or something becomes unavailable, and it's clear it's unavailable.

A latent condition is something where somebody performs an action, or a series of actions, and a system is rendered inoperable or unavailable, but it is undemanded. And so, the condition may lay dormant for some period of time before its effects are observed

CHAIRMAN APOSTOLAKIS: Some of the latent conditions are very difficult to find, though, right, if there is an error in the procedure?

DR. HALLBERT: Well, an error in a procedure is not a latent condition that we would be

-- we would be looking for a latent condition that would be reported in an LER, like an emergency diesel generator was determined to be inoperable six months after some maintenance was conducted on it, and it was later determined that that maintenance rendered it inoperable. That would be a latent condition. The fact that the emergency diesel generator would not function, but was not detected to be inoperable is a latent condition. It would also be an LER.

CHAIRMAN APOSTOLAKIS: Well, that's --

DR. HALLBERT: Yes, but it would also be reported under the LER rule, most probably.

CHAIRMAN APOSTOLAKIS: Okay.

DR. HALLBERT: Okay.

Every time -- we typically break an event down into constituent pieces, and there's some subjectivity involved in this now. We've gone from very objective description to now our own assessment of what happened in what sequence of time, and how it played out.

And, basically, there's a couple of things we look for. We look for human actions, and the kinds of human actions we look for are successful human actions, as well as human errors, or human failure events, as well as equipment actions. Sometimes equipment fails, and sometimes it actuates in a way that it should.

In addition, you see that we have these designations for positive and negative, and, actually, there's a typo on this slide here, it says positive human action and negative human action, really, this should be, you know, positive context and negative context, and what we are talking about are conditions that are either present or produced by equipment and human actions.

We refer to these things as sub events, so we'll decompose, if you will, an LER into a series of sub events.

MR. BONACA: Could you give me an example of the difference between human error and negative human action?

DR. HALLBERT: Actually, it wouldn't be negative human action, it would be negative context, in fact, this is --

MR. BONACA: Okay, so --

DR. HALLBERT: This is a typo, it should have said negative context or positive context. I'm sorry. I just caught that. Yeah.

And, the wording up here also for many of these things also corresponds to similar wording conventions that are used in a PRA model, except for pos and neg, that doesn't occur in a PRA model. So, you can see, as we are going through and we are decomposing a series of human and equipment actions we are trying to render some information about the causes and consequences that can also be related to a PRA model of activities.

Dependency, we know that dependency occurs, there have been a lot of discussions of dependencies, and, indeed, one of the analyses that we perform of human actions is to understand how series in human behaviors influence one another, and whether, in fact, there are dependency between actions, because we are trying to, again, use this information as an evidentiary basis to inform HRA methods and to inform the treatment of human actions in PRAs. And, our definitions of dependency come straight out of the Good Practices NUREG. So again, we are building on -- and you'll see, if you've noticed already, much of what we are doing here is building on work that's already been performed and thoughts and ideas that have been promoted by others. So, we are striving for commonality in our thinking with HERA, that, you know, meets and intersects with different HRA methods and models.

For each human sub event, where human performance is analyzed, we then go into more detailed analysis of the human action itself. We look at, first of all, you know, any cognitive components that we can identify from the report, or from our knowledge of the operation of that plant in the specific action, and we look at the PSFs using the same -- most of the same PSFs as are in this SPAR-H method, but we are also adding a few more PSFs based upon the recommendations of a peer review group.

We'll also look at the action portion of the behavior and do the same thing. We'll identify what kind of personnel were involved in the sub-event, whether it was maintenance people, operations people, engineering, management, supervisors, whoever, and we'll look at contributory plant factors. Contributory plant factors is, essentially, an amalgamation of factors that represent context, and the definition of contributory plant factors were taken from research that was conducted by the Halden Reactor Project, where they studied and worked on defining context, if you will, for our purposes.

Contributory maintenance factors are those things in, specifically, maintenance operations, that are suspected to contribute to human performance.

As I mentioned, a lot of actions that we analyzed will be successful human actions, because it's just a fact that all these LERs have resulted in successful termination of the events, whatever they were. So, there are successful human actions as part of most every analysis. And similarly, you know, where there were errors we'll try to relate the error to something like slips, lapses, mistakes, something like that.

Every time an analyst goes through and performs one of these analyses, and notes where PSFs or conditions contributed to performance, we asked the team to document their comments, in other words, how did they arrive at that conclusion, because we know that this is based upon judgment and expertise. There's no clear criteria for these things yet, and, in fact, this still is a research project. We are trying to demonstrate the principles of doing this.

So, you know, we want to look back through these things and judge for ourselves, you know, a year later do we still agree with that assessment, would somebody else make the same assessment?

And, here's an example. I said I'd give you an example, here's one. The event is designated and the LER from which it came is up on the top here.

Now, this kind of a time line is produced for every LER that we analyze. Our system automatically produces this, automatically now, and what it does is, the way that we store the data, it extracts it and it puts it up on a time line like this.

Now, time lines differ because sometimes we don't know when some things occurred, especially like latent errors and stuff, we just know that some time in the past something happened.

Similarly, not all events will have this fine of a detailed time line in terms of hours, minutes and seconds on the bottom here, because we simply don't have that information. In this event, we do. What was happening was, at this particular plant there was a positive activity that was going on, a positive context. The operators were preparing to perform a monthly surveillance of the Turbine Protection System. Now, they had made their notifications to the control room, they had gotten all their work orders, it was an approved work package, and it was properly done. There was nothing wrong. So, they were doing things the right way. Things were set up for success.

At some point, just after midnight, and this is 12:16 in the morning, so it's just after midnight, two auxiliary operators, non-licensed, went out to work on the system, and one operator mistakenly placed a lever actuating tool on the TRIP lever, not the test lever. Okay, it's a simple mistake. Okay?

MR. BONACA: He did not actuate anything.

DR. HALLBERT: He just put it on, okay, that's what's designated as XHE01, that is an error. It's a mistake.

Now, something positive happened right here, which is that the same operator realized that the tool was on the wrong lever. He himself said, oh, I've gone and put this on the wrong thing.

However, there was a second error that occurred because he attempted to remove it from the TRIP lever reflexively. He could have stopped and thought about, what's the best way to do this, maybe I should notify someone that this is a potential problem. But, what he thought was, well, maybe I can just fix it real quickly here. So, he attempted to take it off himself, without analyzing how to best handle that.

As a consequence, he rotated the lever enough just so it caused a main turbine trip. Okay?

Now, there's a line between XHE01, XHE02, this line on a HERA time line indicates dependency.

It means if you look -- if you now look in the HERA record you'll see that we noted a dependency between XHE01 and XHE02. If you want to know what that dependency is, you just have to go and look it up.

One of the features that we're going to have in the future, actually, is to have a dependency table that prints out along with the time line.

So, there's an actuation here. There's a main turbine trip, and because the reactor was above 30 percent the turbine trip caused an automatic reactor trip, which should be an EQA01 not an EAQ01, it's just a small typo.

CHAIRMAN APOSTOLAKIS: So, this is not to scale, I mean, I see. The turbine trip occurred one, what, 1/10th of a second?

DR. HALLBERT: Yeah, yeah, yeah, that's right.

CHAIRMAN APOSTOLAKIS: Okay.

DR. HALLBERT: That's right.

CHAIRMAN APOSTOLAKIS: So, EAQ01 is really very close to XHE02, huh?

DR. HALLBERT: Yes. This is cognitive time.

CHAIRMAN APOSTOLAKIS: Very good.

DR. HALLBERT: Okay.

So, what happened then was, two successful human actions then ensued. The main control room entered into the emergency procedures of the plant, and they stabilized the unit in Mode 3, and there's a dependency here between the decision to enter into the procedures and to successfully stabilize the plant subsequently in Mode 3.

CHAIRMAN APOSTOLAKIS: Uh-huh.

DR. HALLBERT: So, that's an example of a HERA time line.

And for everything on this time line here there will be an entry in the HERA database.

CHAIRMAN APOSTOLAKIS: How many events like this do you have, roughly, do you remember?

DR. HALLBERT: We have -- we have about 47 LERs and about 700 records in HERA right now, and that was the next slide, George.

MR. BONACA: So, but you are adding LER, there are many, many LERs.

DR. HALLBERT: Yes, yes, and, in fact, we just received funding through continuing resolution that we'll have to discuss with our project manager what to do next.

CHAIRMAN APOSTOLAKIS: Do you -- project?

DR. LOIS: No, we don't have funding.

CHAIRMAN APOSTOLAKIS: Geez.

MR. BONACA: Now, the LERs are a particular subset of events that happen at the plants.

DR. HALLBERT: Yes.

MR. BONACA: They have to have some licensing significance. I can't remember exactly the specifics.

DR. HALLBERT: They are threshold events.

MR. BONACA: That's right.

So, but, they have to be, you know, an event that affects other tech specs or whatever.

DR. HALLBERT: Yes.

MR. BONACA: And, does that limit very much the information you put in?

DR. HALLBERT: Well, it means we are limited to just that one source presently.

MR. BONACA: I was thinking about, you know, INPO has --

CHAIRMAN APOSTOLAKIS: Yeah, the APEX.

MR. BONACA: -- the APEX system that you wonder at some point if that would be a worthwhile as a minimum sampling just to evaluate what kind of information is not being, you know, provided, by the LER system.

DR. HALLBERT: Uh-huh.

CHAIRMAN APOSTOLAKIS: You are also using augmented inspection -- reports.

DR. HALLBERT: Yes.

CHAIRMAN APOSTOLAKIS: These are really very detailed.

Question, isn't one LER giving you one time line, so how can you have 700 data records, what does that mean? If you have 45 LERs, why don't you have 45 data records?

DR. HALLBERT: Well, what it means is that, 45 LERs, you know, if you divide 700 by 45 you find the number -- the average number of sub events that are coded per LER.

CHAIRMAN APOSTOLAKIS: But, the time line included the sub events.

DR. HALLBERT: Yeah.

CHAIRMAN APOSTOLAKIS: So, by data records you don't mean the time lines.

DR. HALLBERT: No, no, what --

CHAIRMAN APOSTOLAKIS: There are 45 time lines.

DR. HALLBERT: -- there's 45 time lines.

CHAIRMAN APOSTOLAKIS: Okay, okay.

DR. HALLBERT: That's right.

CHAIRMAN APOSTOLAKIS: Yes, Erismia?

DR. LOIS: I just wanted to note that we would like to load many data, but these activities are really resource drive, because of the analysis that go in, and we -- so far we were not able to load as many events as we wanted.

The other issue that in actuality we are still striving with quantifying the data the way it's better, so as we are having meetings with our HRA users we are changing the structure, if you wish, of the database constantly, and we have to go back and recode some of the data.

So, so far we are not still in -- we are not in the production mode of the database yet.

DR. HALLBERT: That's true. This is definitely a research project.

However, you know, my perspective, my perspective on that is, is that, you know, we started thinking about this a couple of years ago before this project was really initiated formally by the NRC, and we've done some thinking about it, and over time I would say that the number of changes is gradually reducing to the point where we are not conducting major changes to our structures and definitions, but we are really, you know, improving and refining the information that's in it, I think.

Would you agree with that, Erismia?

DR. LOIS: Yes.

DR. HALLBERT: Okay, yeah.

CHAIRMAN APOSTOLAKIS: Okay.

DR. HALLBERT: Let me just mention also, I'm not going to talk about this very much, but we also have an arrangement with Halden, where Halden has been going through some of their experiments, and you heard about the experiments yesterday, and they sent a visiting scientist to Idaho for six months for the express purpose of working on the HERA structure for encoding information from their experiments into the HERA database, and so that is happening now.

CHAIRMAN APOSTOLAKIS: So now, this is on top of the 45 LERs.

DR. HALLBERT: That's right.

CHAIRMAN APOSTOLAKIS: But, their events are on the simulators.

DR. HALLBERT: That's right.

So, what we are probably going to do is partition the database and have a portion of the database that's from operating experience and a portion of the database from simulators.

And, the other thing I was going to mention is, we have currently one doctoral student, and we are planning on a second doctoral student, and that's partly why Professor Mosleh is here today, a couple of doctoral students working on new approaches for data generation and utilization. We have one doctoral student at Vanderbilt, and she's working on taking information from the psychological literature, and focus on a particular topic, and trying to extract information from that and make it, you know, into HERA records.

MR. BONACA: My reason for asking the question before is the fact that, let's take the example where the operator is placing the lever on the wrong component, okay, and then this results in a trip now.

DR. HALLBERT: Uh-huh.

MR. BONACA: There are many events that are planned to happen that are successful, in the sense that he actually is able to remove the lever, nothing happens, so the first error is occurring there, a second one does not occur because it recovers from it, so, therefore, there's not a reportable event. In fact, the problem isn't -- it would be reported within the plant. I mean, the plant will know.

DR. HALLBERT: Perhaps, yes, yes.

MR. BONACA: And, in most cases that will be a condition that happened and was encoded, but is successful.

So, my question is mostly, are there a lot of successful actions that happen at the plant which are not really listed or recognized into a repository of this type, and, therefore, they are somewhat biased, you know, ultimately, the views that you get of how these events occur.

DR. HALLBERT: If you think about it that way, then yes. I mean, we are only looking at, you know, some portion of the iceberg that's above sea level.

MR. BONACA: Yes.

DR. HALLBERT: There's a much larger portion that we are not observing.

MR. BONACA: And, in the portion that ends up with the negative events, I mean, they are reportable.

DR. HALLBERT: Yeah, that's exactly right.

MR. BONACA: Okay.

DR. HALLBERT: But again, remember also that PRA is focused on unreliability.

MR. BONACA: True.

DR. HALLBERT: And, we are concerned about unreliability, specifically, of human performance in these complex operational contexts. What we are trying to learn from HERA is, you know, how people end up in situations where they make mistakes that are of significance, enough significance that they have to report them, and as well, how people behave in those situations and recover. Both pieces of information, we think, are highly relevant to informing our assessment of how to apply human reliability methods and specific parameters that ought to be included in human reliability methods.

So, this is sort of an empirical source of information give that some failure events have occurred. It doesn't give us everything, it's not a great estimator of the number of opportunities for success.

MR. BONACA: Okay.

DR. HALLBERT: Okay.

MEMBER KRESS: This setting of this lever that started everything off is a relatively simple thing to do, and if I were going to try to program something in a PRA what I'd want to know is how often that happens compared to how many of these simple events occur during the year of a plant, or averaged over its lifetime.

DR. HALLBERT: Yeah.

MEMBER KRESS: How do you get that denominator?

DR. HALLBERT: We don't.

MEMBER KRESS: You don't get that denominator?

DR. HALLBERT: We don't presently.

MEMBER KRESS: So, you can't really use that particular part of the thing. It's what comes after that that you have to use?

DR. HALLBERT: Not in a frequentist sort of way, you know a frequentist approach would say, you know, error probability is the number of errors divided by number of opportunities. We don't have -- well, we are not currently estimating the number of opportunities, although there have been some approaches that have been used in PRA with equipment to estimate, you know, estimate exposure in some way or the other.

We talked a little bit about that. We don't have anything yet to discuss about that, but we have discussed it.

Rather, what we are proposing to do at this time is to use Bayesian methods, to take evidence and information that we do have, recognizing that it's sparse, it's incomplete, and it may be partially in some situations, and to employ that systematically so that we can, you know, estimate the likelihood of specific human actions, and that's what the rest of the presentation relates to.

So, if you are ready, what I'm going to do is I'm going to really breeze -- I think I've explained already sort of what the intent of the Bayesian activity is. We have data and observations from a variety of sources, experience, research, training, the literature, and at the same time we are trying to make improvements about the kind of predictions that we make of human actions in PRA-relevant context.

And so, a number of us decided, with the encouragement of our program manager, to get together to share some ideas, to discuss some ideas, related to how you would use information from a source like HERA and apply it formally to PRA-type applications. So, we had a workshop.

And, this morning, you know, time permitting, and we are going to try to move through these fairly quickly, we are going to talk about a couple of things. The first one, Professor Mosleh is going to give a very important introduction about, you know, what exactly is it that we are making predictions of in HRA? What's the quantity of interest, and what do we know about the quantity of interest, and how could we use information in something of a Bayesian framework to do that?

I'm going to follow up with two presentations, and I'll try to speed through these a little bit. One is on using information from simulator research, and how that can inform the Bayesian process, and then Professor Mahadevan from Vanderbilt University has been sponsored to do some research and thinking on this project previously, and I'm going to talk about his thoughts on how you could use data from a source like HERA to estimate model parameters from an HRA method.

So, with that, I'm going to turn this over to Professor Mosleh.

DR. MOSLEH: Okay, good morning, Ali Mosleh, University of Maryland, and following what Bruce's introduction to the objectives of the workshop, this effectively represents some of the proposal's ideas that I presented at the workshop and following the workshop in a couple of meetings with other researchers in the discipline. And, the idea was to see if there are things in the Bayesian framework of probabilistic thinking that could be used in the HRA field and the estimation of error probabilities.

So, some of the material that I will present are typical things that we have seen in the past, some new material, and the key objective is to relate the two, effectively.

We started by asking questions that you would ask within the Bayesian framework broadly, that we need to identify the quantity that we are trying to estimate, so the unknown of interest as I call it. You have some information about that from whatever source or whatever type, it's just generally the information that you have, that you would put in what is known in a Bayesian framework as a prior distribution of the unknown, and the additional evidence that you have, so that's E. evidence information. You gather this evidence based on some assumption, some process of observing the evidence. So, the model of the process that generates the evidence is called a likelihood function, and then you put all these things together through the Bayesian inference to obtain the posterior combined state of knowledge about the unknown of interest.

So, base theorem here is what we have there. P is the quantity of interest we are estimating, pi-zero is the prior distribution, the likelihood function and the posterior distribution, and base theorem.

So, given that, we started asking the basic questions. What is the unknown of interest in the HRA? And so, well, from a PRA perspective what you are looking at the operator response objectively, what is happening, and what kind of physical action outcome is on the system. And, normally labeled as success and failure using some reference point that is model dependent.

And, particularly of interest in PRA is the probability of action failure or error failure. So, P is defined as such, and we have seen in the HRAs and PRAs that people put the probability distribution over that p, so pi of p is the probability distribution of p, and the questions are why are we uncertain about the action outcome. So, what is the source of the uncertainty in that? Why a p between zero and one, not zero and one, why is it not the kind of a deterministic? And, what is the uncertainty about p? You know, what are the sources of uncertainty about p, and why are we uncertain? Okay.

Now, in form of reference I put a couple of major frameworks that people have used in estimating error probabilities, I call them blue model and yellow model, and I'm sure you recognize what's blue and what's yellow, that in the blue model you are calculating the error probability, probability of response as the condition of probability of error or response given a particular condition or context, and you multiply it by the probability of context, and you sum over, so that's kind of a partition of the space of possibility leading to a particular response.

The yellow model, which is the vast majority of HRA models, you have a probability of response directly as a function of something that kind of characterizes the context and in form of performance influence the factors of performance safety factors. And, examples of the gap in this function, function of relationship that we have, the tabular form, such as in the CREAM methodology, you have a table relating conditions, values, parameters, to response probability, mathematical functions such as the one that you see in the SLIM, which is an exponential function, and expert judgment is, essentially, the approach that is taken in ATHEANA. So, these are different forms of the function relating PIFs to response.

In an attempt trying to kind of describe the meaning of p, by no means this is something that is a consensus, just an idea it is, but trying to clarify some of the elements that contribute to the p being a number between zero and one, and then zero and one, is we thought that maybe if you look at it the following way, that given a very specific condition, external or internal to the operator, cognitive or physical, response is predictable as either success or failure, make this assumption.

Now, one of the reasons that we have a probability is that in reality we can only specify a class of similar but not identical conditions, that the level of modeling and specificity that we have only reflect a class of conditions.

And, in that class some of those are really deterministically leading to failure, some of the are deterministically leading to success, and depending on where you fall in the spectrum a fraction of those that lead to success is 1 minus p, and the other fraction p is the error, error probability. So, it refers to the fraction of conditions that produce the error versus success in a grouped -- you know, the definition of a context.

So, p, therefore, is the product of grouping a spectrum of conditions into one class. And maybe an additional layer or contributor to the p could be this kind of questioning the validity of the first assumption, that the behavior is deterministically determined given the condition, that maybe that's not correct and there is still a residual randomness in one behavior. So, that contributes to p also.

CHAIRMAN APOSTOLAKIS: I don't know why you say that. I mean, if you say that there is p, that implies there is randomness, there is a probability that it will do something wrong, but if you consider many, many contexts like that, precisely because there's a probability, sometimes it will fail, sometimes it will succeed. I mean, the randomness is there in what you call the outcome. I mean, the random variable is the outcome, success/failure, then you have a probability for that outcome, that probability reflects the context. But, it is still a probability.

DR. MOSLEH: Yes.

CHAIRMAN APOSTOLAKIS: So, every time you have that context you are not guaranteed that there will be success or failure, because it's a probability that will go one way or the other.

So, this residual randomness is not clear to me.

DR. MOSLEH: Well, it, effectively, means, for instance, a specified condition, according to your model and resolution of the model. You may still say we do not fundamentally know whether people behave in a predictable way, so you have --

CHAIRMAN APOSTOLAKIS: Well, that's what p is, that it is not predictable.

DR. MOSLEH: -- no, but I think the kind of unraveling is kind of peeling it off layer by layer, so like we have specified a set of conditions that there -- it kind of basically is the version between the subset of those conditions that would lead to success, some subset that would lead to failure, and really don't know which ones are those within the class of conditions that they have defined.

MEMBER KRESS: Would those conditions be a combination of the performance shaping factors, for example?

DR. MOSLEH: Yes, normally the performance shaping factors are supposed to kind of characterize such conditions, so, yes, yeah.

CHAIRMAN APOSTOLAKIS: Are you saying that there is not a single p even if you have the conditions, in other words that --

DR. MOSLEH: Yes.

CHAIRMAN APOSTOLAKIS: -- there is some --

DR. MOSLEH: Correct.

CHAIRMAN APOSTOLAKIS: -- group variability or something like that?

DR. MOSLEH: I tried to separate that in terms of uncertainty about p.

CHAIRMAN APOSTOLAKIS: Yeah, but there is uncertainty about p --

DR. MOSLEH: Right.

CHAIRMAN APOSTOLAKIS: -- even though p has a unique value.

DR. MOSLEH: Yes.

CHAIRMAN APOSTOLAKIS: Or you can say, p itself does not have a unique value, because there is some variability, okay, and then I'm uncertain about the curve itself.

DR. MOSLEH: Right.

CHAIRMAN APOSTOLAKIS: So, this residual randomness is not clear to me. You are saying that even in a very specific condition the operator response is predictable. Well, it's not.

DR. MOSLEH: No, that's an assumption. You say, let's assume that --

CHAIRMAN APOSTOLAKIS: If you have a p,

it's not predictable, because, you know, it's a probability. You may do right or wrong. So, that --

DR. MOSLEH: That assumption -- that is -- the fourth bullet is questioning that assumption, that things re predictable.

CHAIRMAN APOSTOLAKIS: What I'm saying is that, perhaps, it's not stated well. Given a very specific condition, the operator response is not predictable, it's always probabilistic.

DR. MOSLEH: Yes.

CHAIRMAN APOSTOLAKIS: But, given a specific condition, we are making the assumption that there is a probability of behavior, in a certain way, that we can deal with, that we can do something about on variability. But, the behavior itself is still uncertain.

DR. MOSLEH: Perhaps, this points to more like from a modeling perspective, whether some of those conditions that lead to behavior randomness are fundamentally identifiable.

CHAIRMAN APOSTOLAKIS: I think the first bullet, Ali, you have to reword. It says, "Given a condition the operator response is predictable value," and then you give a p. So, obviously, it's not predictable, because there is a p. I mean, it can be ten to the minus three, but if they do it a thousand times one time they will fail.

DR. MOSLEH: Are you referring to what I say in the fourth bullet?

CHAIRMAN APOSTOLAKIS: The first bullet.

DR. MOSLEH: The first bullet says --

CHAIRMAN APOSTOLAKIS: Predictable is the wrong word in my view.

DR. MOSLEH: Oh, no, no, but the first bullet said, let's assume that this is the case, that given well specific condition the behavior is predictable.

CHAIRMAN APOSTOLAKIS: Oh, so the second bullet modifies it.

DR. MOSLEH: Right.

DR. HALLBERT: Or, I think the first bullet is saying, you know, given some condition, and a demand on the operator, their behavior will be either, you can classify it as either successful or unsuccessful.

CHAIRMAN APOSTOLAKIS: Yeah, that's the fundamental around the variable, but then you have a model for that, which has a p in it.

DR. MOSLEH: And then you say, the reason for that p is that you really can't identify that specific very well-defined condition, therefore, we attribute p to the variability of the condition.

CHAIRMAN APOSTOLAKIS: Variability of the condition, plus we don't have enough evidence to say that --

DR. MOSLEH: So, that's kind of basing -- the second bullet says the variability of the condition, and the last bullet says, even then you may say, well, we really don't know or we cannot really predict in a deterministic way, given precisely defined conditions. Therefore, there is --

CHAIRMAN APOSTOLAKIS: Well, your next slide probably explains a little bit more.

DR. MOSLEH: So, if you are putting another layer of uncertainty on p now, the value, then what could be potential sources of that variability.

CHAIRMAN APOSTOLAKIS: So, am I to interpret this that when we say context there is some randomness within the context.

DR. MOSLEH: Right.

CHAIRMAN APOSTOLAKIS: Yesterday we saw results from Halden that showed us that they had seven crews responding to the same sequence, and there was randomness in the response time, right? So, how would I take those times then?

No, it's not the response time, the response time does not define the context. The response time is the outcome, where there was something there that made one crew do the right thing in five minutes, and another crew do it in 11 minutes. So, it's this underlying cause that could be random, right? We don't know what it is.

DR. HALLBERT: Well, in the case of Halden, in the case of the Halden research, this is Bruce Hallbert again, they were attributing the variability in performance, to some extent, to the variables they were manipulating.

However, as you started digging in deeper, what you found was that they were crew-related factors.

CHAIRMAN APOSTOLAKIS: Additional.

DR. HALLBERT: Right.

CHAIRMAN APOSTOLAKIS: And, I think what Ali is trying to say, actually, he is saying, he is not trying to say, he is saying, is that there is variability that is not of the state of knowledge type.

DR. HALLBERT: Right.

CHAIRMAN APOSTOLAKIS: There are some other factors, like maybe the experience of the crew or whatever, which -- anyway, I don't know where you are going with this.

DR. MOSLEH: Okay, stochastic --

CHAIRMAN APOSTOLAKIS: What's PIF, by the way, PIF is what?

DR. MOSLEH: Performance influency factors of PSFs.

CHAIRMAN APOSTOLAKIS: Yeah.

DR. MOSLEH: So, you see, there is stochastic variability in characterizing the condition variability and time pressure and other conditions, crew-to-crew characteristics variability, are the types of things that one could kind of basically use as a basis for varying p for the same class of conditions.

CHAIRMAN APOSTOLAKIS: So, are you going to the two-stage Bayesian here approach?

DR. MOSLEH: Yes.

CHAIRMAN APOSTOLAKIS: Okay. All right.

DR. MOSLEH: And then, the other thing is, uncertainty of the values of specific PIFs, that could be treated as -- PSFs, that could be treated as parameter uncertainty. And, the model uncertainty, whether you have the right set of PIFs, for instance, the PSFs.

CHAIRMAN APOSTOLAKIS: And, how they influence the behavior.

DR. MOSLEH: How they influence, right.

CHAIRMAN APOSTOLAKIS: So, all these factors, the adjustment factors, the other report from Idaho --

DR. MOSLEH: Yeah, all the structure.

CHAIRMAN APOSTOLAKIS: -- this is major model uncertainty.

DR. MOSLEH: Yes.

So, some way, one way you are supposed to cover that under pi of p, the variabilities and these uncertainties. Okay.

Given that, we kind of made an attempt to try to see what is it that we were estimating, and where were the uncertainties go, I see three areas where we could use Bayesian methods in HRA. One is in development of generic reference HEPs, most vast majority of HRA methods we use a reference number, a set of reference numbers, whether they are judgmental or based on experience, or otherwise, generic reference values from different types and sources of information.

And, so far this is still unavoidable because of the questions that you raised earlier, the success database doesn't exist, we don't have the denominator. So, we still need to use some sort of generic reference values.

Expanding and extending the data classification, bringing the Bayesian kind of philosophy into the data classification. Bruce showed an example of how the events were divided into sub events. All these involve judgment. Analysts need to make judgments, assigning, defining PIF values, connections, relations, all these are judgmental to a large extent, based on the evidence that you have.

So, one area is really to bring the -- allow the analyst to record their level of subjective judgment about different parameters or attributes of an event, and also relax the requirement of quality and quantity of data. So, if you don't have success there that can be distribution of a success data, as opposed to a precise point estimate.

So, these are departures from the classical thinking, classical statistical thinking, bringing a little bit of a Bayesian flavor into the data collection and classification.

A third area is the exploration of the role of causal factors. You have -- you record data in HERA or other databases, and then you identify

PIFs, PSFs, you build some cause of connections, relations between the error forcing or error influencing factors and behavior, in construction of such causal models one could see a role for Bayesian methods, in particular, Bayesian Belief Network for constructing causal models, and if you follow that you've got use of evidence to update or estimate the probabilities that you need in the Bayesian Belief Networks.

These are the three top areas.

So, this probably is quite evident by now that, obviously, the nature of the evidence in HRA involves a lot of subjectivity. You have different estimates, estimation forms, estimates based on data from other situations, expert estimate, numbers that are coming from other HRA methods, the non-homogeneity of information source, different pieces from different sources of information, and, obviously, the incompleteness of data sets, whether the -- database is, in fact, a failure-biased database.

Indirect or partially relevant observations, when you observe behavior based on PIFs as an indirect observation, partially relevant information, when you are using, for instance, pilots landing commercial airline data for space shuttle crew landing space shuttles, it's evidence, but it's partially relevant only.

Now, we go to the formalism of base theorem Bayesian methods, you need to consider the fact that in constructing the likelihood function that relates the evidence to the estimate, the fact that you have multiple types of information, you have the issue of dependence among the different sources of information, the issue of credibility of the data sources, data from experts and models, the homogeneity of the data points, the fact that you have multiple conditions for which you might have estimate, and the question is, how do we put all that together to represent a development estimate for a particular condition, applicability of sources of information to the HEP human error if interest, and evidence uncertainty, and then more, perhaps.

Step by step, again now, we are looking at the techniques introduced, Bayesian techniques introduced in other parts of PRA, and see which ones would be useful in this case.

The question of using multiple sources of multiple types of information is a simple one. In fact, you can construct likelihood functions that speak --

CHAIRMAN APOSTOLAKIS: -- conceptually.

DR. MOSLEH: Pardon?

CHAIRMAN APOSTOLAKIS: Simple conceptually.

DR. MOSLEH: Yes.

CHAIRMAN APOSTOLAKIS: Trying to do that thing there in real life is not simple.

DR. MOSLEH: Yes, that's why you still need experts on Bayesian methods.

CHAIRMAN APOSTOLAKIS: They have to make assumptions on the second role of dependence, then it becomes significantly simpler, but --

DR. MOSLEH: Absolutely right.

CHAIRMAN APOSTOLAKIS: -- boy, defending the assumption is -- how many people do you think copy other sources when they report human error probabilities?

DR. HALLBERT: I'm sorry, can you repeat that?

CHAIRMAN APOSTOLAKIS: I mean, there is a number of methods out there, right?

DR. HALLBERT: Yes.

CHAIRMAN APOSTOLAKIS: Do you think these methods are independent? I believe most of them are caught in Swain and Gutman, or at least calibrating what they are reporting using Swain and Gutman.

DR. MOSLEH: In fact, if you look at some of the sections of the reports that talk about evaluation of their method, evaluation means we looked at PRA numbers that they provide.

DR. HALLBERT: Oh, and specifically, you know, some methods have said, and, you know, we heard from SPAR-H yesterday, one of the things that they did in coming up with the base HEPs in the model was, they looked to see -- they looked for conversions between their numbers and what was prevalent out there.

And, I have heard other HRA people at conferences say that they have, you know, validated curve numbers. So, it is still a benchmark.

CHAIRMAN APOSTOLAKIS: Well, the fact, of course, that they decided -- it depends who is doing it, but if they decide to be influenced by Swain and Gutman, maybe make some changes, is also useful information, that means that they find what is in Swain and Gutman reasonable estimates.

DR. MOSLEH: Partially relevant.

CHAIRMAN APOSTOLAKIS: You know, also with the failure rates, it seems to me that there is tremendous dependence to this day of the failure rates on the Reactor Safety Study.

DR. HALLBERT: Yeah.

CHAIRMAN APOSTOLAKIS: I mean, they set the standard, and then, you know, so the issue of independence is a serious one.

But anyway, yeah, conceptually, this is --

DR. MOSLEH: Conceptually, that's the kind of framework. So, if you have multiple sources of information, depending on the type of information you can construct the likelihood function for these different types. So, that's, you know, the --

CHAIRMAN APOSTOLAKIS: And then you have a student do the integral, the denominator, right?

DR. HALLBERT: Or, do it numerically. You know, in Maha's presentation he did go ahead and, you know, we'll talk about that, but, you know, he had to do the denominator numerically through simulation.

CHAIRMAN APOSTOLAKIS: In very rare cases it's analytical.

DR. MOSLEH: We now have algorithms that do a 20 parameter base integration in a fraction of a second, in a very precise way.

CHAIRMAN APOSTOLAKIS: I've always wanted to do that in a fraction of a second.

Now, Ali, I don't know, do you want to show these equations?

DR. MOSLEH: No.

CHAIRMAN APOSTOLAKIS: Okay.

DR. HALLBERT: Why don't you just do --

DR. MOSLEH: Well, in any case, I think those are kind of ways of mapping the techniques to the specific situation.

CHAIRMAN APOSTOLAKIS: Sure.

DR. MOSLEH: So, let me --

CHAIRMAN APOSTOLAKIS: The mechanics of doing it is not of interest today. I mean, we believe -- but the other stuff, how to do what, that's of interest. In other words, to show an equation.

DR. MOSLEH: Well, one area that I mentioned earlier, in terms of bringing -- capturing the subjectivity of the analysis, the data analysis, is what we are suggesting proposing that what we did in the case of Common Cause failures could apply here, there we took data, this is an example from Common Cause, the event, the LER, where data is classified different ways and the analysts put the weight or level of confidence on each data record, and the two Bayesian methods, you know, one can, basically, roll up all that set of weights and subjectivity to the level of estimation. So, there are formulas and methods for doing that.

Another area, the issue of success data

not being available, it's a major issue, major problem we need to wrestle with, and the fact that, again, we can use estimates or ranges of values has helped, you know, that the formalism allows us to be uncertain about success, the amount of data.

Okay, so I mentioned the fact that we can deal with uncertain evidence and data, there's more on formalism, on how to do that, so these have been tried and tested before.

CHAIRMAN APOSTOLAKIS: Wait, wait, wait, come back. The first one there is exact numbers, right?

DR. MOSLEH: Weighted posterior.

CHAIRMAN APOSTOLAKIS: The others are approximate.

DR. MOSLEH: I can't say that. Actually, we can show that the first two -- the third one is definitely an approximate, it's just an ad hoc method. The first two are both correct under different sets of -- different interpretations of interpretation evidence.

CHAIRMAN APOSTOLAKIS: I would say the first one is the really accurate one. Maybe the others, the second one gives reasonable results, too. But, the first one really, because you say this is the evidence, this is what my state of knowledge will be, and I have three different kinds of evidence, so I'll weigh the results and add them up.

DR. MOSLEH: Yes.

CHAIRMAN APOSTOLAKIS: But also in some other cases, I think what you did with the outcome model, even the last one, is --

DR. MOSLEH: Okay result, yeah, for the value, right, yeah.

CHAIRMAN APOSTOLAKIS: -- and it's easier to do.

DR. MOSLEH: Yes, it's a five-second procedure.

CHAIRMAN APOSTOLAKIS: But, I think the first one is really --

DR. MOSLEH: I published a paper showing the relation between the first two theoretically, so I think -- I believe that the two, both are correct on their own set of conditions.

The third one is certainly an approximation.

Okay, well, the issue of using partially relevant information, you can go from standard base theorem to a modified base theorem. This is a little bit ad hoc, but still something that we have done in the component failure probability -- you know, -- failure probability.

Let me talk a little bit about this part, which I think kind of basically is a background and lead to what Bruce will present. We mentioned the possible role of Bayesian Belief Network in developing and establishing causal connections, causal models, so this slide shows the principles of Bayesian Belief Network, for the record, basically. You are talking about fuzzy, uncertain and less than a clear connection between causes and effects. And, therefore, you make probabilistic and non-deterministic kind of statements about the causal relations. That's, basically, the essence of what you do in Bayesian Belief Networks, relating what we call input nodes to output or target nodes to probability relations.

And, you can capture dependencies in a very, very formal way. In fact, initially, it was invented, you know, to, basically, trace interdependencies of the variables.

We have seen the beginning of, actually, a few attempts in formally using BBNs in relating PIFs to performance, so this schematically shows, you know, that we can put the --

CHAIRMAN APOSTOLAKIS: How are the BBN different from inference diagrams?

DR. MOSLEH: BBNs, I would say, are a sub class of inference diagrams, that when you use formally base theorem Bayesian condition of probability to --

CHAIRMAN APOSTOLAKIS: Okay, but in inference diagrams base theorem is the main tool.

DR. MOSLEH: Maybe I'm referring to maybe a broader class of inference diagrams, where the connections are not probabilistic.

CHAIRMAN APOSTOLAKIS: Yeah, it could be done both ways.

DR. MOSLEH: Yeah, that's what I meant.

CHAIRMAN APOSTOLAKIS: But, here -- well --

DR. MOSLEH: Here it is, essentially -- in all application of BBN to our own HRA, that's the -- model, the causal model is not purely BBN, it's a mix of BBN and some other types of relations. But, that's for maybe --

DR. HALLBERT: Aren't these also acyclic, the BBNs?

DR. MOSLEH: BBNs have some restrictions. They are acyclic, so feedback and cross correlations that of a feedback nature cannot be modeled. There are limitations.

This is an example of what we mean by, basically, a rough high-level BBN of different influencing factors. You have the PIFs and the PIFs, and then you show the interdependency, and then you try to put numerical scale on probability for the conditional connection.

CHAIRMAN APOSTOLAKIS: So, this is an example.

DR. MOSLEH: This is an example, yes, yeah. This is what you are using --

CHAIRMAN APOSTOLAKIS: Emotional arousal is something we're never going to use. That's what it says there, "emotional arousal."

DR. MOSLEH: Is not what, sir?

CHAIRMAN APOSTOLAKIS: Most likely this Agency is not going to use that as a PIF. We are --

DR. MOSLEH: Well, it depends on whether you --

CHAIRMAN APOSTOLAKIS: -- to stay away from the mental state of our subjects.

DR. HALLBERT: You know, on the other hand if what you are concerned about is, you know, fatigue, well, maybe take the word emotional out and just say level of arousal or something.

MR. BONACA: It's interesting, looking at p-related factors, yesterday we were talking about control room crews, and we were trying to identify causal factors of why one was successful and the other wasn't. And, what was striking, from the observation I've done when I used to work at plants, was that, you know, among the individuals I knew personally, they were operators who were capable, and yet I saw them working on a crew where the supervisor was dominant as an individual, and clearly the operator accepted this dominance, even on a technical basis. It seemed as if all the calls the supervisor made he accepted blindly, and I really wanted to -- it's something of that, you know, chemistry within your crew represented here, p-related factors, how would you -- because those are very important factors. I mean, the human relationship or how people relate to each other in a crew, it becomes very dominant.

DR. MOSLEH: In my opinion also, they are extremely important factors, and then we are trying to capture those.

Now --

MR. BONACA: Really, the influence among individuals cannot be captured.

CHAIRMAN APOSTOLAKIS: Yeah, I don't think we are really trying to capture the team effects. We are assuming the team acts in a certain way. I'm not sure to what extent the team effects are important.

DR. HALLBERT: We haven't really thought about using that level of data yet, and, actually, to be honest with you, I hadn't really thought about the availability of that.

First, the first that we've even seen any of that is really coming out of Halden over the course of last year.

CHAIRMAN APOSTOLAKIS: I think the Navy has done some work on submarines.

MR. BONACA: You see, to me when I think about, you know, even big events like Chernobyl, I mean, you know, there was an issue within the control room that left open questions, you know, people that are very knowledgeable.

CHAIRMAN APOSTOLAKIS: Well, Ali has the team-related factors.

MR. BONACA: Yeah, that's why I raise that question.

CHAIRMAN APOSTOLAKIS: This is the exception, rather than the rule with the models.

DR. MOSLEH: We have a lot of detail on that category of team-related factors, but I didn't put those in here because I didn't want to make this kind of presentation on a specific model.

CHAIRMAN APOSTOLAKIS: Maybe Halden could shed some light again on this, as we were saying yesterday.

DR. HALLBERT: We really hope to use more of Halden data to --

CHAIRMAN APOSTOLAKIS: Yeah, and there was some work in Germany I remember. It was published in Reliability Engineering years ago, where they actually looked at team effects, and, in fact, there was a domineering person there. I don't remember what the conclusion was.

MR. BONACA: It's startling to sit back behind a glass and they don't know you are there, and see the chemistry.

CHAIRMAN APOSTOLAKIS: Well, I mean, all you have to do is come to an ACRS meeting.

DR. HALLBERT: You think that goes on there?

CHAIRMAN APOSTOLAKIS: I don't know, we'll ask Halden.

DR. MOSLEH: In most recent experiments, we see clear evidence of team effect from the Halden, and they are here, they can tell you more.

CHAIRMAN APOSTOLAKIS: So, you are really going beyond the Bayesian framework. I mean, you are using subjective theory of probability to propose a framework within which people can make judgments.

DR. MOSLEH: Precisely.

CHAIRMAN APOSTOLAKIS: Bayesian -- okay, fine.

Okay, let's go on.

DR. MOSLEH: Okay.

So now, one of the things about the Bayesian inference is, and then if you use BBNs or Bayesian Belief Networks, it gives you a framework to explore the causal connections, causal relations, from data and information. In fact, you can do a backward and forward inference from informational causes given, informational error given causes, to go to kind of understanding of causes when there is a particular error.

This is a very powerful -- I mean, that blue box will give you tremendous flexibility to extract information from limited data, from say 40, or 50, or 700 events that we heard. And, some of the effort I think that has been labeled as Bayesian inference has been basically focused on use of this type of backward and forward inference.

The other thing is that you can really use the Bayesian framework for BBNs to modify generic data, and for that I -- let me go to this particular vu-graph where I have put, say, imagine that you have a Bayesian Belief Network representing the relation between PIFs and performance, and I've color coded them differently, that means it's the same structure but if you go -- if you look at HERA database for instance, and you look at what influences the behavior, you can take that limited data set, this is all in the space of events, it doesn't really need success or failure data, you can classify them into maybe a countable number of sub classes, each representing -- so these BBNs represent those classes.

And, if you have a number of those, 45 or 100 events, you can see what fraction of events fall in context alpha 1, what fraction in alpha 2, and so on and so forth. So, you divide now the space of observation into specific situations which can be characterized by the PIFs that have been observed in here.

Therefore now, we have, these alpha factors could be correction factors, basically. If you have a reference number, you can use these to modify those reference numbers, reference human error probabilities, and get error probabilities that are more specific to context 1 versus context 2.

And so, that's I think something that is quite doable with the HERA database. HERA will provide the top distribution, the alphas, and the generic HEPs need to be developed some other way, and there are some ideas about how to do that, and then you can modify the results using the alpha factors.

This follows to a large extent the philosophy behind like the beta factor in common cause failure analysis. You know, it's kind of the same basic philosophy.

Bruce mentioned some of the limitations of BBNs. They really cannot really capture dynamic factors and the feedback effects, and the quantification is difficult. If you really want to do this quantitatively it's very difficult, because now you are going to deeper layers of the influencing factors when you are revealing with cognitive factors, putting numbers and probabilities on those would be very difficult.

So, what we did, I did, in a different environment, for different application, in fact, for modeling the role of organizational factors in behavior, was to kind of relax the requirement of quantitative scales for BBNs by dividing the BBN into a qualitative segment and a quantitative segment. So, the deeper roots would be qualitative, you know, high, low, medium, and the more observable layers would be quantitative.

And, we have developed some algorithms to kind of connect these two, and therefore they are a little bit more flexible framework.

I guess these are some ideas that we have been basically looking to in terms of kind of basis for the methodology.

CHAIRMAN APOSTOLAKIS: Thank you.

DR. HALLBERT: Thank you.

So, the things that we are presenting here are suggestions or ideas. They are proposals for where we might go with the use of data from a source such as HERA, using Bayesian methods.

Now, one of the things -- I'm going to take a technical pause here for a moment, the computer is --

CHAIRMAN APOSTOLAKIS: Close it.

DR. HALLBERT: Oh, I have to close it, it's trying to do something.

CHAIRMAN APOSTOLAKIS: There is Hallbert up there, the second one, is that Hallbert, too, from the top, that one? What is that one? No. What is that, reliability data, no? We want something that says extending Bayesian methods.

DR. HALLBERT: It's this one right here, by process of elimination I found it.

CHAIRMAN APOSTOLAKIS: All right, boy, look at that.

Does this show the level of stability in organizational terms?

DR. THORNSBURY: This is sort of like a Rorschach test.

DR. HALLBERT: Okay.

CHAIRMAN APOSTOLAKIS: Oh, there's an orbiter?

DR. HALLBERT: It's predictable as well.

MEMBER KRESS: Heisenberg would have liked this.

DR. HALLBERT: All right.

In -- I'm thinking about tying this presentation -- and this is Bruce Hallbert again for the record, tying this to some of the preceding discussions that we've had.

One of the things that we've heard from all the people who are either developing HRA methods or applying HRA, is that they are looking for information about how to take information about the environment and use it in a human reliability analysis.

The question is, is where does that information come from? And, what we know about performance shaping factors, and I'll use that word throughout my presentation, not because there's a specific technical distinction between what I'm saying here and any methods that are used today, but where is that information coming from, and how do we use it?

Now, I'm focusing here more on the Bayesian context. Ali was talking earlier about one form of representing outcomes as being either success or failure, and that's true. If you were to look at many instances in performance, you could classify them as successful or unsuccessful, success or failure. That's only using part of the information that we have available to us.

What I'm going to propose, and what I propose at the workshop, is in addition to looking at whether performance is successful or not, we look at the factors that contribute to successful as well as unsuccessful performance, and try to systematically relate that through a formal model.

And, what I propose at the workshop is

that, our likelihood models should try to become as parametrically elaborate as they need to be, in order to represent the dynamics of human behavior in these operational contexts.

Now, I know that's challenging and daunting from a mathematical standpoint, but we've done some preliminary work on this sort of stuff, showing that it can be done, although you have to do it numerically. But, what we want to do is take information, relate it to the success and failure model, but include parameters about PSFs systematically so that we can talk about where performance shaping factors drive performance to success and failure.

In a sense, it's common, it's similar to like limit state determination for systems, except we are talking about the human system.

And, the issue is that, you know, most methods today come with causal models, and most of the causal models that they use at human performance are related to an information processing model. Everybody is using information processing models to represent cognition.

CHAIRMAN APOSTOLAKIS: So, this is one of the equations that Ali showed, right, on the bottom there?

DR. HALLBERT: Yeah.

CHAIRMAN APOSTOLAKIS: The blue one, that's the blue equation.

DR. HALLBERT: This is the blue equation. Yeah, and --

CHAIRMAN APOSTOLAKIS: How did you manage to make it illegible?

DR. HALLBERT: I think I had a little help from Microsoft.

DR. MOSLEH: It was an error producing condition.

DR. HALLBERT: It's error producing software.

CHAIRMAN APOSTOLAKIS: Or, is it the operator himself?

DR. HALLBERT: It actually looks okay on my computer.

CHAIRMAN APOSTOLAKIS: Actually, you know, the problem is that you probably tried to enlarge it after you typed it, and you have to do it symmetrically, otherwise the letters go on over the others.

DR. HALLBERT: I think it's a problem between Mac translation to --

CHAIRMAN APOSTOLAKIS: It could be too.

DR. HALLBERT: -- because I did this on a Mac.

CHAIRMAN APOSTOLAKIS: Well, that will show you.

DR. HALLBERT: That's what I get for trying to do something on a Mac.

CHAIRMAN APOSTOLAKIS: Okay.

DR. HALLBERT: Okay.

What this is showing is that, these different models acknowledge the role of a variety of different performance shaping factors, such as individual history, learning, heuristics, biases, and the situation, to produce conditions that they describe as either error forcing, or shaping and influencing, or common performance conditions. If we look beyond those terminologies, what we find is that the quantification process is trying to relate information about the performance environment to a model of human behavior to predict a metric of risk. That's all they are doing. And, these are different ways down here of doing that.

In terms of my proposal at the Bayesian meeting, it was to look at extending the Bayesian framework to address more parametrically elaborate solutions to making predictions about the likelihood of success and failure, using many of the things that Ali was talking about previously.

The goal is to be able to systematically, as I was saying earlier, systematically relate information about the environment to our model of human behavior, and use that model of human behavior that's empirically calibrated to make predictions of outcomes that are related to our metrics of interest for the HRA.

We'd like to, in doing this, in using information from relevant operational contexts to inform our models, we'd like to improve the accuracy of the HEP estimates. We know there are right now sources of data that are used in HRA commonly come from look-up tables, or from sort of a static deterministic or predictive model in some way.

The question is, is could we use data to feed into our model, and at the same time attempt to try to account for the multi variate nature context, and then the SPAR-H presentation yesterday, you heard about the -- or we discussed the APSFs. That's, essentially, a multi variate explanation of context, it's saying that at any given time of writing things, it could be influencing the behavior of a crew.

And, what we are talking about is trying to figure out comparatively-based ways of developing those models of likelihood.

The illustration that I'm going to talk about here is something that I discussed with the ACRS probably three years ago, maybe four years ago now, and it's related to data that I've been collecting over the years on performance shaping factors.

One of the recommendations from the ACRS was, well, why don't you write up this research, and I've done that. I have a draft manuscript about what I'm presenting here, and my plans are to submit that to the Reliability Engineering Journal.

The study focuses on different kinds of PSFs and how they relate to performance, and how to use them in models such as Bayesian models is what I discuss at the workshop or suggest at the workshop.

So, in order to do that, you need to have a set of PSFs that are in some way predictive of crew performance, and there's been a lot of discussion about, how do you take performance related data, or performance data, and relate it to a failure model? And so, we discuss that, and we're in the process of discussing that still with Halden, for example. How do we figure out how important some PSFs are relative to one another. How do we express the relationship between PSFs.

So, in this research we've collected data on performance shaping factors, and the kinds of PSFs that we were looking at at the time were procedures, training, stress, workload, information, system feedback and other elements of the human-machine interface. And, during a series of experiments over a number of different years, we asked operating crews at different reactors, as they went through a set of similar scenarios, to rate the effects of these PSFs on their own performance.

And --

MEMBER KRESS: High, medium, low?

DR. HALLBERT: No, we used a Licard scale, it was a 7 point Licard scale.

And, we were interested in knowing how they believed that these things influenced their ability to carry out some critical tasks related to mitigation of the transient that were clearly a PRA-relevant gate. In other words, if you failed this gate, you'd go down a leg in your event sequence. So, it was a critical path.

We had the operators rate these PSFs on their performance after the scenario, and we collected the data now, you know, in the U.S. and abroad, with licensed operators in simulated operational settings, and the types of scenarios that were used were broadly representative, over heating, over cooling, loss of coolant accidents, and these were specifically pressurized water reactors.

And, George, as I was relating this information to the other members I was saying that, when I was here three or four years ago I presented this study, and one of the recommendations from the committee was, well, why don't you write that up, and why don't you come back and consider how to use this in a reliability way, so that's what I'm doing. I have written it up, and I'm preparing to submit it to the Reliability Journal.

CHAIRMAN APOSTOLAKIS: This is a paper?

DR. HALLBERT: Yes.

CHAIRMAN APOSTOLAKIS: Okay.

DR. HALLBERT: And then, there's an excerpt of the paper in what you have.

CHAIRMAN APOSTOLAKIS: Yeah, I know this. This was, basically, your presentation to the workshop, right?

DR. HALLBERT: Yes, exactly.

So, we used a linear model of multiple

regression.

CHAIRMAN APOSTOLAKIS: Some people made some interesting comments there.

DR. HALLBERT: I'm sorry?

CHAIRMAN APOSTOLAKIS: Some people made some interesting comments in their presentations there. I purposely went back to the ACRS transcripts and found the guy that says this is incorrect.

DR. HALLBERT: I think he was sort of lambasting me, wasn't that me?

CHAIRMAN APOSTOLAKIS: No.

DR. HALLBERT: I think it was. I thought that it was -- well, we can talk about this later, but I thought it was your question to me about if we had 20 trials from Halden could we use that to estimate a reliability metric, and I don't remember the exact context, but --

CHAIRMAN APOSTOLAKIS: I'd have to go back to it.

DR. HALLBERT: Yeah, I think it was our conversation anyway.

But anyway, we looked at this linear model and related to performance, where Y was the critical mitigation time, and so we were looking at time from initiation of the event sequence to mitigation. And, the different Xs in this model were the different PSFs, and the ratings of the PSFs made by the operators, and the Bs are the weights that were determined through multiple linear regression.

And, what we found was that this model was predictive of performance and accounted for the majority of variability in -- majority of crew variability in mitigation time.

We found that the model became more predictive on a scenario-specific basis, so if you aggregate the data the model would be predictive across scenarios, but became more predictive at an individual scenario level.

And, furthermore, differences in the model predictions were observed across plants. So, really, what it's saying is, there is predictive ability in these PSFs. The operators understand their influence. They are able to express their influence. We find an association between their perceptions of the PSFs and their actual performance, and we find differences in the PSFs themselves across scenarios and across plants.

CHAIRMAN APOSTOLAKIS: This point is extremely important, of course, and you were here yesterday when we had the expression of two different viewpoints from Susan and David. Somehow, I mean, you are arguing now that, yes, the PSFs do make a difference, and they probably define a lot of the context.

DR. HALLBERT: Uh-huh.

CHAIRMAN APOSTOLAKIS: What do you think, Ali? The difference was -- you have to go with the concept of context from the beginning, because if you don't do that, you know, you go with the PSFs, you will never really manage to define -- to describe the context.

The counter argument from David Gertman, which I believe you support, is that, look, the PSFs are an approximation, but if you have a good set of eight, or ten, or whatever, maybe you capture 80 percent or even more of the context.

Have you thought about it at all?

DR. MOSLEH: Oh, yeah, a lot.

CHAIRMAN APOSTOLAKIS: Just, give us a short answer.

DR. MOSLEH: I think what we -- I have done in the past ten years about modeling, called development of the causal model, you know, basically, had to address this issue, what do we mean by causal kind of relations, and whether you need the level of detail that, say, you see in real events from kind of a predictive perspective.

And, I think the truth is really somewhere in between. This really highly abstract that some PSFs, with no clear correlation or connection with performance, probably are not sufficient to identify the types of errors people make, the specific errors.

At the same time, I don't think that the extreme view that you really need to know the precise set of conditions in order to make any meaningful prediction is kind of extreme view.

CHAIRMAN APOSTOLAKIS: Well, I asked, if you were to do an HRA tomorrow, or a PRA, okay, somebody comes to you and says we are going to do a PRA for this plant, Professor Mosleh, we want you to help the HRA part, but the high level, how would you proceed? You would say, gee, you know, and this HRA is going to be used before the NRC, those guys are going to review it, you know how picky they are. Would you immediately go to ATHEANA, would you go to SPAR-H, would you go to IDAC, what would you do? They want HRA contribution to the PRA, which has to be done in a year and a half.

DR. MOSLEH: I think the right balance between the two would be the right solution in the short term.

CHAIRMAN APOSTOLAKIS: Between the two?

DR. MOSLEH: Between the two, and I would say extremes, I'm not characterizing SPAR-H or ATHEANA as really extremes, but, you know, they could think of two extreme points of view, where you really look at correlated factors correlation as indication of B, and the other one really looks at a very detailed context analysis.

CHAIRMAN APOSTOLAKIS: But, we should do them in parallel? Would you do one first and then the other?

DR. MOSLEH: The framework would be --

CHAIRMAN APOSTOLAKIS: I'm sorry, Bruce.

DR. MOSLEH: -- yeah, I think the framework would be something that is closer to a set of PIFs through some causal model that is -- I think has the right level -- like the type that I showed earlier, you know, that kind of 10, 15 factors, but causally connected, not just a linear list. I think that's dangerous, meaning everything --

CHAIRMAN APOSTOLAKIS: Taking the sum.

DR. MOSLEH: -- yeah, right, that is absolutely -- I think that is incorrect.

But, some model of causal relation between the PIFs, their interdependencies, and performance, on top of kind of a review and analysis of the situations under which the operators will have to deal with an accident, which is part of, actually, any PRA or credible HRA, people look at the scenarios, look at the conditions, but quantitatively you characterize those and then come up with some mapping between the situation and analyzing the set of PIFs and the causal model, and you go to to the quantitative procedures that the model presents, somewhere in between.

I think the hybrid methodology would be most meaningful at this stage of the state of knowledge that we have.

CHAIRMAN APOSTOLAKIS: Okay, thank you.

Okay, Bruce.

DR. HALLBERT: Okay.

So, this slide sort of illustrates the point I was making on the previous slide, simply showing that the correlations between, you know, predicted mitigation time and observed mitigation time, you know, can be largely accounted for by the PSFs, at least in the scenarios we studied and the crews that we collected data from.

I know that Halden is continuing to collect data on PSFs, I can't remember exactly how they are using it, or what analysis they might perform, but that's something I will be talking more about.

CHAIRMAN APOSTOLAKIS: What variable are you plotting there?

DR. HALLBERT: Observed values predicted.

CHAIRMAN APOSTOLAKIS: Of what?

DR. HALLBERT: Oh, actual mitigation time versus predicted mitigation time.

CHAIRMAN APOSTOLAKIS: Mitigation time.

DR. HALLBERT: Yes, mitigation time.

I wanted to use an objective measure, because, you know, we already had subjectivity in the ratings of the PSFs, and we weren't exactly sure whether, you know, whether we could get people to understand the definitions of those PSFs, and whether they would agree upon the PSFs, and so we wanted to have an objective measure to see if these things could be systematically related to some outcome.

CHAIRMAN APOSTOLAKIS: Now, the time, of course, is the actual random variable, right?

DR. HALLBERT: Yes.

CHAIRMAN APOSTOLAKIS: It's a random variable.

DR. HALLBERT: Yeah.

CHAIRMAN APOSTOLAKIS: But, in our -- well, I guess if you have an allowed or available time that you compared it to, then you get 01, which is the success or failure of the human performance, right?

DR. HALLBERT: And, we didn't actually look at that in these situations.

CHAIRMAN APOSTOLAKIS: No, but I'm saying --

DR. HALLBERT: Yeah.

CHAIRMAN APOSTOLAKIS: -- in the big context that you and Ali have described, you have the observable variables, right?

DR. HALLBERT: Yup.

CHAIRMAN APOSTOLAKIS: Here we have two in this context, one is the actual time that it can take compared to available time, and say that was a 01, which is a fundamental variable of human action, then I start saying I don't know whether it's 01, or p, p itself may have a distribution in it, the whole thing that Ali described.

DR. HALLBERT: Right.

CHAIRMAN APOSTOLAKIS: It's very important for people to understand these things, by the way. I don't mean you two, but it's really important.

DR. HALLBERT: Yeah.

CHAIRMAN APOSTOLAKIS: Because these variables and all that, and then Ali introduced as randomness in p, it's not an easy thing to comprehend.

DR. HALLBERT: No, it's not, and one of the questions you ask yourself is, so, where does the data come from for me to initialize those models, or to define the relationships among PSFs? How do I develop those Bayesian Belief Networks, and what I'm suggesting is, you know, there's a variety of sources of information. Some of that information can come from a source like HERA, where we have information that's retrospectively available.

In addition, a source like simulator studies, where you can actually observe the dynamic interaction and interplay between PSFs, and study the correlations and the causal connections between PSFs on performance, help you to develop a more empirical basis for developing your reliability model.

CHAIRMAN APOSTOLAKIS: Absolutely.

DR. HALLBERT: In that particular context.

CHAIRMAN APOSTOLAKIS: Here's a thought that just occurred to me, though. What I just described, there is a response time of the operators, available time, and compare it til you get to 01.

But then, if you look at the models, the available time, and one of the PSFs in SPAR-H is that, or time pressure, that is used as a performance shaping factor, not as a fundamental random variable.

DR. HALLBERT: Right.

CHAIRMAN APOSTOLAKIS: Why not? Why not as a fundamental variable, and then all the PSFs affect the length of that time.

DR. HALLBERT: Well, I believe you are right. I think it could be a dependent measure.

CHAIRMAN APOSTOLAKIS: Could be actually the outcome.

DR. HALLBERT: Yeah.

CHAIRMAN APOSTOLAKIS: Instead of saying success/failure of the operator, you are looking now at the time the operators take to do something, like in the example you showed earlier.

DR. HALLBERT: Yes.

CHAIRMAN APOSTOLAKIS: The guy realized there was an error within a period of time, and that was before something bad happened, but then he made a mistake again. You know, he put the thing -- the lever at the wrong place.

I wonder whether that would be a more reasonable way to proceed.

DR. MOSLEH: To the extent that it's an

observable, of course, is one of the more favored kind of elements of this model, because you can measure, you can talk about it, you can see. But, I don't think that it is really the most representative parameter or characteristics for all situations.

CHAIRMAN APOSTOLAKIS: This is the fundamental outcome. In other words, all the PSFs, all the BBNs that you presented and so on, ultimately feed into how long the operators will take to do something.

DR. HALLBERT: Or, perhaps, the quality of -- the quality of the behavior that they --

CHAIRMAN APOSTOLAKIS: The whole ting.

DR. HALLBERT: Yeah.

CHAIRMAN APOSTOLAKIS: The quality of the behavior might be poor, so they take a long time, like we saw yesterday, 11 minutes or so.

DR. HALLBERT: Uh-huh.

CHAIRMAN APOSTOLAKIS: But, the problem that I see with that, and it would be nice to have someone trying it, but the problem you see with that it negates all the models that are out there.

DR. LOIS: Well, this is the typical TLC.

CHAIRMAN APOSTOLAKIS: No, it's not a TLC.

DR. LOIS: Why not?

CHAIRMAN APOSTOLAKIS: No, because EPRI said, well, EPRI, everybody would, they said if the available time is five minutes this is the probability that they will not do.

What I'm saying is, forget about that, I'm not talking about TLC, I'm talking about all the models that Bruce, and Ali, and others are developing, should serve as a fundamental focus the time it takes for the operators to do something. So, you are going to have PSFs, you are going to have the whole works, like we saw yesterday from Halden. So, all these PSFs now will lead to some probability distribution at that time, and if that time exceeds the available time then you have a failure.

We're not talking about TLCs at all, nothing, TLCs are out.

MEMBER KRESS: I think you want distribution on both of those, available time --

CHAIRMAN APOSTOLAKIS: The available time distribution comes from thermal hydraulic observation or something.

MEMBER KRESS: Yeah, and the overlap will give you that.

CHAIRMAN APOSTOLAKIS: And, the overlap will give you that, yeah.

DR. MOSLEH: If the time to perform something is really what is impacting the PRA model, or the sequence of events, of course, time is the underlying parameter of everything, obviously.

But, whether that's the real anchor parameter, or factor, to look at --

CHAIRMAN APOSTOLAKIS: Look at what Bruce presented a half an hour ago, how does he present the information, he has a time line.

DR. MOSLEH: Yeah, but he chose to --

CHAIRMAN APOSTOLAKIS: They did it here, they did it there, thank God the core damage was down there. You know, it's a fundamental random variable, that the operators do something, and then all the models will try to figure out what are the factors, what are the contexts, if you go ATHEANA will be the context that affects, that influences, that length, because this is really -- and I mean in practice, too, you see, when you do -- because this ties -- there are several reasons, first of all, I'm not really saying do that, I'm saying here is another way of approaching human error that may or may not be better than this, by this I mean, you know, all the models.

DR. MOSLEH: Yeah.

CHAIRMAN APOSTOLAKIS: It ties the human

performance to the thermal hydraulic and neutronic behavior of the reactor, because the available time comes from thermal hydraulics, right, only sometimes, but usually thermal hydraulics, and then you have a benchmark. I mean, if they do it within ten minutes or so.

The other thing is that in licensing actions, when you -- when a licensee asks to raise the power, the allowed power, by 20 percent, the main impact on human performance is the shortening of the available time, right, the available time. So, the crews do take 15 minutes, but before the available time was 18, now it's 11. Okay?

So, you are really dealing in all applications with the actual times. Now, what the licensees say and what the reviewers agree with, is that, okay, so there is a shortening of the time, but it's not very big, so even though we may not know the probability of doing something wrong, the new probability is not that different, so accept. But, nobody really knows what happens to that probability, and probably the reason is that the focus is not on the time.

I don't know, this is something to think about.

MEMBER BONACA: We will go through some of the sequences they have tested on the simulators, they know how long it takes for that parameter to respond to it. So, they have some indications there, but you are right.

CHAIRMAN APOSTOLAKIS: There are some indications, but also --

MR. BONACA: The time --

CHAIRMAN APOSTOLAKIS: -- for some reason, from the beginning, and maybe that's an interdependence on Swain and Gutman, we have all focused, including me, on the probability of the operators doing the right thing or the wrong thing, but it seems to me the fundamental random variable underneath is really the time --

MR. BONACA: Especially some fundamental operator action, for example, clearly, -- for PWRs is a fundamental decision.

CHAIRMAN APOSTOLAKIS: Yes.

MR. BONACA: Now, if you look at how -- for certain plants, like the C plants with small PRBs and small charging flow, very small actually, you have a very narrow window for success. Either you enter -- within, I believe, it's like two hours or one half hour, or you just don't succeed, there's no way to succeed.

CHAIRMAN APOSTOLAKIS: Yea.

MR. BONACA: Okay.

Now, the pressure on the operators to make the right decision, the right call, is tremendous. I mean, you know, because they know that once you get into bigger figure, the whole containment, and you may just -- and that's a big issue, time is just a factor.

CHAIRMAN APOSTOLAKIS: This is critical, I can't think of a case where it's not. So, you should go back to Ali's presentation, for example, the second slide, the amount of interest now will be the time, and the formulation still applies, but you have a different amount of interest in this third slide where you say, the amount of interest is the operator response failure or success, you said it was 01, now you say time to do something, and then I have some estimate from the hard sciences, thermal hydraulics and so on, to tell me how much time I have.

And, ideally, as Dr. Kress said, if you also have uncertainty on the available time, then you do this convolution thing there.

Anyway, I mean, that's just a thought.

DR. MOSLEH: The operator action could change the time scale of things, basically, by the action, the error, they could change the sequence, they could change the time scale.

CHAIRMAN APOSTOLAKIS: But, that's the reality of it, and also it would be very consistent with HERA.

HERA cries for time, it says, you know, here is a sequence, so how do we use that information, and maybe your Bayesian calculations would be, in fact, easier.

DR. HALLBERT: Yeah, we --

CHAIRMAN APOSTOLAKIS: Because you are having evidence directly on time.

DR. HALLBERT: -- no, time is clearly the stream in which all behavior occurs. I mean, everything unfolds in time, over time.

CHAIRMAN APOSTOLAKIS: Right.

DR. HALLBERT: And, I suppose to some extent all the data that we have reflects outcomes in time. The question, though, of conversion, or considering the use of time as the performance metric, and what that tells us, I think we have to --

CHAIRMAN APOSTOLAKIS: Oh, I'm not saying it's obvious, far from it. I'm not saying, no, here is a good way to do it, you dummies haven't thought about it. No. All I'm saying is, perhaps, we should be focusing on that as the random variable, which would be completely consistent with what you are doing in HERA, and again, I'm not using it as a performance shaping factor, no, this is now the true random variable. And, if it's very short, if the available time is very short, then a performance shaping factor will tell me the stress level is high.

Dr. HALLBERT: Well, let us consider that.

CHAIRMAN APOSTOLAKIS: SO, I still have these things.

MEMBER KRESS: Yes, I think you would find this to be a unanimous recommendation from the ACRS, even though they are not all here.

DR. HALLBERT: Okay.

CHAIRMAN APOSTOLAKIS: It's the domineering effect.

Anyway, that's a thought, Bruce.

DR. HALLBERT: See, we have team factors that play here even.

CHAIRMAN APOSTOLAKIS: Yes, that's what I say.

Can you wrap it up in ten minutes?

DR. HALLBERT: Sure, sure, actually, I've got one more slide.

CHAIRMAN APOSTOLAKIS: Summary, well, gee, I like to see that.

DR. HALLBERT: And then, I'll just -- I don't know, should I wrap it up?

CHAIRMAN APOSTOLAKIS: We have read the other stuff, this is from Mahadevan?

DR. HALLBERT: Yeah, right.

CHAIRMAN APOSTOLAKIS: Yeah, it's in the book.

DR. HALLBERT: Yeah, so we could just --

CHAIRMAN APOSTOLAKIS: This is not HERA.

DR. HALLBERT: We can wrap this up, and then if you have some questions on Mahadevan's stuff, otherwise we can come back to that.

DR. MOSLEH: It's a fundamental variable we need to finish in ten minutes.

CHAIRMAN APOSTOLAKIS: And, I'm trying to influence you.

DR. HALLBERT: I think the likelihood is high because I've been drinking so much water here this morning.

CHAIRMAN APOSTOLAKIS: We can stop for a few minutes if you need.

DR. HALLBERT: Ali can talk, but I have to leave.

Summary, we have -- I mean, we have

demonstrated this link between performance shaping factors and performance. You know, we argue, we discuss, we question, we disagree about whether performance shaping factors are important, how they are important, whether we should or should not use them. I think fundamentally we are missing a huge piece of very informative information if we don't account for these things we call performance shaping factors. And, I don't care whether your method calls them PIFs, or common performance conditions, or context, you know, it's important that we account for these things in some systematic way.

Moreover, it's important that we collect data on these variables. It's important that we collect data and have a source of information that we can use to understand the interactions between PSFs and performance, so that we can have, in fact, predictive models.

It's also important that we start thinking about collecting human performance data in ways that allow it's direct use in reliability type models. The information I've been showing you in this presentation is simply performance, it's not related to any model of reliability. And so, when we do research, when we do collect data, it is important that the performance domain also include occurrences of failure, and that way it's possible that we can start using information about human performance outcomes in ways similar to what we do with structural reliability and system reliability, and that is, to develop formal models of human performance, like limit state conditions. And, I say that with a full understanding of what we do in developing limit state conditions.

We need to have data that relate human performance, elements of the context, and be able to derive limit states, so that we can make some real statements about human reliability, and not just estimates.

But, that has to be based in data, it has to be -- and even if we disagree about what to call specific elements, we can agree on specific outcomes.

The other thing that I would say that I feel strongly about in this research is that we didn't begin with any assumptions or models about PSFs in performance. What we allowed to happen was for those relationships to emerge as they naturally occurred. This was naturalistic research. We observed the operators, we collected the data after the fact, and we used, you know, standard parametric statistical analysis techniques to identify or to understand how the evidence supported model development, and that's what I'm a proponent for.

CHAIRMAN APOSTOLAKIS: Good.

Finished? Any questions?

MR. ELAWAR: I am Zouhair Elawar from the Apollo Nuclear Power Plant, I work with HRAs extensively. I just want some clarification or, perhaps, a recommendation to your project.

I hope you don't intend on giving the industry work for them to do, that Bayesian updating, to keep doing -- you know, improving their HRAs. I hope your bottom line will be maybe generating some generic HEPs by whichever means you reach them, and they will be given to the industry with guidelines as to how to use it, because I believe, and I know for sure that at least I am of that frame of mind, there is a deep skepticism about using Bayesian methods on human performance. People don't believe in it, frankly, to say to you.

So, you need to come with a very convincing reason as to why you think it applies, first of all, and secondly, you will need to make it so simple for the end user to be able to apply it without being a mathematician, let's say.

Thank you.

CHAIRMAN APOSTOLAKIS: Okay.

Any other comments?

Well, thank you very much, gentlemen.

As usual, this was very interesting, always incites interesting comments and debates, appreciate it.

Ali, thank you for coming.

I think maybe we can go now around the table and get some views, although we are not writing a letter on the whole program. We will be briefed in February, and we will write a letter on the evaluation of the models against -- you don't have to leave, by the way, it's up to you -- go if you like.

DR. HALLBERT: No, well, are we done with our --

CHAIRMAN APOSTOLAKIS: We are done with your presentation.

DR. HALLBERT: Okay.

CHAIRMAN APOSTOLAKIS: This is now among the committee members.

What was I saying? Oh, yeah, we will review and write a letter on the evaluation of the models against the best practices, but what we are planning to do, as we discussed with the staff yesterday, is in a few months afterwards -- well, actually, not even a few, one or two months afterwards, we can review SPAR-H or other pieces of work that Dr. Lois feels are ready for review, and write comments, I mean, a letter, because as I said yesterday, the committee communicates its views only through letters, what individual members say, you know, their individual opinions.

But anyway, from what we heard yesterday and today, what are the views?

MEMBER BONACA: Well, first of all, I appreciated very much the evaluation of the HRA methods against the best practices. I thought that it was a very clear document, lot of information, it allowed me to really understand much better the HRA tools available.

And, I must say that also it gave me a sense of the value of the best practices document which was developed a year ago, or whatever, I don't think I appreciated it as much before, until I saw the comparison perform, and that was very helpful. That's my sense, and I think that probably the whole committee will recognize the value of this NUREG.

Second, I think I was quite impressed, I must say, by SPAR-H. I mean, from the various indications, it seems to be an effective tool. I also appreciate much more the importance of having separately SPAR-H and ATHEANA.

Now, ATHEANA I made some statement yesterday about, you know, -- motivated by the fact that I would like to see some application of ATHEANA to better understand how it is being applied.

CHAIRMAN APOSTOLAKIS: We can ask the staff to come and show us some of their applications that they have.

MR. BONACA: Because again, I mean, you know, at least to my sense it has been kept hidden and is not used.

CHAIRMAN APOSTOLAKIS: HE called it a nuclear weapon.

MR. BONACA: Well --

CHAIRMAN APOSTOLAKIS: You are threatening people here, but you never use it.

MR. BONACA: -- you never use it, you know, and you also threaten people with it.

But, I think, you know, I can see now there is an application that is being made, and I understand better the differences between what you get out of ATHEANA versus SPAR-H, okay, given that SPAR-H seems to be already a tool that is being used by, I guess, LLR, right, for evaluation. So, we need to look at it more closely as we go and review SPAR, as we've done in the past, we are going to do more of that.

By the way, I was impressed by all the presentations to the subcommittee. I mean, that was a very valuable subcommittee meeting that we had.

The Halden project, it was very interesting to me, I mean, again, particularly the work -- the people working in crews, and that's opening in my mind the question of, you know, how to model this issue of temperament, personalities, how people relate to each other, and those are issues which are dominant in the team. And, you know, it's a mystery to me right now how you are going to effect that, or to use that, although I believe the work from Halden may shed some incite on that, and we don't really pursue that enough.

But, maybe the crews that they used were not -- didn't have that make-up, but I've seen some U.S. crews where, you know, you can see a dominant individual on a crew, and that seems as if -- and people follow pretty much the leader. And so, you are losing the ability of having the crew individuals separately thinking about it and feeding back information.

And, I don't know about using Bayesian methodology for human reliability analysis. I heard this comment here, but I thought that some of the incites were valuable, I appreciate the presentation that you provided us.

That's pretty much my comments.

CHAIRMAN APOSTOLAKIS: Thank you, Mario.

Tom?

MEMBER KRESS: Well, first off, I think as far as risk assessments are concerned, the quantification of human reliability is very important.

CHAIRMAN APOSTOLAKIS: It's very important.

MEMBER KRESS: We are showing that it tends to dominate a lot of the sequences.

And, I think the view that the whole reliability is greatly determined by the performance shaping factors is the right view.

CHAIRMAN APOSTOLAKIS: Is the right view?

MEMBER KRESS: It's the right view. I think definitely those things are what are going to influence it.

Now, I've known -- this thing for quantifying the influence of the performance shaping factors, I think the Halden project gives some incites on what these might be and how they might be, but I don't see how to go from what they have to a quantification, but what I saw by this Bayesian process in the University of Maryland work, using the influence diagrams for example, can have promise, I think, of determining the interdependence of the various performance shaping factors, and come up with a way to actually quantify their influence, and their own particular conditions, how many of them -- how many of these are available, and to what extent they are available in a given call for operator action for some critical action.

So, I'm really encouraged by what I saw from this Bayesian approach, and I encourage them to keep on with it. I think George's recommendation that the focus should be on the time was a marvelous one. I think that should influence how they think about these things, and how they look at them and all, I think they need to take that very seriously.

I guess my final comment was working with the various models, there is a need to put the ISPER study to bed, and I don't know how we do that, but we need to do that.

So, I guess those are my major thoughts right now, George.

CHAIRMAN APOSTOLAKIS: Very good, thank

you, Tom.

Well, I also was very -- I'm very pleased with what I heard yesterday and today. I think we are making significant progress. The staff is to be congratulated for running a very good program in the human reliability analysis.

I appreciated the presentation from EPRI and the industry. I must say that I wasn't too favorable towards the calculator before I heard you, you know, based on what I knew, but now I think there is a lot of value to it. The fact that you are developing software that attempts to make the process more systematic and so on, the benefits that you have when you have a software package I think this is very good.

I was also pleased to hear you say that not too many people, in fact, possibly none, are using HCR. Then the staff presentations were excellent, all of them actually. You know, we learned from all of them, you know, we made comments in the spirit of being constructive.

I'm still now -- not still, I mean, now I am a little bit disturbed by the exchange yesterday between Dr. Cooper and Dr. Gertman, especially since as Mario said the Agency is, in fact, using SPAR-H in regulatory actions as we speak, and then to hear that, you know, the right way to go is to use context is disturbing. Maybe people in the heat of the debate took extreme positions that they really didn't mean, I don't know, but that has to be resolved.

I did appreciate, as always, the presentation from Halden. They are doing very good stuff. A few years ago, I thought you couldn't do anything about human reliability in terms of experiments and so on, but they are doing good stuff. We are getting good incites. I mean, when you see all the crews responding in a short period time, and then one crew is way out there, you ask why, which is a good step, actually, you know, but you ask questions that you might not have asked without this evidence.

And, today's presentation, too, from Bruce and Ali, I thought were very good. I mean, the framework is the proper one, I think. That doesn't mean that, you know, you are going to tell the industry or the staff, go out and use these formulas that Ali showed and do it, no, I mean, the research project has to produce results that are usable by somebody who is not a mathematician, I agree with you, but that's the intent. I don't think they intended to throw out a huge Bayesian thing and say --

MR. ELAWAR: Just want to make sure.

CHAIRMAN APOSTOLAKIS: -- you guys, okay, it doesn't hurt to do that, reduce the uncertainty, very good.

So --

MEMBER KRESS: It seems to me the Bayesian updates on the probabilities, seemed to me like the repository for this should be NRC.

CHAIRMAN APOSTOLAKIS: Oh, absolutely.

MEMBER KRESS: Yes.

CHAIRMAN APOSTOLAKIS: But, the NRC --

MEMBER KRESS: That's their job to do that.

CHAIRMAN APOSTOLAKIS: -- everything we do is public.

MEMBER KRESS: Right.

CHAIRMAN APOSTOLAKIS: Unlike some organizations.

And, the first presentation by the staff yesterday, and its contractors, on the evaluation of the various models against the Good Practices Document, I thought was excellent.

MEMBER KRESS: Yes.

CHAIRMAN APOSTOLAKIS: The document is excellent. I was so pleased when I read it, especially when I read the criticism of -- the commentary, not criticism, the commentary on models that have been developed by the NRC. I thought that was a brilliant move on your part, to have outsiders review it, and then, of course, it's your document, I mean, you have to edit it and so forth, but Jeff did not object to anything.

So, I thought this was an excellent piece of work.

MR. BONACA: -- because, you know, I mean at some point I questioned in my mind how much the EPRI calculator is influenced by the fact that the presence there is influenced by the fact that the industry has been using this approach for a long time, there are other ways maybe to skin the cat, and I think the document of this NUREG offers you some incite.

CHAIRMAN APOSTOLAKIS: It's the very first step, but a very significant step, what I've always wanted, a benchmark exercise.

MEMBER KRESS: Why would you want to skin a cat?

MR. BONACA: That's the point I wanted to make, by the way, that, you need, we need -- one of the things that we need to see is more a benchmark exercise.

CHAIRMAN APOSTOLAKIS: Yes.

MR. BONACA: I mean, there's always some between different users with the calculator, but also different ways of doing this work, because one is then left with a question of, you know, when you put the right numbers in, like with the calculator, and you add factor, after factor, after factor, how credible all these factors are, what kind of numbers do you get at the end of the process.

CHAIRMAN APOSTOLAKIS: Yeah.

MR. BONACA: And, you know, you may train people to come out with the same numbers, that doesn't give me the comfort that the number is the right number, and it may be simply that we are all thinking that and moving in one direction, but comparing with different kind of code or model, that would be a benchmark that would be significant to me.

CHAIRMAN APOSTOLAKIS: Yeah, and I -- as Tom said, I mean, even though that benchmark exercise from the European Union is now, I don't know, 20 plus years old, we can't ignore it. You can't have a paper like that out as that infamous table, that shows the results being all over the place.

MR. BONACA: Yes.

CHAIRMAN APOSTOLAKIS: You can't do that.

You have to respond, either the benchmark exercise was wrong, or inadequate, or whatever, or we have progressed now and we have better results or something.

Now, doing benchmark exercises is not a trivial matter. I mean, it's expensive, it requires the cooperation of many, many groups, but, I mean, we have to do something, and I think the NUREG that the staff is about to issue on their evaluation is really an excellent first step.

So, unless there are any other comments from the members, public.

DR. LOIS: I have --

CHAIRMAN APOSTOLAKIS: Yes.

DR. LOIS: If you -- it will help us if we know which of the projects that we presented you would like to present to the full ACRS committee.

The methods evaluation has been planned.

CHAIRMAN APOSTOLAKIS: SPAR-H.

DR. LOIS: SPAR-H.

CHAIRMAN APOSTOLAKIS: And then you decide. The others is not obvious to us how ready they are, so you decide that.

DR. LOIS: Okay. It doesn't have to be this year, are you looking forward to --

CHAIRMAN APOSTOLAKIS: Not just this year, I mean, but certainly SPAR-H, some time March or April. You know, are there any other projects that are near to --

DR. LOIS: We intend to --

CHAIRMAN APOSTOLAKIS: You can always come to us without the request for a letter, if you feel that, you know, you are going to take a major step in one of the projects, and you would like to have the subcommittee's input. I mean, we do that a lot.

DR. LOIS: Absolutely.

We are developing the ATHEANA users' guide, which addresses some of the --

CHAIRMAN APOSTOLAKIS: Absolutely.

DR. LOIS: -- committee's concerns.

CHAIRMAN APOSTOLAKIS: Absolutely, yes.

DR. LOIS: So, that may show up some time.

CHAIRMAN APOSTOLAKIS: You know, following up on Mario's comment, maybe we can have a demonstration, a presentation on the actual application, to PPS, or -- or both, and that can be in the context of the guide you are developing. And, when do you think we will be ready for that?

DR. LOIS: Summer, not before summer.

CHAIRMAN APOSTOLAKIS: Okay, that's fine.

MR. BONACA: PPS will be very interesting, because the PPS --

CHAIRMAN APOSTOLAKIS: I mean, the PPS, I --

MR. BONACA: -- the human performance, you know, will change the whole dynamics of that issue, I mean, all the crew members, will be eliminated as contributors, okay, because --

CHAIRMAN APOSTOLAKIS: And, the paper that Bye and others, -- and John, on how to use expert judgment in this context, I mean, oh, absolutely, that can be the third one, that can be the third one, and you judge what.

DR. LOIS: Okay, thank you.

CHAIRMAN APOSTOLAKIS: Okay?

DR. LOIS: Thank you.

CHAIRMAN APOSTOLAKIS: But also, don't hesitate to come when you are about to make a major decision, because we've done that with the Regulatory Guide 1174 and after that a lot of the groups come here and they say, look, this is what we are thinking, what do you guys think, rather than coming at the end and having us disagreeing or whatever.

DR. LOIS: Thank you very much.

CHAIRMAN APOSTOLAKIS: Okay, thank you.

Okay, so this subcommittee meeting is adjourned.

(Whereupon, the above-entitled matter was concluded at 10:41 a.m.)