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

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

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

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

MR. KRESS: Is this a linear time scale?

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

MR. SIU: Each of those are equal time?

MR. HALLBERT: Pretty equal, yes.

MR. SIEBER: Did you measure error rate?

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

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

MR. HALLBERT: No, we --

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

MR. HALLBERT: No, we didn't.

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

MR. SIEBER: Thank you.

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

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

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

MR. HALLBERT: Eight crews all together.

MR. POWERS: Right.

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

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

MR. HALLBERT: This is averaged.

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

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

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

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

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

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

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

MR. POWERS: Right.

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

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

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

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

MR. HALLBERT: That's correct.

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

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

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

MR. HALLBERT: Okay.

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

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

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

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

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

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

MR. KRESS: Fifty-fifty chance.

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

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

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

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

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

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

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

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

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

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

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

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

Next slide.

MR. SIEBER: One second.

MR. HALLBERT: Yes.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

MR. KRESS: What were the other three?

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

The ones that did have effects and were demonstrated through statistical analysis techniques included aspects of procedures, training, stress, workload, information available to the crew, the way that the system provided feedback to the crew on their actions, and the human machine interface in general.

MR. KRESS: Is time required to do an action? Is that a performance shaping factor or is that something else?

MR. HALLBERT: That was actually the dependent measure.

MR. KRESS: It's a dependent measure?

MR. HALLBERT: Yes.

MR. KRESS: How long it took him to actually do the --

MR. HALLBERT: The important thing that they had to do in that particular scenario was what we actually measured. I'll explain how we did this just a bit more here now.

We had a data collection instrument that we developed to measure how the operators experienced these performance shaping factors. In their own terms, how they affected their ability to carry out the critical mitigation tasks in a particular scenario. We asked them to rate these performance shaping factors just after the completion of a transient, a scenario study if you will.

MR. KRESS: The instrument could be a form that they fill out?

MR. HALLBERT: It was a form. That's exactly what it was.

We asked them to consider each of these performance shaping factors that we had discussed and defined prior to their running the scenario. Then we afterwards asked them to rate on a scale how these things had influenced their ability to take the appropriate mitigation action, which was specifically defined.

In the case, for example, the loss of feed water, it was to restore the condensate booster pumps. In the case of the LOCA, it was to isolate the hot lag or something like that in a particular scenario.

After the simulator trials were done, these operators rated the affects of the PSFs on their performance of the critical mitigation tasks. The data that I'm going to present today is essentially the result of collecting data at different times with different crews and different locations.

We had four crews in the US plant and that's documented in this NUREG reference here. We had four crews at Loviisa like I was just describing, and then four crews at Halden. And, we had three common scenarios: undercooling, overcooling, and a loss of coolant scenario. Again, we had the thermalhydraulic references for all these scenarios. We thought they were comparable in nature.

Next slide please. The results are that we used a linear model to assess the effects of the performance shaping factors on operator performance. Whereas I mentioned previously, the prediction of "Y" in this formula here was the critical task mitigation performance time. When, after the initiation of the scenario, were they able to complete their critical mitigation task?

We collected data on these performance shaping factors across these scenarios, crews, and plants, and even countries I suppose. What we found --

MR. POWERS: What does it mean when you use a linear model like that with a constant term? It becomes an adjustable parameter in this model.

MR. HALLBERT: It actually was empirically driving. What we found was that -- and you'll be able to see on the next graph, the next slide -- that typically the prediction of performance would intersect with the "Y" axis, and the effects of these performance shaping factors were over and above, or were around, that intersection point.

So let's say, for example, that the average mitigation time was 18 minutes after the initiation of the scenario. You could have the intersection point being at 14 or 12 minutes. Then the PSFs basically predicted up and around -- or the weighting of these factors predicted up and around that time.

What we found through these studies and

the data collection was that the linear model was sensitive to scenario differences. And I'll show you on another slide how we found that. It was sensitive to plant differences, and it also demonstrated predictive ability.

Next slide. I talked about being sensitive to plant differences. Here is the sum total aggregation of the normalized critical mitigation times. These are the predicted values.

We see, overall, that the multiple correlation in the multiple regression model here was 0.36. What that means is that about 14 percent of the variability in the scatter of the actual mitigation time can be predicted by that model.

MR. WALLIS: Now it's predicted based on data? It isn't a prediction from something else?

MR. HALLBERT: It's a prediction from the best fit of that linear model.

MR. WALLIS: So when you have a limited amount of data and a number of coefficients, you're going to predict something even if it's --

MR. POWERS: What he actually is looking at is what fraction of the variance in the data can be explained with this linear model?

MR. HALLBERT: And the unique contribution of the individual performance shaping factors is measured through the beta weights.

So, this is all crews, all plants, all scenarios. Fourteen percent of the variability was explained through this linear model.

When you looked at just one plant for example, all scenarios, the multiple correlation coefficients were significantly higher. And, you found the same result for all the other plants. So what you see is that the predicted model has greater predictive ability when looking at specific scenarios as opposed to all scenarios. We went from explaining 14 percent of the variability up to about 47 percent of the variability.

We found the same thing in plants. In other words, the closer you got to specific scenarios within a plant, the greater the predictive ability of the model was. So this is suggesting something. It's suggesting that individual differences and how operators experience the scenarios is significant. They are truly different. For example, an implication of this might be that how would we recommend people incorporate performance shaping factors into a particular scenario.

Next slide.

MR. POWERS: What you're really saying is that there's not a uniform PSF for every scenario, that I just can't put a constant in there?

MR. HALLBERT: That's right, and it seems to be different across plants.

MR. ROSEN: It doesn't seem to be that surprising, does it? That operators would react differently to undercooling than they would to overcooling, that they would react differently to loss of power? But within those three scenarios, that operators would feel more challenged by undercooling for instance.

MR. HALLBERT: Or more along the lines of what aspects of their procedures and training and other performance shaping factors contributed to their ability to mitigate that transient, and how then in the future to best incorporate those performance shaping factors into the estimation of human error probabilities.

Again, this is part of establishing a technical basis for how performance shaping factors should be addressed in an HRA.

MR. POWERS: Yes. I mean that's what he is really -- he hasn't got anything definitive here, but he's building an information base that's really calling into question the way we do things now. As George said, we do get smarter with time. It's not always obvious we get smarter. All it says is life is more complicated than we thought.

MR. APOSTOLAKIS: Let's not be unfair. People do consider different performance shaping factors for different scenarios in existing models. And, it's nice to have confirmation of --

MR. POWERS: But see, what he's saying is that if you take a specific performance shaping factor and say it's affect is to double the time, that may be true for one scenario, but it may not be true for another scenario.

MR. SIU: That's right. Some HRA methods, indeed, they do allow you to adjust and others they don't. Now for guidance, it raises immediate questions.

MR. APOSTOLAKIS: I mean you see more clearly that --

MR. POWER: More pertinent is that he's also demonstrating that you can actually get something useful out of these studies, which is really excited.

MR. APOSTOLAKIS: I don't think anybody else has done this, have they?

MR. HALLBERT: No, not anything like this.

I kind of combed the literature. Again, the reason why it's been sort of a passion of mine over the years has been because there is such apposity of information about these things. The other things is that it really is needed I believe.

MR. APOSTOLAKIS: So all we needed was a passionate guy.

(Laughter.)

MR. POWERS: That's what's needed in everything. I mean if you hadn't had runners cruising down the mile, we would not understand anything about the momentum of the equation.

MR. HALLBERT: I'm actually more passionate about other things, but this is very interesting.

The other thing that I wanted to mention is that there would be some intrinsic value to not only understanding about the performance shaping factors' relationship on performance, but for example, how important certain of these performance shaping factors are in certain kinds of scenarios. Now I haven't done that analysis yet. I'm interested in looking at it, but I haven't done it yet.

For example, we talked about: are there any properties that are unique to undercooling scenarios that are demonstrated through these performance shaping factors. I don't know. I don't know yet.

MR. APOSTOLAKIS: So now you're really creating the context within which the HRA modeler would develop the models, the general shape of the models. I think this is great.

MR. HALLBERT: Yes, hopefully. And even eventually to provide some insights and better guidance.

MR. POWERS: To be precise George, the context with which they will evaluate the plethora of models, we'll see if they're useful or not.

(Laughter.)

MR. HALLBERT: And perhaps even from a regulatory perspective, eventually to be able to asses the HRAs that are done and to find out whether all the appropriate PSFs have been taken into account.

MR. POWERS: Yes.

MR. HALLBERT: And why they believe so or not.

MR. POWERS: But let us not forget, if you're seeing -- this is not unusual in this stage of understanding to have a substantial amount of the variance that remain unexplained.

MR. HALLBERT: Yes.

MR. POWERS: It's terrible, but I happen to know in a lot of physical fields that that's where we start, with huge amounts of variance and discovering where that increment of variance is --

MR. HALLBERT: Yes. I mean this to me is very exciting because what you're describing is very applicable to this stage right here. There has not been a lot of data collection yet and it's very informative.

MR. POWERS: From a statistical point of view, the problem with your model and your procedures is that what you're treating as well known variables for themselves have a substantial amount of uncertainty in there, and you've used a liner regression analysis in which you're assuming that those things are all precise. You shouldn't have done that. But unfortunately, the regression algorithms for the right way to do that are pretty hairy to work with.

MR. HALLBERT: Yes, and also in the social sciences, these liner regression models have been shown to be fairly robust to certain violations of assumptions and mathematical properties. So, we start there and at least try to establish that there is a relationship and try to understand better the appropriate models eventually.

MR. POWERS: Try a min-max routine against this and see if it doesn't give you -- first of all, it'll eliminate a certain amount of your variance.

MR. HALLBERT: Yes, min-max or stepwise approaches. Good recommendation.

MR. APOSTOLAKIS: You should keep everything in context. You're not producing a --

MR. POWERS: He's looking for a variance that can be explained and what not. Now some of his variance comes from the fact that his independent variables are just themselves uncertain.

MR. HALLBERT: Thank you. I'll summarize now the presentations of both the embedded study and the overall point of my presentation.

First of all, in the embedded study --

MR. APOSTOLAKIS: Excuse me. Can you tell also at some point what is the most important performance shaping factor or the top three?

MR. HALLBERT: I hate to answer your question this way, but we did some exploratory analysis into the relationships among the performance shaping factors, and we found some stability through factor analytic reduction techniques. Essentially, you could sort of define three overarching performance shaping factors in the set of seven if you will. Stress and workload basically comprised one factor that we'll call demand or maybe even workload. But they loaded negatively overall in these scenarios. So what they did was they kind of worked against the operator.

The other ones were procedures and training, procedures and training loaded together. And, that seemed to be best described as preparedness, how well prepared they were to deal with the specific demands of the transient.

The other three were information available, system feedback, and the HMI, which is probably best described as the HMI. So, features of the control room design, features of the crews' preparedness, and the control room systems designed for the scenarios, as well as the crews own experience of the transient and it's negative effect upon their ability to match with the demands.

MR. APOSTOLAKIS: But is available time and performance --

MR. HALLBERT: I didn't define -

MR. POWERS: He has taken that out of his study because that's what he's measuring in "Y".

MR. HALLBERT: I agree with you.

MR. APOSTOLAKIS: If I have a task that needs to be completed in 20 minutes verses another one that's 42 minutes, should I asses the impact of the time difference on these preparedness performance shaping factors and then do my analysis, or do I have guidance as to how the 20 minutes verses the 42 minutes will effect it? Should I go indirectly through the three that you mentioned or it is from the factor itself?

MR. HALLBERT: I don't know.

MR. APOSTOLAKIS: Again, I don't expect you to have all the answers. But, these are the kinds of questions I think that are important.

MR. HALLBERT: It's a limitation of the approach that --

MR. POWERS: The way that he has done his study, he can't really answer the question.

MR. APOSTOLAKIS: That's fine.

MR. POWERS: He didn't say you were wrong. It just said, I have to look at a --

MR. APOSTOLAKIS: He would never say that even if he thought it.

MR. POWERS: We will say that for him.

(Laughter.)

MR. POWERS: I think we can move on.

MR. APOSTOLAKIS: These things are things that we ultimately have to face in certain regulatory actions.

MR. WALLIS: We have faced already. We have some data.

MR. POWERS: I think we can congratulate you on a pretty well defined study. I can quibble with your data reduction techniques, but I know what you're trying to do. I think it's interesting that you're getting insights out of this thing, which is all you can ask for right now. The actual percentages, that will have to come with time.

MR. HALLBERT: Yes.

MR. POWERS: I think we can -- unless you have some particular points you want to make here.

MR. HALLBERT: The think the final slide was just essentially what I've already covered. To date, there have been some studies conducted and there is some data available right now. And, we're looking through those sources of data to see what is relevant for HRA.

New studies offer great promise because whatever we learn from these other studies could be taken into account for the design of future studies to collect specifically, HRA -- or these kinds of questions from the outset that you're asking us today.

Then I kind of end up where I started, which is that I believe these simulator studies are valuable, and they provide useful data for HRA.

MR. POWERS: I would put a caveat on that. I think simulator studies carefully designed, well conceived, appropriately done, and cautiously used can yield insights that perhaps give us an idea on what we ought to be doing.

MR. ROSEN: Just like thermalhydraulics studies.

MR. HALLBERT: I agree with those points you just made.

MR. POWERS: I mean I think that's the step that this committee has never seen, people coming in and doing simulator studies very carefully, very well designed with particular objectives. They may well have done that, but we just have never seen it.

MR. APOSTOLAKIS: They keep it a secret.

MR. POWERS: Well, there's always a problem when you present to this committee that doesn't pretend to be specialists in this field. But this was nice. You could understand it and what not.

What I would like to do now is quickly ask the members what kinds of topics they want to pursue further in the discussions this afternoon. I think we're done -- am I correct in thinking that we're largely done with the formal presentations and now we want to discuss what the research program is going to be?

I myself very much want to go into this topic that showed up on both Erasmia's slide and Jay's slide called tools and tool development. I'd like to understand what the objectives of tools are, what the vision is, who those tools are for, what they're going to look like. And I invite the other members to make comments on what they want to talk about when we come back from lunch.

MR. ROSEN: I'd like to talk about the issues of organizational performance, safety culture, and indicators.

MR. APOSTOLAKIS: Seconded. Also, in addition to this, I would like to understand a little better the development of the plants to develop an HRA model that will actually give distribution. I mean is there a conceptual design at this point or that kind of thing? I know that it's still early.

MR. POWERS: You get the chance to name your topic, not discuss your topic right now.

Graham, are there any other issues that you'd like to pursue?

MR. WALLIS: Nothing more than yours that really asked the questions so far. I'm an interloper on this committee anyway.

MR. POWERS: You are never an interloper. You are a very welcomed participant.

MR. WALLIS: This is a very tough area to quantify. It's much tougher from the hydraulic. And I don't quite know what tools could be useful and how they would be validated. So, I've asked questions like that already.

MR. POWERS: Dr. Kress?

MR. KRESS: No. I'm interested in it too.

MR. POWERS: Okay. Jay?

MR. PERSENSKY: I'm also interested in tools and safety culture issues.

MR. POWERS: George?

MR. APOSTOLAKIS: Can I add one more?

MR. POWERS: Yes, you are unlimited to the topics.

MR. APOSTOLAKIS: The view of existing models and what plans there are to use them in the development of your own model would be of great interest to me.

MR. WALLIS: What's the state of that?

MR. APOSTOLAKIS: Yes.

MR. WALLIS: You had those four operators in so many situations or something. Now, that's an interesting study. But there must have been a lot of things like those before in some other context.

MR. POWERS: Well, as far as care of design, this is one of the best I've ever seen.

MR. WALLIS: Like human performance in flying airplanes.

MR. POWERS: Now let me interrogate our speakers. What would you guys like to talk about this afternoon?

MR. SIU: Actually before we get to that, I think one point to make is that Bruce and Dave Gertman have a flight and they to leave here by about three o'clock. So, any questions that you have relating to I think the last point -- well, I guess you'll obviously have something to say about existing models, but also if you wanted to talk about experiments that would be good right after lunch to make sure those get done.

MR. POWERS: Okay, the experiments right after lunch.

Are there topics that you need to

communicate? Recognize that our intention is to write a letter that says here are the aspects of the research program that we like and what not. So if there are things that you think we need to understand better, don't be shy about it.

MR. WALLIS: I have a question. We had some very general presentations about the program then we had something very specific from Bruce. There must be other specifics that are going on that would illustrate the generalities for me.

MS. LOIS: So then the intent was to give you an overview of where the program has --

MR. WALLIS: But it seemed to be that we went from one pole to the other.

MS. LOIS: But we hope that this will be the beginning of probably several follow up meetings with the committee to tell them in more detail. On the things that we've done in detail -- I guess those that are still in the planning stage, we're just struggling with that, some things.

MR. WALLIS: In the case of thermalhydraulics, we have some sort of big scheme of needs and then we have the framework, which is codes, and then we have individual projects fitting into the codes. And because of an individual project, we've got some kind of general scheme. What the code of mechanism that --

MR. POWERS: The grand vision is what you're talking about, and that's where I want to go with the tool development and try to understand that a little better in the grand scheme.

I think we are going to get an opportunity to see the applications that showed up frequently. I'm much more concerned right now about the underlying technology we're developing that supports all these applications, the PTS, and things like that that are going on, and the strength of that program. And, we'll discuss that.

In that case, I propose we go ahead and break until 2:00.

(Whereupon, the committee recessed for lunch at 1:00 p.m.)

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

(2:02 p.m.)

MR. POWERS: Let's come back into session.

We concluded the last section by saying here are the things that we want to talk about. It looks to me like the topics, the big scheme of the program, what we mean by tools, organization, safety culture, indicators, development of HRA models, and the view of existing models and the state of the art are the topics.

It does not look like we are going to go into any great detail further on data collection and data manipulation and digest. Though, I will emphasize to you the concluding talk on which we did there was illuminating and gives us new insights on the importance of various elements in the program book, the human factors and HRA.

At this point, I'd like to understand better the program, what's in place, what's in just the planning stages, what we're trying to endorse here exactly. Okay?

MR. SIU: Let me start off by saying that we've asked John Forester and Dave Gertman to join us at the side table. I hope they'll chime in with comments as the discussion moves along. Of course, Bruce Hallbert is sitting up front with us. And Dave and Bruce, again, have to leave at about three o'clock.

A number of questions have come up about the vision of the program. I guess I'd like to get to that a little bit. We tried this morning to give you some sense of how we saw things. Obviously, it wasn't detailed at all and it wasn't intended to be.

Let me start by saying that I think that there are two aspects of vision. One is, if you will, organizational, and one is technical. The organizational vision is pretty much what you were seeing this morning. We have needs presented to us from other parts of the agency. From our understanding of what's going on in other parts of the agency, we try to our best to help address those needs through the activities that we perform, which are analyses, reviews, and developmental activities.

This seems trivial, but actually it's not because this is one of the areas where we got good comments from NRR in their review of our research plan. They talked about the need for much more focus or emphasis on issues like HRA guidance. We had it in our original plan, but we hadn't perhaps put sufficient emphasis on that. So, this is one place where we think we're going to strengthen, to support folks who are faced with particular applications.

MR. POWERS: What's your view? It seems to me that there are two models for the support that you could provide to the non-specialist in this area that has needs.

One is that you can say, "Here is my telephone. Anytime you need an HRA analysis, give me a call and we'll get it done for you." Clearly, that's the mode you operating in now and it may well be the mode you have to operate in.

The other vision is to say I'll live that way for a while, but eventually I want to have tools in these guys hands so when they have an HRA question, they can pull up this tool that will act like an expert system, it'll walk them through the questions, and they'll get their own answers.

MR. SIU: I don't know that we actually fit into either model right now. I think what we would like to do is more towards the second. Where we are right now is actually, in the case of reactor operations, NRR is doing the HRA reviews. We are not doing HRA reviews.

What we haven't done, and NRR pointed this out, is we haven't taken the results of our research over the years and boiled it down to something that -- for example, a review of the use when looking at an application. And by "use", it doesn't mean necessarily redoing the analysis. You might just say, "What are the questions I should be asking?" These are things that I think in the short-term we need to be working towards.

MR. CRONENBERG: This morning the power uprates came up as an issue that the PRAs are coming in saying that there's no effect on human performance or little effect on the power uprates. Yet, they have the study where one of the principle impacts was the reduced operator time for reaction to accident scenarios.

And so, we had the conflict there on one -- it was a study, and then the licensees come in and say there is no effect, and this committee had to struggle with these types of issues in the last year and a half on power upgrades.

Have you had any user needs from NRR to answer questions like that or have you given them any support? They are not risk informed, licensed amendment requests. They are traditional licensed amendment requests, so risk information is kind of supplemental to those requests.

But still, this committee has struggled with conflicting -- and their gut feeling is conflicting with what the licensee is telling them.

MR. SIU: And quite literally, we do not have the user need to provide support there, at least in the HRA realm.

Jay, I don't if you guys have been?

MR. PERSENSKY: Not specifically to that. I mean the work that we were doing on the changes to the operator action was in fact in part related to the power uprates. In that, if it is a risk informed submittal, there is a way of dealing with the risk aspect of it. If it's not, we can still apply risk to it. But the basis there was more to look at the level of review.

As I understand it -- Dick was here earlier, and he's been one of the people that I know involved in that from NRR. Most of what they've been looking at for the power uprates, they've actually looked at simulator trials and requal trials and they found that the actual error rates, not HEPs, but error rates have been very low in that kind of a situation. So, they've been basing their approvals on that.

I just saw Dick walk in if you want to

follow up anything on that.

MR. POWERS: Well, let me follow it up with a question here. In the course of this morning's discussion, we had a variety of questions raised on the adequacy of simulator data to reflect what goes on in the actual plant. How does one take those questions and look at the simulator trials with a jaundiced eye?

MR. PERSENSKY: Some of the things that you indicated were problems. For instance, bringing in different people. Just like any other experimental situation, especially when you're dealing with people, you can do a very large, multi-variant experiment, but the time and resources and ability to do that is very limited.

From the standpoint of the situation that we're talking about here for the uprates, it's their plant, it's their operators operating primarily in their mode of operation rather than separate modes of operation. It's their normal mode. So that's what we asked them to demonstrate. The whole point is being able to demonstrate that they can do it with sufficient cushion I believe.

MR. POWERS: The question I'm not asking is, it's not a question of really power uprates. The question is one of research and what Nathan said. What kind of questions should we be arming these guys to ask when they look at that information, and what are we doing to develop those kind of questions?

Like I said, we came up with some questions on the fidelity of simulators for actual plant operations. I mean they're kind of a qualitative sense so it be difficult to defend that as proof. You just couldn't use that information at all. It was just totally inapplicable based on the discussions we've had, but it's enough of a question that shouldn't the research program be addressing that kind of question?

MR. SIU: Yes. And again, I think that was the intent of the guidance task in various areas. We would start relatively modestly in terms of taking what we've learned to date and then trying to if not make a formal guidance, at least provide some useful information to users. And later on, of course, start getting more formal in terms of guidance for specific things.

Erasmia had mentioned the HRA standards activity, for example, and we intend to play a more active role in that activity.

MR. ROSEN: To refer to that comment that

I made earlier was to me, the way you would handle that properly was it's just another performance shaping factor. It's a crew performance shaping factor. Whatever number you ascribe to the likelihood that the crew performs successfully as it is constituted normally, you modify that number with some shaping factor. But a third of the time, the crew is not going to be in its normal configuration.

MR. SIU: And research again, whichever way they answer laws could provide a basis for deciding when you can take a certain degree of credit or under what conditions you can take a certain degree of credit.

MR. POWERS: I think, I mean we've had licensees, or in this case the applicants, come in and say we go through THERP on this thing and we get 1/100, but when we look at our database we see it's more like 1/1000. Could we go ahead and use 1/100 to cover this? And Professor Wallis says, "How do you know that factor 10 is good enough?" That's the question that's really answered here in this guidance program. The other guys he has downgraded his information by a factor of 10. Yes, that's probably more than enough or it's half of what he should've been or something like that.

I mean that's what you mean by "guidance", how to ask a question and what kind of answer is a reasonable answer. It will never be out to two significant digits because every plant is different and every environment is different and what they can tolerate is different.

MR. FLACA: If I could just follow up with a comment on that. When we look at a number though, it really represents something. What's behind the number, of course, is what's important: the procedures, the framing, and so on, how likely the event is going to occur, and what the operator is going to be prepared for. So, I think it really represents the way one thinks about it. I think that's what George was saying before.

And the question is, as far as our programs are concerned, do we have the infrastructure to be able to think about these questions, and be able to answer other questions that might evolve from the pursuit of these changes that are going on out there? Whether we have the tools and ability to do that I think is very critical. If we don't have them, we're only kidding ourselves. We're just not asking the right questions. We don't know if we've got the right answers.

But in all that context, I think it's more than just a quantification. It's really looking at what that means in the context of what you're giving credit for. If it's 1 in a 100, we expected than there should be a certain level of backup, a technical basis for that 1 in a 100. That comes down to doing some analysis based on what procedures and so on is in place. And, we need the tools to do that.

Now the question I guess is do we need certain tools, do we need to develop new to come and address new issues? One of them is the changes in risk as we see them as plants are making changes. Some of this is maybe due to manual actions verses automatic actions or changing things in that way. And how do we go about doing that, and do we have the tools in place to do that?

Isn't that really the issue on the tools? Again, I'm sorry. I came in a little late and I didn't really hear the beginning of it.

MR. POWERS: Well, the issue in tools is -- you certainly hit upon an important aspect on the issue of tools. My particular interest is one of vision of what the tools we want to look for -- not in the next three years, but say in ten years - when we actually get advanced plants coming in here to be certified, what kinds of HRA and HF tools do we have available and for whom? Are they tools for the specialists in these activities or are they tools for the non-specialists in these areas.

MR. SIU: If I could just add to that point. We were talking about organizational vision and I think that was something that we had shared with human factors. As we indicated earlier, PRA and human factors provide different sets of tools for different problems. Clearly, we have to address needs presented to us by the agency users. From a technical vision standpoint -- and this is where we're going to split a little bit because we have different areas of coverage, different domains. On the HRA side, if you want to talk about a very long-terms vision -- and it may not be all that long-term. I hate to think of 15 years out. Five years is kind of our current planning horizon now. I think it's reasonable to hope that we will have a common high level HRA model.

I think there's reason to believe that we can get there. When you listen to different developers talk about what they're doing, the concepts they're using are very similar. We have differences in terminology. We have some differences in scalp of particular modeling elements, but they all share very similar features.

Furthermore, I believe that there's a sense at least in a good numbers of members of the community that there is a need to drive towards some sort of common goal.

MR. POWERS: When you say "common model", you mean common with the agency or common within our nuclear community?

MR. SIU: Within the HRA community, at least the ones that perform assessments for nuclear power plants and similar facilities.

So, we would like to work towards that. That gets to George's point about knowing what others are doing. We're trying to go beyond that. We're trying to work with these others to develop this common high level model.

It's still a very high level description. You're talking about the notion of, for example, the importance of context and modeling the context explicitly. You still have to get it drilled down to what specific elements of context are you talking about. For example, are you talking about it in a static context, a dynamic context, and so forth?

My belief there is that, as now, in a few years we will still need ranges of methods and tools.

Sometimes very simple tools are good enough for the problem at hand, and sometimes you need a much more sophisticated tool. Our job would not only be to develop those tools, but also of course develop the guidance of when do you use one tool verses another?

Again, if you want to talk in terms of vision, this is I think where we might head. Obviously, there's a notion of validation involved here as well. And what Bruce talked about this morning, point us in the direction that we're going to start using -- we believe we're going to start using existing data and we can start generating new data to support at least some limited validation of these models.

I think, as I indicated in one of my answers to I think George's question, it's unlikely that we'll be able to validate these models in all performance areas. But at least for those areas where we think we can collect data, by all means, we'll try to do that.

Obviously as John Flaca indicated, we have to have a capability to address emerging issues. So the methods and tools that we're working towards now, and we have a laundry list of those, we tried to present those in that two-dimensional matrix. But we've also had a list of the issues and that appeared in that paper that we distributed before the meeting.

So those are issues we recognize that we have to deal with now, and we're trying to deal with. Certainly, things come along the path that we haven't anticipated, and we have to have the capability to address those. So, that's kind of the high level vision.

In terms of quantification in particular -- again, the HRA involves qualitative and quantitative aspects. On the quantitative side, we've been talking internally for a while about the notion of reference values, and perhaps interpolation schemes can think of it conceptually. Once we've identified what are the important factors, you define some sort of phase space, and you can hopefully through experiments or super sophisticated analysis develop some reference points to use as a basis for some sort of scheme to say what should the probability be in another part of the phase space for which you don't have those reference points.

So, that's conceptually a notion that I think we're trying to pursue. You won't see much of that in current discussions on quantification because again, we were trying to make sure that we had a good wrap-up of the expert elicitation process that we're using in ATHENA. But there are some place in, I think, the conference papers where we do take about the notion of reference points.

So, again, that's the direction of where we're heading. And I don't know if you wanted us to through the laundry list of activities that we've got to give you a sense of the breadth of applications and the particular technical challenge areas that we think we need to address.

MR. POWERS: I think your slides this morning provided a pretty good inventory on your current applications, and less of an inventory on where you think you ought to be applying HRA. For instance, we raised the issue of Option 2, if replaced, that maybe there was a rule for HRA to apply.

In some sense, I think that NRR generates user needs based on their thinking about things. I'd be equally interested in the user needs you think they should be sending to you. Do you think there's a richer field there that can be explored now, and is there yet another even richer field once you have these tools that you've been talking about?

MR. PERSENSKY: If I may, from my

perspective at least?

MR. POWERS: Sure.

MR. PERSENSKY: As far as the user needs are concerned, most of the user needs are in fact things that we as a staff talk about together. So, it's not like we're over here. In fact, we need to draw on their experience and the kinds of things that come up in the application of what tools they currently have and where those weaknesses might be.

On the other hand, similar to when I was talking about the study we did on the ROP, we indicated there that here are some things that we think might be helpful. So, it's not that we're not already doing that. It may not be to the extent that you'd like to see it, but in fact we do have that process in place and we talk a lot amongst ourselves as far as how we address that.

As Nathan had indicated, there is somewhat of a difference in what you might consider the vision between HRA and human factors though they are very related. He talk about guidance documents, and that's what we do. But I've been envisioning and I've said in the SECY that what we probably need is some sort of toolbox. With current technology, we can move a lot of this stuff that we now have on paper and pencil onto even something as simple as a palm pilot to take inspection modules with various links to be able to get into the technical basis.

So, it would be something that is useable that addresses all of the various documents that are out there right now. Human factors, as I said, is not just vanishing interface. It has all those same elements, elements that we talk about in terms of PSFs for instance or context.

So, there's that aspect of building something. The vision is trying to put everything into one place so that you don't have to carry around a bunch of paper, but also that there be an infrastructure in place that allows us to continue to develop those that need to be improved upon.

We've taken a lot of heat for 0700 in the past. Yet, it is one of the most used documents, not only by the NRC but in the industry. When it comes to control room design, the EPRI meetings, most of what they're doing in developing their stuff is based on that. Nonetheless, it could be a more useable, more useful kind of document. There are still gaps in it. There are still things that we don't have good guidance from.

Most of the guidance that was developed or put into that document came from things that we stole from the military. This is not that we did a lot of research, in terms original research in a laboratory to develop that guidance. Most of that guidance was taken from other places, but we went through a validation process.

The few things that we were able to do in a laboratory type setting, we've made use of the Halden project and whatever we could to get simulator data and develop the guidance and the criteria that are established in those documents. So the infrastructure is really something that -- whether it's our simulator or Halden's simulator or some other simulator, we need access to that kind of thing for operations.

The thing that we have somewhat ignored by spending a lot of time on simulators is that a lot of the errors, and one of things that we found in some other studies that we did with INEEL, was the issue of latent errors. Those errors were being made by maintenance people, not by the operators.

MR. ROSEN: That's my opening.

(Laughter.)

MR. ROSEN: In the context of tools, what you spent more of your time on, and I think appropriately, is the focus on control room operator performance. But what Davis-Besse tells us, and what a lot of other stuff tells us, is that personnel outside the control room, including top managers, maintenance people, supervisors, and engineers can make mistakes too. Mistakes they make become latent errors, and those are the cases that come out and bite your leg.

So the question here, in the context of tools, what tools do you need to look at the performance of other people who are not control room operators? And this gets to the question of organizational performances or rich literature, which I'm sure you know better than me. There's rich literature on organizational development in psychology and how that factors into the personnel performance of engineers and managers and all kinds of people in the organizational settings, and what sort of tools should we be using.

It think that this is the opening. This is the area that can have the single biggest incremental value to the agency. I know it's controversial. If it wasn't controversial, we probably wouldn't be interested in it.

MR. APOSTOLAKIS: If you do it in the context of how to these things affect human, I don't even think you need to go to the Commission.

MR. ROSEN: Well, that's what I'm doing about. Organizational performance is safety culture. And organizational performance is simply the sum over the integral of all the individual performance.

MR. APOSTOLAKIS: You're doing it because you're trying to understand human performance. There would be no objection. That's the way I understand it. I'm serious.

MR. HALLBERT: Part of the --

MR. APOSTOLAKIS: But if you say, I want to establish a program of safety culture, you might as well not even call. You shouldn't start it by itself. You should start it in the context of something that is immediately useful to the agency.

Yes, Bruce. I'm sorry.

MR. HALLBERT: That's okay. Part of the insights from that work that we performed on the errors in power plants that contributed to these risk significant events did identify that maintenance errors were important contributors to many events.

One of the questions that we entertained when we were back here at a meeting on that particular project was if we could just eliminate maintenance errors, could we make a substantial improvement in reducing the number of risk significant events. In other words, if you needed all the failures that occurred in this event for this event to have occurred, if you just removed maintenance errors, you would thereby reduce the number of total events that had occurred.

Part of the quandary in an approach like that is recognized in that maintenance failures for maintenance contributions to significant events don't occur in a vacuum of maintenance. They occur in a context of the overall plant division of responsibilties and mission activities. They're linked to engineering activities, they're linked to operations activities, and it seems like -- and this is just maybe just my opinion right now that I'm saying -- but it seems like if you want to get reductions in the overall rates of some of these kinds of events, you have to understand those contexts and go into some of the causes of those maintenance errors, just like the kinds of causes that contribute to corrective action program failures.

MR. APOSTOLAKIS: They're not just maintenance errors.

MR. HALLBERT: True.

MR. ROSEN: The reason that David-Besse didn't find the problem was because there was an error repeated several times in putting in the access ports. That was an engineering or a management error. If they had put the access ports in, then maintainers would have gone and said that stuff is coming from something other than the flanges.

MR. APOSTOLAKIS: As I said earlier I started reading this root-cause analysis, which is very good. To make it interesting, I started making notes.

If this deficiency can be identified, what is it telling us? Some of them are telling us that the work processes were not very good. They were not required to do certain analyses after they found something, you know. That's a relative easy fix.

I think where the main difficulty will be when they know of the problem and they don't take action. Because, I don't know how to model them. I think that's going to be more difficult. They say it very clear, "the plant restarted without taking correction action for identified problems." This is the utility speaking now.

But these are the kinds of insights that are beginning to address these questions of organizational questions and so on. I believe that as a community we spend too much time trying to model errors during accidents sequences. It turns out that pre-initiating events are much more threatening.

MR. ROSEN: Well, I don't exactly agree with that. I think we spent an adequate amount of time on operating sequence. But, we spent almost nothing on the other piece. I could not do what we've done. We had to do that. But we spent almost nothing on looking at errors.

MR. APOSTOLAKIS: When people talk about errors of commission, automatically they think of a sequence or something that's happened already.

MR. SIU: Just as a comment here -- actually, this is one nice case where feedback from the human factors work led to a task in HRA. We have a task on latent errors, which doesn't get to your point George about the cause and initiating events, but the notion there was to start exploring again the issue of latent errors.

There were some beliefs -- in fact we talked about this issue in Stockholm back in '95 or something like that -- that we have at least HRA tools to deal with the likelihood of, for example, maintenance errors. There was some feeling that the THERP methodology was just fine for that kind of application. Now that was stated without any strong technical analysis but it seemed to be reasonable to the people attending.

What wasn't covered there was the notion of the dependants between multiple errors. Now you start asking about the underlying causes, whether it's culture, whether it's work processes. We intended to look at work processes as part of this work.

We haven't gone as far as safety culture. But now that we heard from Scott this morning, we'll probably open that up and see if we should approach the Commission on that.

MR. APOSTOLAKIS: Again, I don't think it would be wise to say we want to study safety cultures.

MR. SIU: Right, but as a contributor to --

MR. APOSTOLAKIS: Right. We are doing this, we have started it, and now we have to move into this area. You know, that kind of thing.

MS. LOIS: I just want to mention although it's in a past life, the University of Minnesota had done some work in the early 90s, and the early indications were that learning and management commitment were very good predictive indicators.

MR. APOSTOLAKIS: No, I understand that. But I really think you have to test these things against what they found in Davis-Besse.

As I say, some of it is just "all I have to do is fix the process". Some other things though, the knew of problems and didn't take action --

MR. ROSEN: Well, there's a corollary here, George. Just looking at Davis-Besse is not enough. One needs to take some hypothesis out of the Davis-Besse circumstance and then apply elsewhere. And one of the place was Indian Point.

If you think about Davis-Besse, they didn't put the access ports in and they could've. Now Indian Point didn't replace the steam generators when they could have. And so again, you come to the question that there's some commonality.

MR. APOSTOLAKIS: Absolutely. I just mentioned Davis-Besse because it's a hot issue, and I just happened to get the root-cause analysis a week or so ago and I was going through it.

But even there, you say your talking about the access ports, that they didn't do it. Maybe they didn't do it for a long time. They were deferring it from outage to outage for three, or four, or five times or whatever. Is that the indicator or the fact that they didn't do it at all?

These are the kinds of questions that I think the researchers will have to answer. Some people are saying a good indicator of a safety culture -- not the total of course, but a good indicator -- is the number of items deferred. They were planned to do it and they were not done during the outage. So, there may be ways to approach it and get some indication.

MR. POWERS: Let me see if I can summarize what we've said about tools.

We have not a great deal of schism between HRA and HF here, but some. That in the HRA, you're looking to develop tools of varying levels of sophistication and the guidance for selection among those tools, that you're looking to validate these tools both by using existing data and Dr. Bonaca has suggested that we look to see if we can use the data for development of symptomatic procedures.

I'm less persuaded that we will have access to that data or even that this data is readable to this point. It seems to be a common problem when getting the data collected over a decade ago that it is no longer readable by any machine that we have. In some respects, what we may be discovering is that we've gotten sophisticated enough that the controls in that data were too loose to make it very useful to us. So I'm less enthused about that, but it's worth looking for.

But more importantly, you're looking at can we develop data to develop new data to provide some sense of validation recognizing that validating these tools that they use strictly in an interpolative fashion is a pipe dream and it's never going to happen. You may be able to find some reference points in a space that you have some confidence in, and you're hopefully no extrapolating vast distances.

Now what we learned just before lunch, that phase space you will of has dimensions that perhaps we haven't explored yet. We don't know what they are because we have variants in the data and you can look upon variant data as projections from the space that has a high dimensionality.

In the HF area, we're looking at a somewhat different kind of tool, more user-oriented, more delivered to the frontline kind of tool that's the implementation of a vast amount of technology that's in hand now. Is that my understanding?

MR. PERSENSKY: That's part of it, yes.

MR. POWERS: You go on and say you want an infrastructure that allows you to build upon that, but the tool that you're producing is one that would be used not by a specialist but by a non-specialist.

MR. PERSENSKY: In the end, yes, that tool, as well as the training that would go with it.

MR. POWERS: And the training. You still have a guidance aspect to this?

MR. PERSENSKY: Right.

MR. SIU: If I could just add to what you said, Dana. Again, it's not that we're not going to also develop guidance for non-HRA analysts. Again, someone who's reviewing an application wants to know from an HRA perspective, so we're also trying to address the user.

MR. POWER: Yes well, that point that you made, that I took a lot of notes on that I don't see right now, we are tying to support NRR, who are doing the -- I really put that under your guidance category rather the tools category.

MR. SIU: Okay, parse anyway. But there's one thing that says here's guidance, how to use this set of HRA tools. Here's the guidance which might support or review of somebody else's --

MR. POWERS: That's right, and I made a

distinction between the two.

MR. SIU: Okay.

MR. POWERS: Now a question that was raised in connection with your data, do we need our homegrown simulator? You know, the simulator for a US plant run by US people doing the kinds of experiments now done by a Norwegian simulation of a Finnish plant with Finnish plant and Swedish scientists.

But the question posed to you is: without thinking of cost benefit right now, could the research program make bigger use of that kind of a facility?

MR. APOSTOLAKIS: I'm a bit confused. How is this facility different from the simulators that exist right now in this country?

MR. POWERS: This is a research simulator. They go do these wonderful tests and things like that. They invite crews to come spend a wonderful week in Chattanooga running experiments for them, wired up like Ginny pigs with stress measures and stuff like that. I mean to develop data, to develop an understanding, to develop a science.

MR. PERSENSKY: The issue is it's a reconfigurable simulator that you can change things around, which you can't very well do at existing plants or even at our own simulators in Chattanooga.

In addition, there would be much wider data collection opportunity, the kinds of things that they do have at Halden and other facilities, NASA facility, FAA facilities. They'd collect a tremendous amount of data. We never even talked about the data they'd collect. That might get into much more finite kinds of things.

But to answer your question in the best of all possible worlds, having a simulator like that I think would be helpful to human factors. It would be helpful to HRA and it would be helpful to Digital I&C at least. I don't know really that it's that practical.

MR. POWERS: The answer is unequivocally "yes" to the question that's posed. But the follow up question is: do you have a strategy that would make use of this, and would it make use of it 60 percent of the time, 70 percent of the time, 100 percent of time?

MR. APOSTOLAKIS: You are asking uncharacteristically an unrealistic question there. I can't believe my ears.

MR. ROSEN: It's not his question. It's mine.

MR. APOSTOLAKIS: Divorce always from cost. Maybe it's cheaper to fly US troops -

(Laughter.)

MR. APOSTOLAKIS: I mean you're asking would it be nice to have this extra research capability. I'm not going to say "no". It would be nice.

MR. HALLBERT: I guess, you know, from another research perspective also, it depends upon the kinds of questions you want answers to.

For example, you talked earlier about the data available from EOP studies for relicensing and requalification exams. If part of what you want to do is collect a larger baseline on operator performance in different contexts, there probably is a large amount of suitable data there.

If what you want to do is something more unique that requires modification of the operating environment, then you have to start looking at the extent of modifications and finding out can it be accommodated in the existing facility.

If, for example, what we were talking about doing -- and I'll use an example here - evaluating how well a new electronic procedures system would work. Well, you wouldn't actually have to have your own dedicated plant to do that because a number of plants considering doing that right now. You might try to find a plant that was interested in that and say I've got a couple of candidate systems we want to research, can we use your training facility.

MR. ROSEN: The answer would probably be "no" because it's used 24 hours a day, 7 days a week. You kept coming back to your own point that those simulators are heavily used.

And licensees, it's crucial that they get the training done that they have scheduled. They can't afford to have somebody in there messing around with their simulator because at seven o'clock in the morning, their crews are coming in.

MR. HALLBERT: So you'd like to piggyback on efforts that are already going to try to take advantage of data that they're already generating. But unfortunately, the problem that we've always had in the past was something like this, that it is not a regulatory issue.

Very few plants want -- well, I'm not sure how many or which plants like to volunteer for that because if something happens during the simulator exercises that they don't like, then it immediately raises issues for them.

MR. POWERS: And you're never going to find a plant that has an appropriate simulator for looking at a modular plant.

MR. ROSEN: But the point is if we don't ask these questions, if we don't ask them now, they will not be asked. Here we are at the verge of perhaps a new generation of reactors, we all hope -- are we just going to do it the same way we did the last generation, or are we going to do it a little differently?

MR. POWERS: Well, I'm kind of impressed with the last generation lately.

MR. ROSEN: I think we ought to do more. It look 50 years to get to the point where the old generation -- it's pressingly talked about.

MR. POWERS: And now you want to put in another new generation to get me depressed again. You're playing with my sanity here.

MR. APOSTOLAKIS: This is not the only way.

(Laughter.)

MR. ROSEN: To start off, this generation of machines, if we're going to build advanced reactors, highly integrated control rooms, passive safety, it seems to me that an investment upfront of what it takes to build a reconfigurable machine where we can test some ideas and test these things is not entirely out of the question. It shouldn't be.

MR. POWERS: If we're going to ipso facto attack the issue of errors of commission, I don't know how you do it if you don't go get some exploratory data. I mean everybody just throws up their hands for error of commission, and I think exploratory studies may be the only way to broach that subject.

MR. PERSENSKY: If I may, one of the efforts that I put into the advanced reactor plan, the first effort in that included sort of a scooping study of what might be the problems with advanced reactors that we should be addressing, where the gaps between what we know, what guidance we have available, and where we might be going if there's a need to chance. For instance, for advanced light water reactors, we may not need to make many changes to the current guidance. For modular reactors, we might.

But in that, we included an element of looking at the need for a simulator. One of the things that we talked about in that particular element of the plan was that currently we've got "X" plants or units out there. Each plant has its own plant specific simulator, but they're all different; whereas for the future plants, we're looking at more commonalities.

So, there might be a real possibility of joining with the industry or with DOE in developing a simulator that we can all use. Not unlike the kinds of things that they did with some of the test facilities with some of the vendors, where we were jointly funding and working towards that.

So we are interested in that, and we plan to look at that as a matter of fact.

MR. BONACA: But I think you want to have a simulator of a plant with a matching set of procedures for that plant. If you build a new simulator that maybe wonderful as a concept -- but you don't have the procedures which are tied to the machine.

One suggestion. A number of plants have been retired, but they had plant specific simulators. They're probably still effective and can be used.

MR. PERSENSKY: They've all been bought up or trashed.

MR. BONACA: Okay.

MR. PERSENSKY: Because we purchased a couple of them for the TTC as a matter of fact. Some of the others had been purchased by other vendors.

MR. GRIPMAN: I'm Dave Gripman. I wanted to comment on -- Jay stole my thunder there, but I think this idea of looking for synergy with the Department of Energy is a petition to cite a pebble-bed reactor. They have a lot of operations experience and operators available.

I think that might be a way to do some cost sharing because I think the use of this research simulator is a very powerful one. I think having one in the US in addition to whatever else we can learn around the world is a good concept. We can full scope. We can look at test simulators and extract general principles and behavioral profiles as well for crew performance. So, I think that's one way we want to go.

I think the other challenge has to do with the issue that was raised a little earlier on maintenance. When we talk about a simulator, I think if we're talking about simulation, we almost have to go to analytic type simulation if we want to talk about maintenance performance, looking at work processes, and what happens when you disrupt time. Can you force common cause failures across systems and look at what those failure rates might be like to see if those shaping factors were the same?

That's a more challenging type of simulation I think, and that's something that maybe ought to be pursued as well.

MR. POWERS: Peter?

MR. FORD: The answer to my question has been partly answered at least to advanced reactors. If we believe the schedules we're seeing, within the next two or three years, we'll be looking at applications for designing new reactors. We don't have simulator for these new concepts. Therefore, you have to rely on the synergy between the conventional reactors and the new reactors that are coming down the line.

When you look at your needs over the next two years, what's keeping you awake at night? You have no way of knowing how you're going to tackle a particular problem in both the human factors and HRA. What keeps you awake, the sufficient lack of knowledge?

MR. SIU: You know what keeps me awake at night? Nine-eleven.

MR. POWERS: I'm jumping to speak here, which is silly on my part, but I do rather silly things. But when I see massively automated plans, I put on an HRA or a human factor hat, and it's the errors of commission.

I probably should probably worry about

latent errors in the maintenance process. The committee definitely heard the story that they were four times as important as the errors following an initiating event. We got that message last year and we quote it frequently.

MR. APOSTOLAKIS: I'm a little surprised though that some committee members seem to be more enthusiastic about getting the simulator. Rosen and Powers are saying this is great.

MR. FORD: Hold on, George, before you get into that particular topic. Nothing keeps you awake at night?

MR. APOSTOLAKIS: I'm not going to say it now.

(Laughter.)

MR. HALLBERT: I not sure it keeps me awake at night, but it's in my thoughts in the daytime when we think about HRA and we're going this work. I have children so they keep me awake at night.

(Laughter.)

MR. POWERS: Wait until they become teenagers.

MR. HALLBERT: We have that too. They wake us up at night when they come in.

(Laughter.)

MR. HALLBERT: Just a couple of things. One is just trying to reconcile the notion of reliability and validity in the approaches that we currently use. I'll give you some examples.

Reliability is different analysts being able to replicate the results, looking at the same scenarios with the same information. There have been some benchmark studies in which the orders of magnitude difference in results is really bothersome. You know, where they did try to benchmark.

MR. POWERS: There is a really nice paper which I had read, but I cannot refine, in which they compared some of these analytic techniques to each other, and it -- human reliability analysis, and it virtually --

MR. APOSTOLAKIS: It was all over the place.

MR. POWERS: Yes. I mean there was no correlation whatsoever.

MR. HALLBERT: The other thing is just the validity for -- I'm not sure if I'm characterizing this correctly, but at least to me, an apparent lack of a process in which methods become validated. In other words, a group of people produce a method and it's then just released.

I'll say this for ATHENA, ATHENA at least has gone through a lot of very systematic attempts and efforts to try to achieve some kind of validation of the principles of the method. Just given all the methods that are out there, there are some methods that have done that to a much less extent, so you really wonder about different analysts using it. You wonder about the validity of the results that come about as a result.

I then think about the NUREG on lessons learned from the IPEs. And in the appendix, I think there's a very -- I think in fact you wrote it Dana if I'm not mistaken or at least you talked about it at the EHPG in Norway I think when you came over there. There are certain criteria to a PRA completeness. And with regard to HRA, there should be the same criteria. So, I don't think we're there with HRA yet.

MR. APOSTOLAKIS: The thing that really bothers me, and it comes to my comment earlier, is that, as I said earlier, I read one model and they seem to be focusing on decision analysis. Another model is focusing on time. Another model, it says PSS. Another one is expert opinion. And, they operate in parallel with apparent interaction. I think it's time to stop that.

MR. FORD: So follow up then, on both the HF and the HRA, you've got data collection analysis. And you're saying that keeps you awake as you go forward on the current fleet using it in its entirety going to advanced fleet. In the prioritization of tasks for the next five years, is that item high on the prioritization list, data collection?

MR. SIU: Practically number one.

MR. FORD: I haven't seen it yet, so --

MR. POWERS: Nathan says it's number one on their list.

MR. FORD: Great.

MR. SIU: That and guidance are the two tasks that we are really focusing on.

MR. POWERS: To follow up on George's point, my understanding of your program is that you know have, you have number one, guidance. Number two is this data collection. Somewhere down a little lower is to look at all these models, distill which are the good aspects, which are the bad aspects, and come up with some judgment on what a desirable tool would be. Now that may be one that already exists, or may be one that you have to invent, or it may be that you can change a Greek thing into a Latin thing.

(Laughter.)

MR. SIU: He says that, but it's a step backwards.

MR. POWER: Okay, a Greek thing into an Anglo-Saxon thing, which is clearly a step forward, and have a new model.

Is my understanding correct there?

MR. SIU: Again, I think we're talking about, as you indicated earlier, is a range of methods and tools suitable for different applications and guidance to support the appropriate application of those methods and tools.

George, I don't think you were in when we were having a little bit of discussion about driving towards some sort of common model. That's something I think that we would really like to do.

MR. APOSTOLAKIS: Good.

MR. SIU: Some of the discussions we're going to have next week are along those lines.

MR. APOSTOLAKIS: Very good.

MR. POWERS: I very much appreciated your presentation. The information was enlightening to us and extraordinarily useful. I wish you well on whatever follow-on efforts you're taking.

MR. APOSTOLAKIS: Keep your passion burning.

MR. KRESS: And get some sleep.

(Laughter.)

MR. POWERS: I appreciate Nathan sharing that material with us because it was helpful on many, many scores.

MR. SIU: While Bruce and Dave are packing up here, another thing I wanted to mention by the way, you had asked about, if you will, the gaps in our program.

MR. POWERS: Yes.

MR. SIU: What you see in Erasmia's slide I think are, most of those are anticipatory activities. For example, the latent errors, we talk about extended applications for LOPAR, and shut down long-term recovery actions, level two HRA. These are things that we are anticipating that we're going to need to improve methods and tools for. Obviously, we've got stuff being used now. But the question is can we do better.

So the list you see in the table that was displayed is our shot at what we think the needs are. We have something that's very global on upgrade and advanced reactors. Maybe it's not specifically enough --

MR. APOSTOLAKIS: On page 19 of the

plan, you have a number of tasks.

MR. SIU: That's right.

MR. APOSTOLAKIS: These are the same?

MR. SIU: Those are the same. We just tried to map those into different needs.

MR. APOSTOLAKIS: Okay.

MS. LOIS: Except, a few tasks are not there such as standards development, vulnerability, or --

MR. SIU: That's right. So, there are a couple of things that have been added on the table.

MR. APOSTOLAKIS: There is also some acronyms at the end WSMS 1-2.

MR. SIU: Yes.

MR. APOSTOLAKIS: RSWER 1-3. Is this a secret code?

MR. SIU: No, this is our risk informed regulatory --

MR. APOSTOLAKIS: That's the RIRIP. I understand that.

MR. SIU: Okay. And it has specific activities in it, so these are teed to those activities. So when there are activities that need HRA support --

MR. APOSTOLAKIS: I have two questions

regarding this table, appreciating the fact that it's in a document dated May, 2001.

MR. SIU: Yes.

MR. APOSTOLAKIS: One is, would Davis-Besse or the Indian Point incidents, among others perhaps, change these tasks because that was done under a different context?

And second, I understand you plan to have an updated version early next year. I think that developing performance indicators for human performance is important. Maybe you can try to accommodate this somewhere there because the reactor oversight process is in desperate need of this. It does relate of course to Davis-Besse and Indian Point again.

Again, I don't mean performance indicators in the sense that they are already in the ROP for reactor safety like the frequency of transients of the frequency of this and that because you may not be dealing with frequencies.

But when the guy there to inspect, is there an indicator that he can look at? Like I mentioned, a number of items deferred for example. Does it make sense universally? But I really think these are what the issues are these days. So other than that, it seems to be a fairly comprehensive list of various tasks and theories.

And one last comment I keep forgetting. Jim Riesen I think makes a distinction between latent errors and latent conditions. I think latent conditions is probably more appropriate because they're not necessarily errors. They create the context within which -- it's a broader term. I think conditions is a little better.

I have a few other comments on the report, but the report seems to be obsolete anyways. For example, on page 20, there are some deadlines.

MR. SIU: Yes.

MR. APOSTOLAKIS: "Develop HRA research lessons to support risk informed regulatory applications", September, 2001. Has that been done?

(No response.)

MR. APOSTOLAKIS: "Develop initial guidance" -- well, there are certain things that are supposed to be done by now.

MR. SIU: Right.

MR. APOSTOLAKIS: And I wonder whether they have been done and if we could get copies of them.

MR. SIU: And as Erasmia indicated, the

two things that are coming down in terms of quantification uncertainty, and that's in the context of ATHENA, and what was the other one? Oh, PTS was the other one.

But yes, the plan will be updated. Obviously, one of the motivating factors behind that is because the dates need to be updated.

MR. ROSEN: When Scott came at the very beginning, he tantalized us by saying we may need to reengage the Commission on Davis-Besse, based on the Davis-Besse experience. Is there more that you can say about that? Is there a whole piece of this presentation that hasn't been given or what?

We have said a lot about it. George has spoken, I have spoken, and people have said things around the table, but you haven't said anything.

MR. SIU: We haven't done significant work in the area. The decision that we would think about reengaging was a very, very recent decision. This is a statement of intention I think, and we're going to start looking at that.

MR. ROSEN: Perhaps you might need some input, more than we've given you already.

MR. SIU: Sure, yes.

MR. ROSEN: One of the pieces of input I've mentioned before was the leading indicator program at EPRI. And the offer that the EPRI management made to me at least was that they would be pleased to come here and brief the staff and the ACRS if we wanted to and the subcommittee on what that program does.

To me, in looking at it and talking to one of the leading utilities that's using it, it's the first piece of data collection that in mind the industry has done that actually has a chance of getting us an early signal that the decision-making environment in a utility is degrading, that tasks are not being done well. I think that's a piece of this problem, an organizational performance problem, that we're labeling safety culture.

The other thing is we talked about the need for indicators. Well, even leading gives you these indicators, to sum it up and look at things. But George mentioned the modifications that are preferred. To me, just corrective actions that are preferred that are significant is another one of those indicators that are important.

Of course, the classic one in corrective actions is the failure to preclude recurrence. The very essence of a corrective action program is that when something happens, you do enough to make sure it doesn't happen again. And when it does happen again, if it does, there ought to be a big signal to management that something is wrong with the corrective action program.

And the third one is a question of, in an environment that is degrading, in a place where there are a lot of good people, those people begin to come forward. In a safety conscious work environment, those people come forward with complaints that we're not doing a good job. How many there are and what management does with them is another indicator of the degrading environment or an improving one.

So, there are some rich data sources to mine. To me, working on how good the operators do in a known transient -- and it's a good thing to do, but it's working on a problem that we've worked, and worked, and worked. We haven't worked at all hardly on this other end of the real risk spectrum.

MR. POWERS: I expected you to -- I mean you certainly mentioned this leading indicator program and its value. But I expected you to go on and comment on this whole business of cross-cutting issue, and how is the HF and HRA program addressing this?

I mean you've got this statement. This is a cross-cutting, and it just kind of sits there.

What do we do with that? I mean is there nothing that can be done?

MR. APOSTOLAKIS: In fact, there was a hypothesis which the ACRS several times in its letters said it's an untested hypothesis. That is there is a problem with any one of these three cross-cutting issues, we will see it in the performance of the hardware so why worry about it.

MR. ROSEN: To my view, that is exactly correct. If there is a problem with cross-cutting issues, you will see it in the hardware. The trouble with that is that you will see it too late.

MR. APOSTOLAKIS: Too late --

MR. SIEBER: The other problem with that is you're not going to find just one issue. You're going to have a whole series of latent defects in the plant that will take you millions of dollars to correct and years to correct.

MR. ROSEN: And the other point that you will apply but didn't make is that if you have a whole raft of these defects, on a bad day they'll all line up wrong. Then, you can have a very serious circumstance.

MR. APOSTOLAKIS: Like Swiss cheese.

MR. ROSEN: The barriers all have holes,

and then one day the barriers all line up exactly right and you get this light --

MR. APOSTOLAKIS: When they say "model", that's what they mean.

MR. POWERS: The ROP people, when they respond to us -- and this is untested hypothesis -- said "yes, we're going to test it", I don't know how they can test it without you people being involved.

MR. PERSENSKY: To some extent, the report that I mentioned that talked about the ROP study, which is NUREG CR-6775, was a response to that question. They did look at how performance was characterized in the reactor oversight process and how it lined up ASP events in the past. That did identify a number of issues.

The one that seems to have the highest payoff right now is the improvement to the corrective action program inspection module. What we're doing is looking at the inspection module.

It did mention some other issues that came up. For instance in the area of latent errors, the possibility of some changes to the sampling under the maintenance program, the maintenance rule. There are certain things like we look only at certain high-risk equipment. Whereas if you look back at some of the accidents, there were other pieces of equipment, that when they lined up properly, caused the problem. So there may be some other changes. We proposed that we look at that.

Also, the issue of communications is one of things that came out as a major problem. But we do have in fact right now, since that work was done, we have come out with a couple of reports in conjunction with NRR on trying to improve the communications' look at things. So, we didn't go back on that.

We also mentioned what might be called safety culture. We made the point in our letter that there is a current restriction on doing much work in that area. But as Nathan said, there's very recent direction that we may be going back and looking at that.

So, there are a number of things that came out. If you look at the three cross-cutting issues -- one is the corrective action program, one is human performance, and the other is safety conscious work environment -- they're all human factors.

MR. POWERS: They're all one thing.

MR. PERSENSKY: They're all one thing. They all come down to a human or organizational or whatever factor.

But, we have done some work in that area. We haven't done perhaps the definitive work, and I think we need to follow it up with more recent looks at things like Davis-Besse.

MR. APOSTOLAKIS: I have a question to that regarding the plan. There was a conceptual problem I had with this.

It says that the methods for modeling or post-initiate actions are in not fairly good shape, but they are more advanced than methods to treat organizational factors. Now we all agree that organizational factors, as the report says, strongly affect those actions.

So how can a method or action be more advanced than methods for dealing with something that's necessary to understand the actions themselves? If I do organizational factors poorly, don't I automatically do human error modeling for which organizational factors are important?

MR. SIU: Or put it another way. Perhaps you're dealing with some sort of an average level. I mean you're able to distinguish between the characteristics of different organizations other than how they affect things that we do try to address in the analysis. Like when we make observations of crews and see how they actually respond to a particular event, or you look at past history and factor that into your analysis. But that's not a direct analysis of --

MR. POWERS: I think you see it in a great deal of variance in the data that you collect on human performance. If you don't understand everything and you project it under the space that you understand, you're going to see a large amount of error. And that's what they see.

MR. ROSEN: They do not understand what the source of the variance is.

MR. POWERS: That's right.

MR. APOSTOLAKIS: What I think really is said here is that there has been a lot of attention paid to modeling human actions. There are a number of models. In that sense, it's more advanced than the other stuff where you have maybe a couple of models. But, it's causing effect. If the cause is not modeled well, the effect is not modeled well. But again, I do bring it very serious.

I have a question for the Chairman.

MR. POWERS: Yes.

MR. APOSTOLAKIS: What time does the coffee shop downstairs close?

MR. POWERS: I believe you will not be able to get coffee after four o'clock.

MR. APOSTOLAKIS: Okay.

MR. POWERS: Let me ask this question. I had five categories of questions that we posed after lunch: the big scheme of needs, tools, organization safety culture, and indictors, development of HRA models and view of existing models, and state of the art. I think we have addressed those in our discussions.

Do you want to take a break for 15 minutes, get your coffee, come back, and do a roundtable for the points that we want to make?

MR. APOSTOLAKIS: Sure. I think that's good.

MR. POWERS: Or do you want to interrogate these gentlemen and lady further?

MR. APOSTOLAKIS: No, but I'm sure they're going to stay.

MR. POWERS: They're more than welcome to stay because I think we're going to need their continuing help.

But I will emphasis that on the time that I have been on the ACRS, this has been the most enjoyable, pleasant, and well thought out meeting in the area of human reliability and human factors that I've ever attended. It comes off with a more optimistic note than I've ever enjoyed.

So, I congratulate you on an excellent presentation to the subcommittee, which almost amounts to the full committee. You will be surprised to find that Dr. Shack, who is not here, has strong views on this subject and will probably take an orthogonal view on everything.

We do need to chat a little bit about what to present to the full committee.

We're done. I think at this point I'm going to close the meeting, and adjourn this transcriber at this point. We'll come back after coffee and discuss a little bit about what to present to the full committee and what we think ought to appear in the letter. So why don't we reassemble at twenty-five of the hour.

The meeting is closed.

(Whereupon, the above-entitled meeting concluded at 3:19 p.m.)

CERTIFICATE

This is to certify that the attached proceedings before the United States Nuclear Regulatory Commission in the matter of:

Name of Proceeding: Advisory Committee on Reactor Safeguards Subcommittee on Human Factors

Docket Number: N/A

Location: Rockville, Maryland

were held as herein appears, and that this is the original transcript thereof for the file of the United States Nuclear Regulatory Commission taken by me and, thereafter reduced to typewriting by me or under the direction of the court reporting company, and that the transcript is a true and accurate record of the foregoing proceedings.

/s/ Matthew Needham

Matthew Needham
Official Reporter
Neal R. Gross & Co., Inc.

ACRS Issues with Risk-Informed Regulatory Activities

Discussion with Advisory Committee on Reactor Safeguards Subcommittee

September 9, 2002

PURPOSE

- To initiate dialog with the Committee regarding their recommendations on Regulatory Guide 1.174

- To fully understand the Committee's concerns

- To come to a common understanding on a path forward

STAFF'S UNDERSTANDING OF COMMITTEE'S CONCERNS

(Based on letter, Commission briefing, and Committee meeting)

- Two types of concerns:
  - Policy/Technical Issues
  - Public Confidence Issues

SUMMARY OF COMMITTEE CONCERNS

(Staff understanding based on letter, Commission briefing, and Committee meeting)

- Policy/Technical Concerns:
  - Regulatory guidance incomplete in addressing all sources of risk of nuclear power plants
  - Uncertainty not adequately addressed
  - Risk metrics incomplete
- Public Confidence Concerns:
  - "Rigorous" PRAs are needed for public confidence

NEXT STEPS......

- Continue dialog with ACRS
- Hold stakeholder public meetings
- Revise regulatory guidance where appropriate

Ron Frahm RF
Donald Hickman DH
Douglas Coe DC
Cindy Carpenter CC

DC
F
H

ACRS ROP Subcommittee Meeting

September 9, 2002

Ronald Frahm

Donald Hickman

Douglas Coe

SRM Dated 12/20/2001

"The staff, with ACRS input, should provide recommendations for resolving, in a transparent manner, apparent conflicts and discrepancies between aspects of the revised reactor oversight process that are risk-informed (e.g., significance determination process) and those that are performance based (e.g., performance indicators)."

Oversight Process

- ROP Regulatory Framework Includes Seven Equivalent Cornerstones of Safety
- Staff Actions are Based on Plant Performance per the Action Matrix (PIs and Inspection Findings)
- Assessment Reviews Performed on a Continuous, Quarterly, and Annual Basis For All Plants
- Plants Appear to be Receiving the Appropriate Level of Oversight

Staff Approach

- Having Both Risk-Informed and Performance-Based Thresholds Provides a Balanced Approach
- Remain Objective, Risk-Informed, Understandable, and Predictable, and Meet the 4 Strategic Performance Goals
- Seek Continued Improvements Through the ROP Self-Assessment and Feedback Processes and Interactions with Other Stakeholders

Potential PI Improvements

- Conducting a Pilot Program for the Mitigating System Performance Index
- Continuing to Improve and Develop Other Performance Indicators

Potential SDP Improvements

- Developed and Started Implementation of the SDP Improvement Plan (Emphasis on Improving Timeliness)
- Recently Formed the SDP Task Group To Address Ongoing SDP Concerns
- Continuing to Improve and Develop Specific SDPs

Staff Conclusions

- ROP is Working, Though Continued Incremental Improvements are Expected
- Recognized the Need for a Clearer Basis for PIs and SDPs and are Creating a Detailed Basis Document
- Plan to Work With RES to Explore the Use of Decision Theory for the ROP

Proposed Future Plans

- Full Committee Meeting in December 2002 (or February 2003) Followed by an ACRS Letter
- Annual SECY Paper in March 2003
- Continued Information Exchange Between ACRS and the Staff

Back Up Slides

Exhibit 1: REGULATORY FRAMEWORK

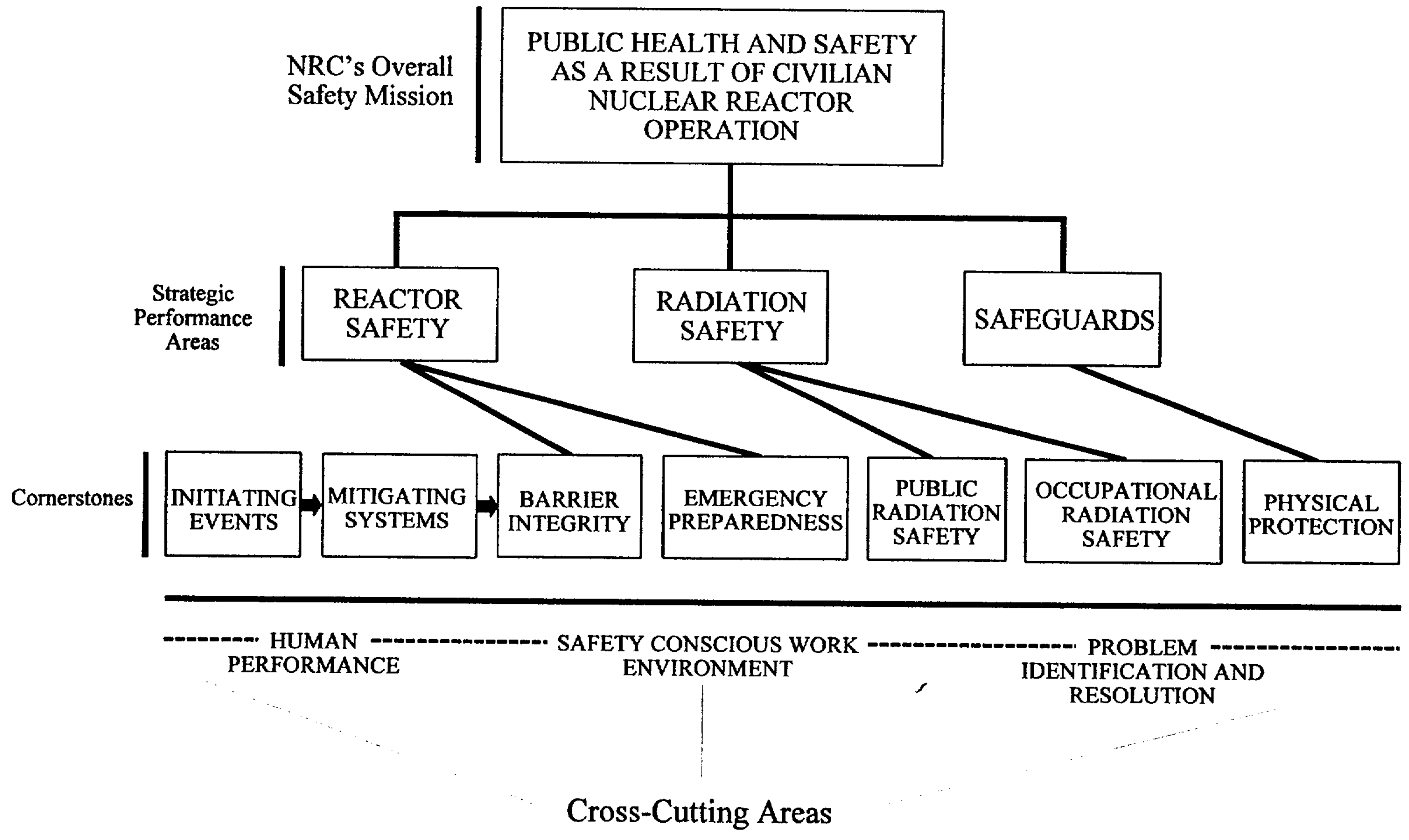


Exhibit 5 - ACTION MATRIX

		Licensee Response Column	Regulatory Response Column	Degraded Cornerstone Column	Multiple/ Repetitive Degraded Cornerstone Column	Unacceptable Performance Column
RESULTS		All Assessment Inputs (Performance Indicators (PIs) and Inspection Findings) Green; Cornerstone Objectives Fully Met	One or Two White Inputs (in different cornerstones) in a Strategic Performance Area; Cornerstone Objectives Fully Met	One Degraded Cornerstone (2 White Inputs or 1 Yellow Input) or any 3 White Inputs in a Strategic Performance Area; Cornerstone Objectives Met with Moderate Degradation in Safety Performance	Repetitive Degraded Cornerstone, Multiple Degraded Cornerstones, Multiple Yellow Inputs, or 1 Red Input; Cornerstone Objectives Met with Longstanding Issues or Significant Degradation in Safety Performance	Overall Unacceptable Performance; Plants Not Permitted to Operate Within this Band, Unacceptable Margin to Safety
RESPONSE	Regulatory Performance Meeting	None	Branch Chief (BC) or Division Director (DD) Meet with Licensee	DD or Regional Administrator (RA) Meet with Licensee	RA (or EDO) Meet with Senior Licensee Management	Commission meeting with Senior Licensee Management
	Licensee Action	Licensee Corrective Action	Licensee root cause evaluation and corrective action with NRC Oversight	Licensee cumulative root cause evaluation with NRC Oversight	Licensee Performance Improvement Plan with NRC Oversight	
	NRC Inspection	Risk-Informed Baseline Inspection Program	Baseline and supplemental inspection procedure 95001	Baseline and supplemental inspection procedure 95002	Baseline and supplemental inspection procedure 95003	
	Regulatory Actions[1]	None	Supplemental inspection only	Supplemental inspection only	-10 CFR 2.204 DFI -10 CFR 50.54(f) Letter - CAL/Order	Order to Modify, Suspend, or Revoke Licensed Activities
COMMUNICATION	Assessment Letters	BC or DD review/sign assessment report (w/ inspection plan)	DD review/sign assessment report (w/ inspection plan)	RA review/sign assessment report (w/ inspection plan)	RA review/sign assessment report (w/ inspection plan)	
	Annual Public Meeting	SRI or BC Meet with Licensee	BC or DD Meet with Licensee	RA (or designee) Discuss Performance with Licensee	EDO Discuss Performance with Senior Licensee Management	
	Commission Involvement	None	None	None	Plant discussed at AARM	Commission Meeting with Senior Licensee Management
	INCREASING SAFETY SIGNIFICANCE --------->					

Note 1: The regulatory actions for plants in the Multiple/Repetitive Degraded Cornerstone column are not mandatory agency actions. However, the regional office should consider each of these regulatory actions when significant new information regarding licensee performance becomes available.

ROP Program Documents

- MD 8.13 Reactor Oversight Process
- MC 0608 Performance Indicator Program
- MC 0609 Significance Determination Process
- MC 0305 Assessment Program
- MC 0307 ROP Self-Assessment Program
- ROP Basis Document
- Annual SECY Papers

Performance Indicators

Cornerstone	Threshold Method
Initiating Events	G/W - PB, W/Y/R - RI
Mitigating Systems	G/W - PB, W/Y/R - RI
Barrier Integrity	PB using Risk Info
Emergency Preparedness	Performance-Based
Occupational Rad Safety	Performance-Based
Public Radiation Safety	Performance-Based
Physical Protection	Performance-Based

SDP / Inspection Findings

Cornerstone	Threshold Method
Initiating Events	Risk-Informed
Mitigating Systems	Risk-Informed
Barrier Integrity	Risk-Informed
Emergency Preparedness	Performance-Based
Occupational Rad Safety	PB using Risk Info
Public Radiation Safety	PB using Risk Info
Physical Protection	Performance-Based

United States
Nuclear Regulatory Commission

NRC Human Reliability Analysis Research Program

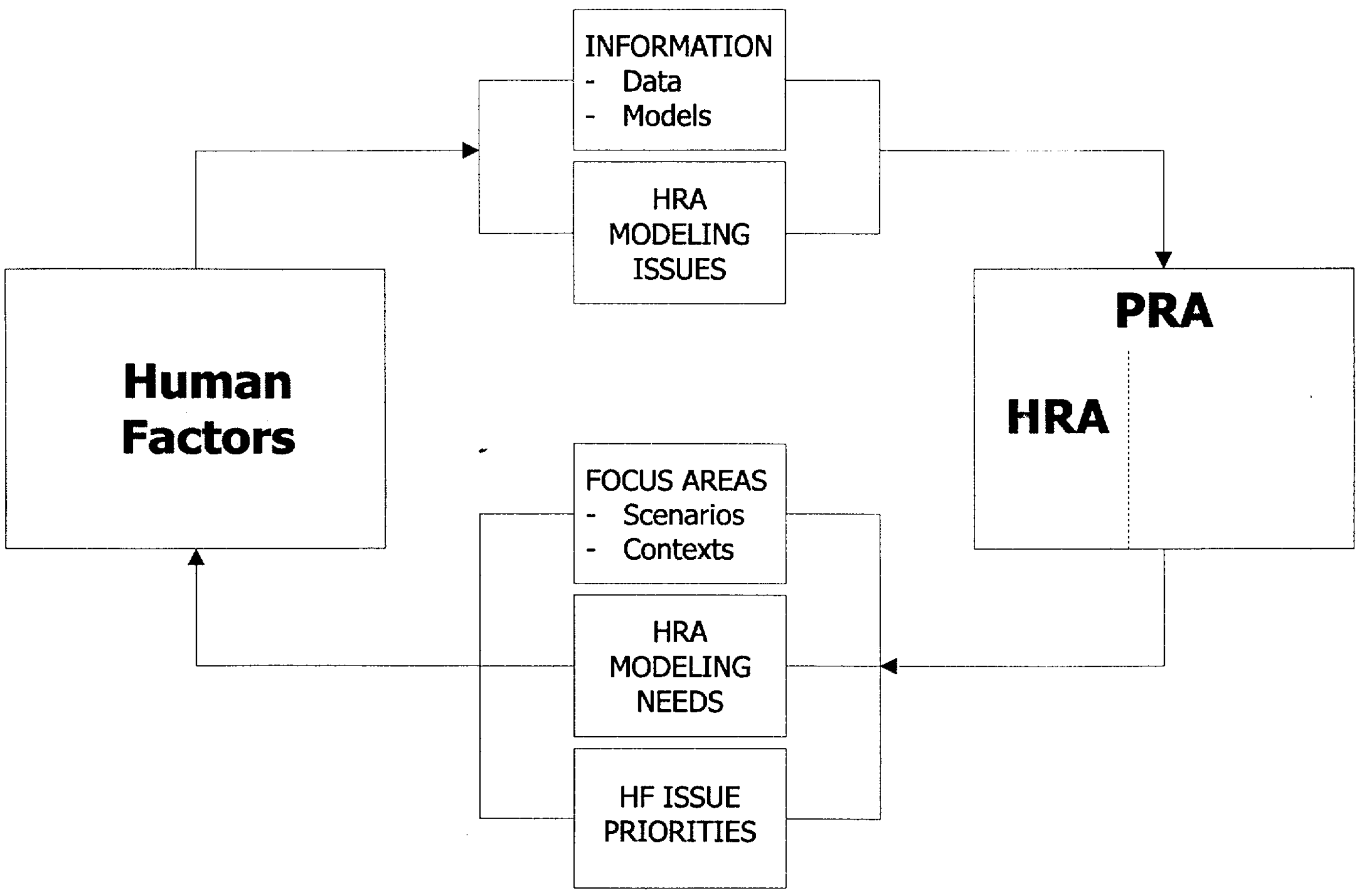
Erasmia Lois
Probabilistic Risk Analysis Branch
Office of Nuclear Regulatory Research

Presented to
Subcommittee on Human Factors
Advisory Committee on Reactor Safeguards
USNRC Headquarters • Rockville, MD • 10th September 2002

Briefing Outline

- **HRA and HF relationships and interactions**
- **Overall HRA Plan status**
- **Currently planned activities**
- **Specific activities**
  - **Advanced reactors**
  - **Data collection and analysis**

HRA and HF Relationship



Overall HRA Plan status

- **Last update: May 2001**
- **Covers 2001-2005**
- **Some activities near completion**
  - **PTS HRA**
  - **Quantification including uncertainty**
- **Remaining activities underway or planned**
- **Expect plan to be updated, by January 2003**
  - **date/milestone updates**
  - **projects deleted/added**
    - **Vulnerability Assessment**
    - **HRA standards**
  - **5 year**
  - **Broader in terms of activity description**

HRA Activities

	Conventional Reactors	Advanced Reactors	Materials And Waste	Security and Safeguards
Rules	PTS			Fitness for Duty
Licensing	▪Fire ▪SGTR ▪Aging Cables	Upgraded & Advanced Reactors	▪Dry Cask ▪Other support	Vulnerability Assessment
Monitoring (e.g., ROP Event Analysis Issue Identification)	SPAR Models			
Infra-structure	<u>Methods and Tools</u> ▪Data Collection and Analysis ▪Quantification and Uncertainty ▪Latent Errors in HRA ▪Extended Applications: ▪Reactor Synergisms and HRA ▪Formalized Methods: Screening, Individual and Crew Modeling <u>Implementation</u> ▪ Guidance, Standards			

HRA for Upgraded and Advanced Reactors

- **Objective:** Determine if any improvements are needed to incorporate the influence of human performance in PRAs for upgraded and advanced reactors
- **Potential technical issues**
  - reduced staff, the changing role of the operator
  - new control room design
  - multiple modules
  - long-term recovery
- **Products:**
  - issue identification
  - methods and tools
  - guidance
  - HRA
- **Plan:** initiate work in 2003

Data Collection & Analysis

- **Objectives**
  - Determine data needs for HRA
  - Collect and analyze data to support HRA model development and quantification
- **Work performed at INEEL**
  - Funded by HF and HRA Programs
  - Currently focuses on needs of the Quantification Task (Sandia)
  - Supports/interfaces with CSNI activities on data collection
  - Collaborates with Halden

Data Collection and Analysis

- **Approach**
  - Characterize the information neede to apply HRA methods and to estimate human er r probabilities
    - Identify concepts and terms used in t various methods
    - Determine concept/term commonalties
  - Identify and evaluate data sources for usefulness
  - Develop methods to utilize information/data
  - Develop methods to estimate develop human error probabilities

Data Collection and Analysis

- **Current activities**
  - development of glossary is underway
  - data sources under examination
    - data in open psychological literature
    - simulator data
  - One specific source--data generated for the advanced reactor staffing study will be discussed in some detail today

United States

Nuclear Regulatory Commission

Human Factors Research at the US Nuclear Regulatory Commission

J. J. Persensky

Office of Nuclear Regulatory Research

Presented to

Subcommittee on Human Factors

Advisory Committee on Reactor Safeguards

USNRC Headquarters • Rockville, MD • 10th September 2002

Role of the Human Factors Research Program at the USNRC

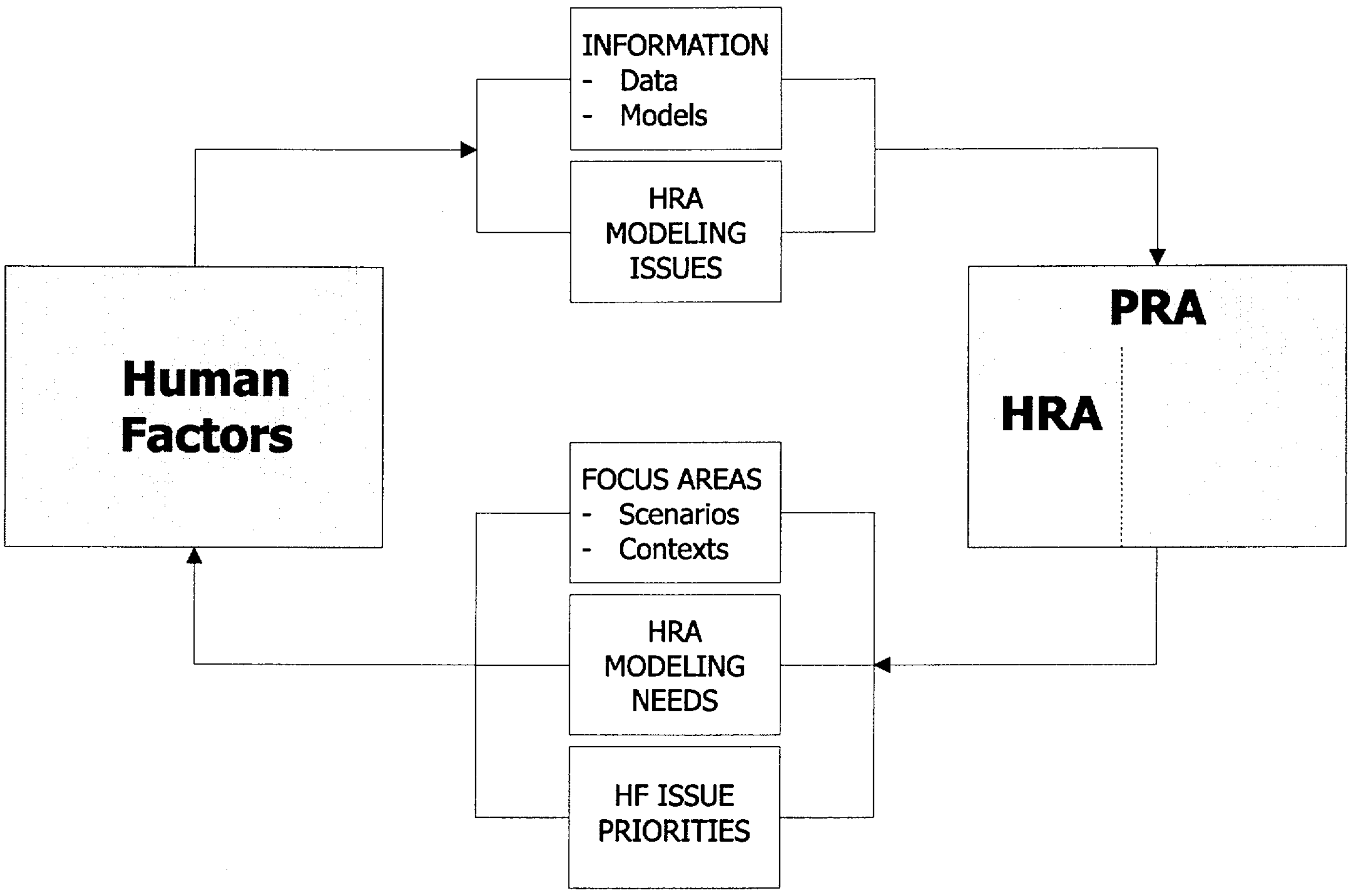
- Provide NRR, NMSS and NSIR staff with tools, developed from the best available technical bases, necessary to accomplish their licensing and monitoring tasks.

- Ensure that nuclear facility personnel have the tools, knowledge, information, capability, work processes and working environment (physical and organizational) to safely and efficiently perform their tasks.

SECY- 01-0196

- Sunset the "Program on Human Performance in Nuclear Power Plant Safety" as an independent document
  - RES participation only
  - Limited Resources
- Integrate activities into Human Reliability Research Plan or Digital I&C Research Plan
- Presented status of efforts from SECY-00-0053

HRA & HF Relationship



Human Factors Activities and Needs

	Conventional Reactors	Advanced Reactors	Materials	Security and Safeguards
Rules	Fatigue			Fitness for Duty
Licensing	▪SRP Chpt. 18 ▪Staffing	▪Staffing ▪Licensing and Training	SRP ▪Development ▪ Review	
Monitoring	ROP: Risk-inform CAP		Inspection Manual Update	
Infrastructure	▪Data Collection and Analysis ▪Halden Reactor Project ▪Risk Communications ▪HF infrastructure for Advanced Reactors ▪Human Factors Role in Security and Safeguards ▪Consensus Standards ▪International Activities			

SFN 5

United States
Nuclear Regulatory Commission

NRC Human Reliability Analysis and Human Factors Research Programs: *Overview*

Scott Newberry and Farouk Eltawila

Office of Nuclear Regulatory Research

Presented to
Subcommittee on Human Factors
Advisory Committee on Reactor Safeguards
USNRC Headquarters • Rockville, MD • 10th September 2002

Briefing Objectives

- **Provide overview of NRC's human reliability analysis (HRA) and human factors (HF) research programs**
  - **Activities**
  - **Relationship and interactions**
- **Obtain feedback to inform ongoing planning activities**

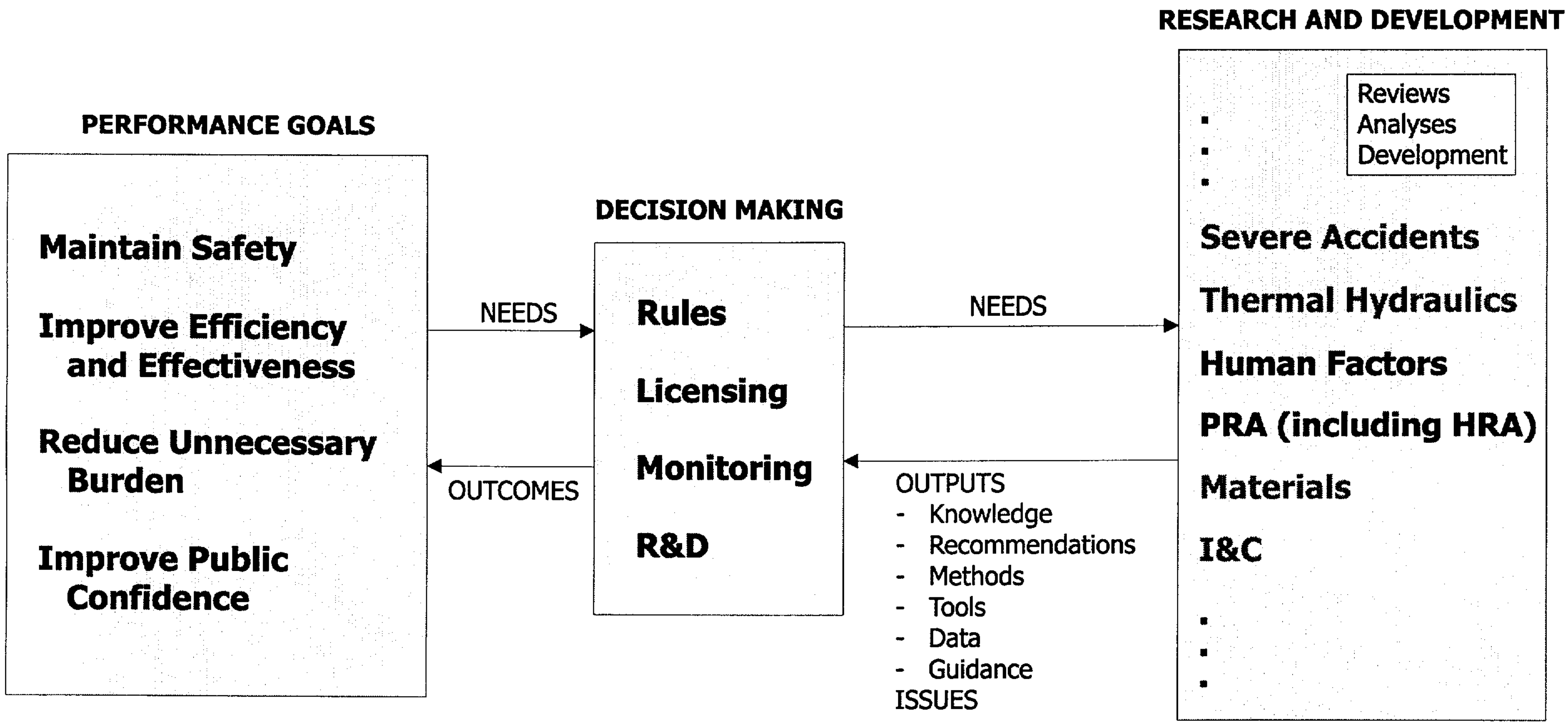
Briefing Outline

- **Why HRA and HF research and development?**
- **Discipline and program relationships**
- **HRA needs and activities**
- **HF needs and activities**
- **Joint research: a data-collection example**

Why HRA and HF R&D?

- **Agency needs**
- **Operating event experience**
- **PRA experience**
- **Trends and future events**
- **Typical questions**
- **Activity types**

Supporting Agency Needs



Programs

	Human Factors	HRA
Organization	RES/DSARE	RES/DRAA
Technical Lead	J. Persensky	E. Lois
Plan	SECY-01-0196 (Nov. 2001)	HRA Research Program Plan (May 2001)

Idaho National Engineering and Environmental Laboratory

Using Simulators in Human Factors Research

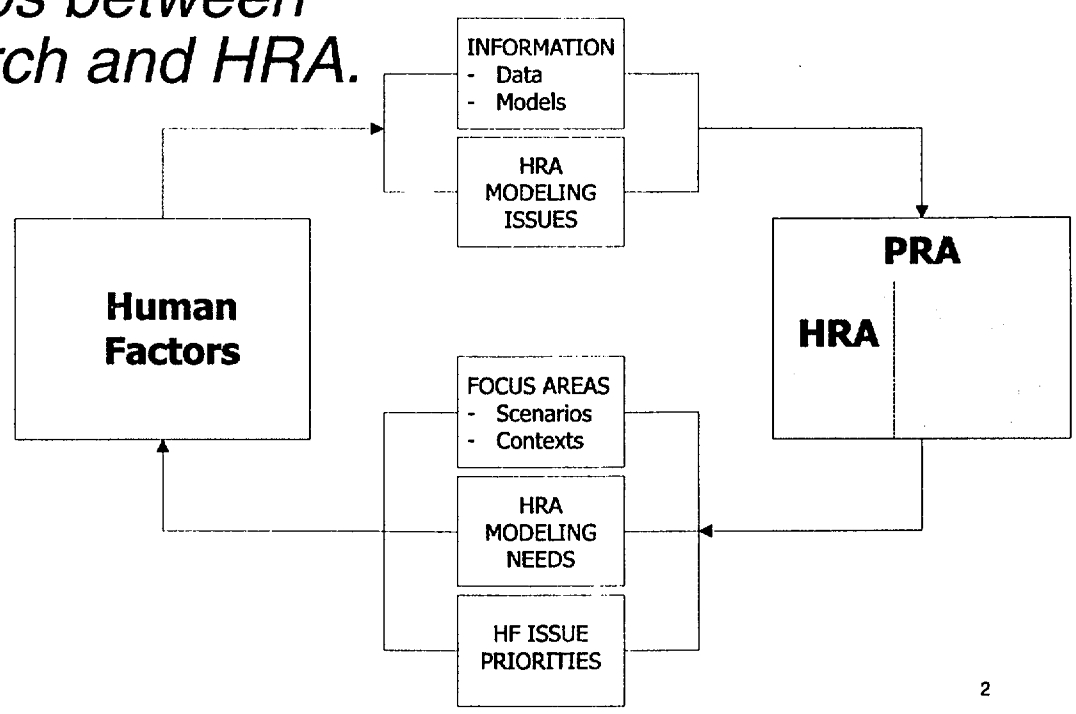
Linking Human Factors and Human Reliability Analysis

Bruce P. Hallbert
Department Manager
Human and Intelligent Systems Research

September 10, 2002

Purpose

- *Presents a study of human performance in which data are present to inform HRA activities*
- *Illustrates relationships between human factors research and HRA.*



Outline

- *Discuss the potential for simulators to support HRA.*
- *Overview of simulator-based research project*
  - *NRC-sponsored staffing study*
  - *Preliminary exploration of PSFs and performance*
- *Summarize results*
- *Discuss potential for HRA*

Developing HRA-Relevant information

- *Simulator studies can provide useful data for HRA,e.g.,*
  - *Relationships between PSFs, performance, and error*
  - *Hypothesis testing and model development*
  - *Benchmarking HRA methods*
- *Current HRA methods do not make full use of simulator data.*
- *Protocols are needed for collecting data and making inferences to support HRA (number of observations, types of plants, degree of realism, etc.)*

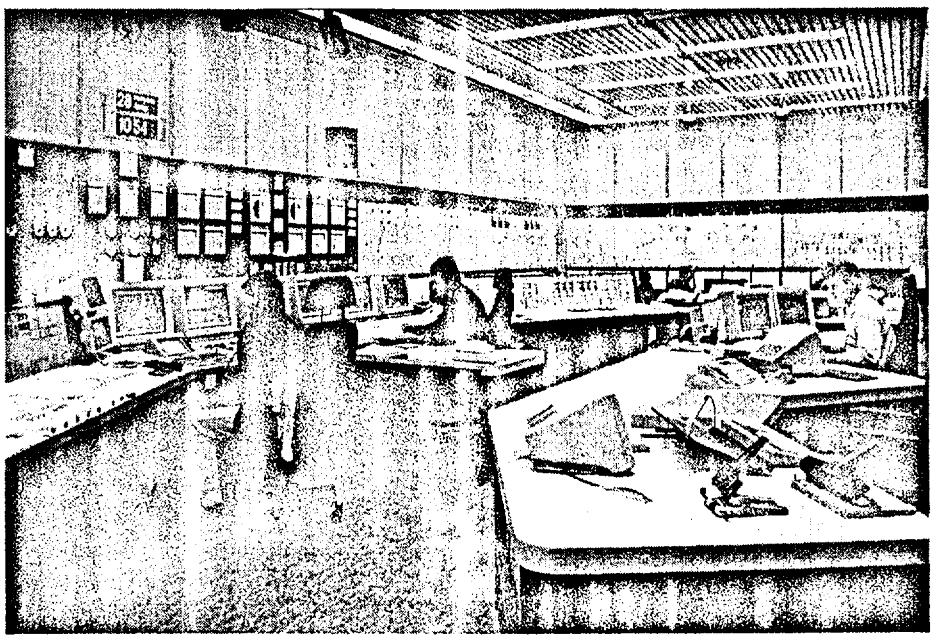
A Study of Control Room Staffing Levels for Advanced Reactors*

- *Study focused on 10 CFR 50.54 (m) and potential changes to CR staffing of future plants.*
- *Improvements in ease of performance through redundancy, passivity, diversity and automation.*
- *Need to better understand the performance implications of staffing and advanced plant performance.*
- *Conducted study of control room crew performance.*
- *Advanced and conventional plant benchmarks; crew staffing; T-H performance.*
- *Design basis scenarios: SGTR, ISLOCA, LOFW, LOOP, SG overfill.*
- *Evaluate two different CR staffing configurations (normal, minimum)*
- *Carried out at operating NPP training simulator (Loviisa) and advanced plant simulator (HAMMLAB)*

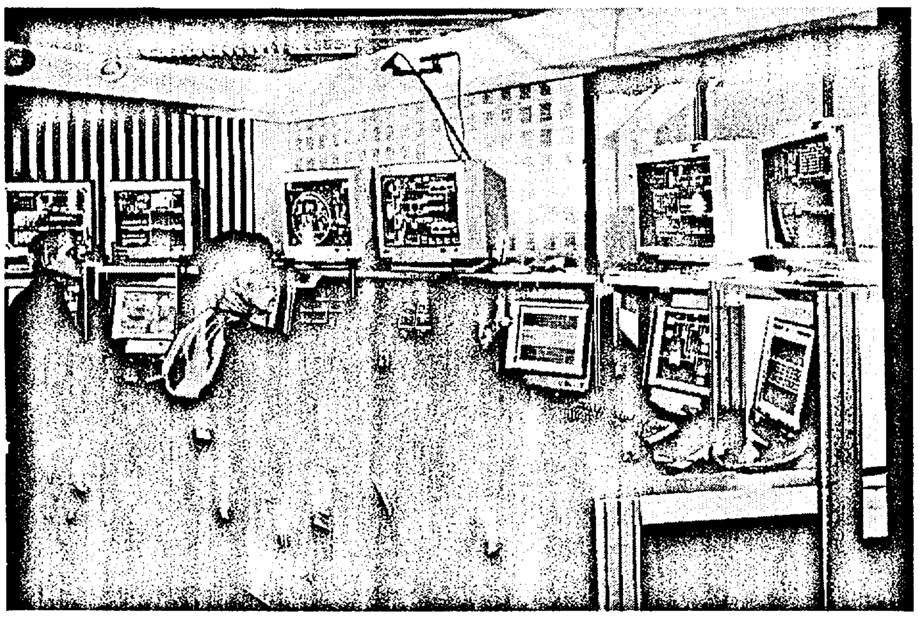
*NUREG/IA-0137 (2000)

Loviisa study phase

- Scenarios maximize similarities to Western PWRs (T-H, accident progression)
- Crews in study operate as crews in plant.
  - training
  - role.
- EOPs use symptom based approach.
- Normal crew = 4
- Minimum crew = 3



Halden study phase



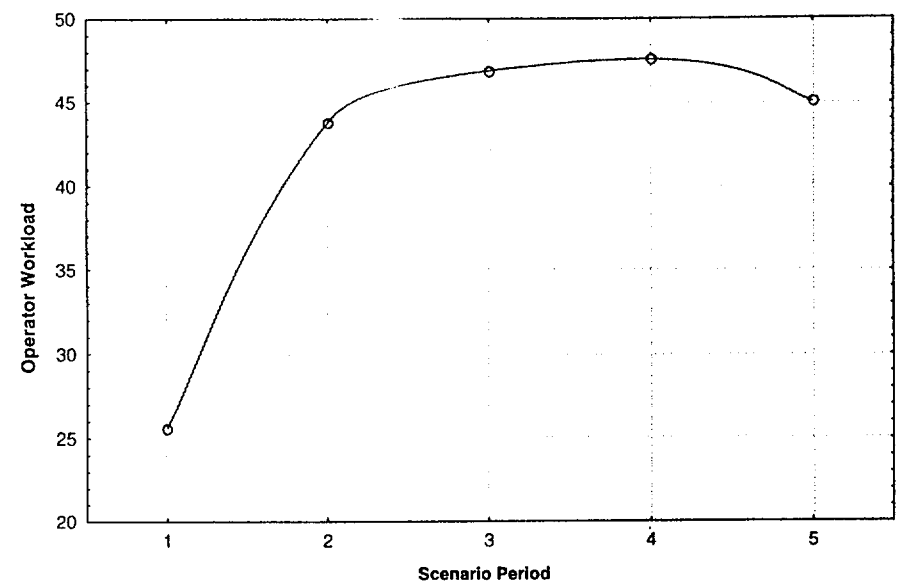
- Simulated plant based upon Loviisa with added automation to simulate passive system performance.
- Digital I&C – Common Overview, Alarms, process displays, SPDS
- Workstation arrangement following CR division of labor
- Normal crew = 4
- Minimum crew = 2 (dual role SS/RO-BOP)

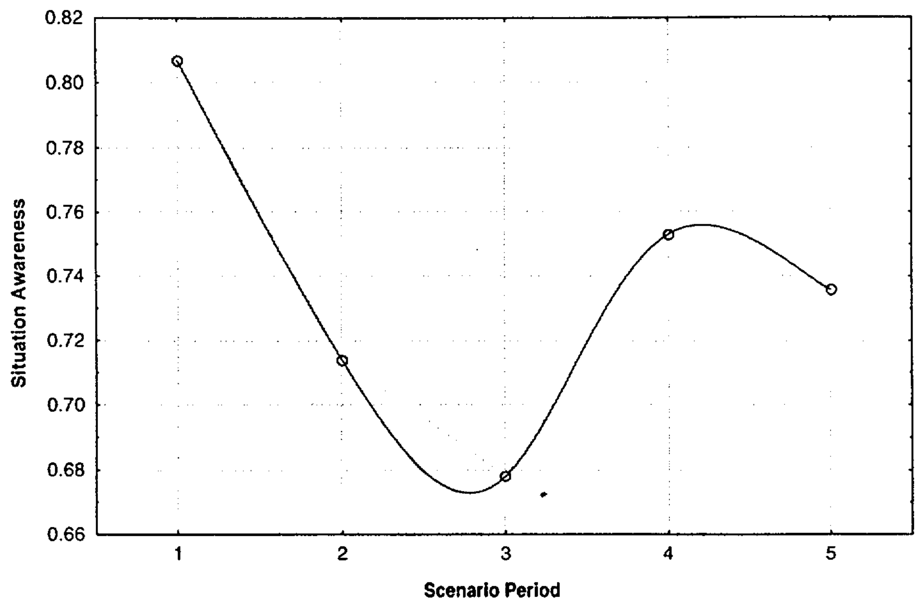
Data Collection

- *8 crews presented with 5 scenarios; 4 crews served in normal, 4 crews served in minimum staffing configuration*
- *Data collected on:*
  - *Subjective Workload (NASA TLX)*
  - *Team Performance (BARS)*
  - *Situation Awareness (SACRI)*
  - *Rated crew performance*
  - *Task completion, Plant parameters*
- *First 4 measures collected 4-5 times during each scenario*

Results

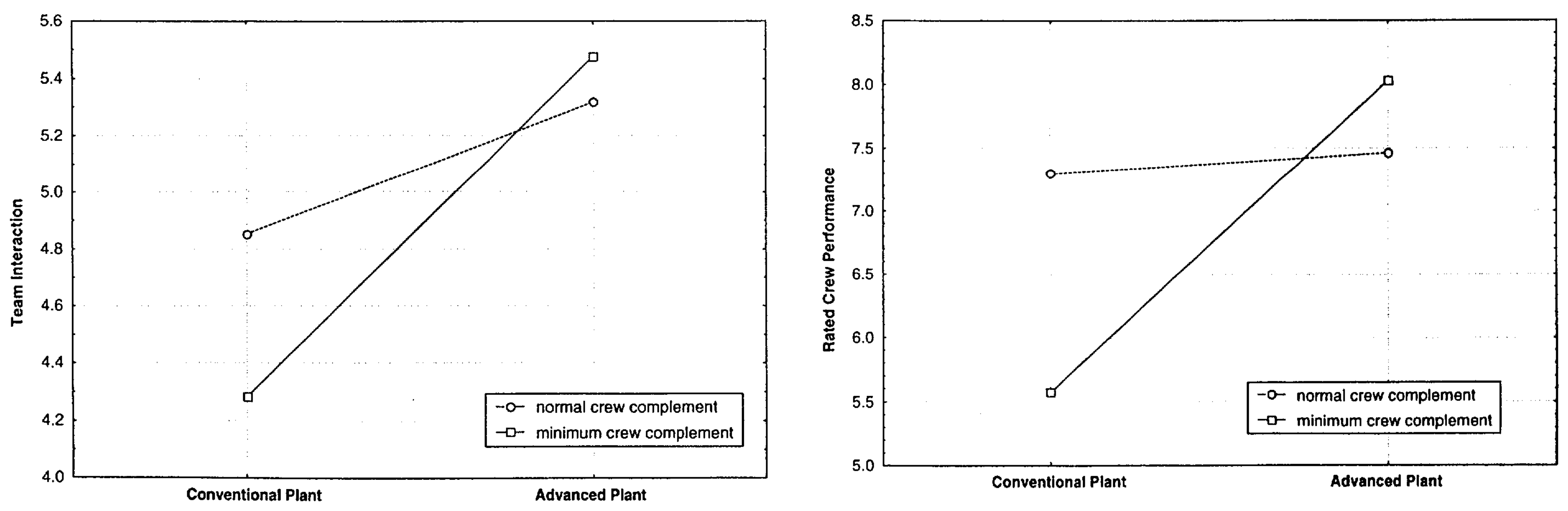
- Crews experienced high workload for extended periods of time.





- Situation Awareness dropped, followed by gradual recovery

- *Rated crew performance on transient management activities paralleled ratings of teamwork*



Embedded Study : Operator Performance and PSFs

- *Intuitive linkage between PSFs and operator performance.*
- *Types of PSFs and their effects on error rates vary among HRA methods.*
- *Assessment of PSFs estimated; uncertainty remains high in most applications.*
- *Need for better benchmarking and understanding of PSF relationship with performance.*
- *Linkage needed to build better models of failure.*

Purpose – Embedded study

- *Explore how data collected in human factors studies could support HRA.*
  - *Identify a set of PSFs that are predictive of crew performance.*
  - *Determine the weighting of these factors relative to one another.*
  - *Demonstrate a general model in which the PSFs can be expressed.*
  - *Measure the factors affecting the predictive validity of PSFs.*
  - *Replicate the results and model developments at different plants and at different times.*

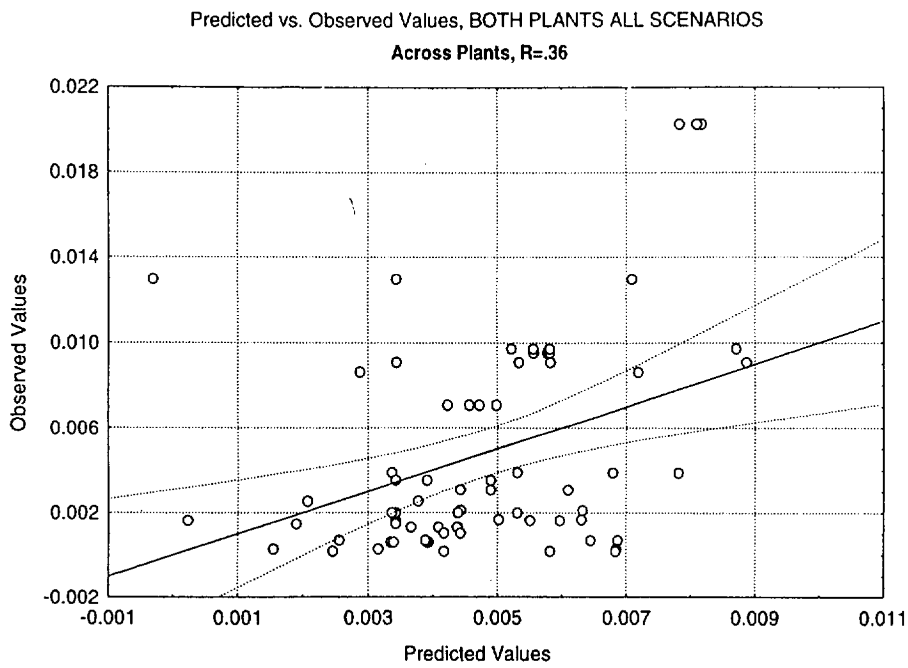
Approach

- *Set of 10 PSFs tested for use in predicting crew performance:*
  - *7 demonstrated predictive power: Procedures, Training, Stress, Workload, Information Available, System Feedback, HMI.*
- *Data collection instrument developed to measure "experienced" effects of PSFs.*
  - *Critical Tasks (mitigation)*
  - *Simulator trials*
  - *Rating by operators on the effect of PSFs on performance after scenario.*
- *Data collected on:*
  - *4 crews in U.S. plant (3 Scenarios used: LOFW, SG overfill, SB-LOCA) NUREG/CR-4966*
  - *4 crews in Loviisa and 4 crews in HAMMLAB*
  - *3 common scenarios: overheating, overcooling, loss of coolant*

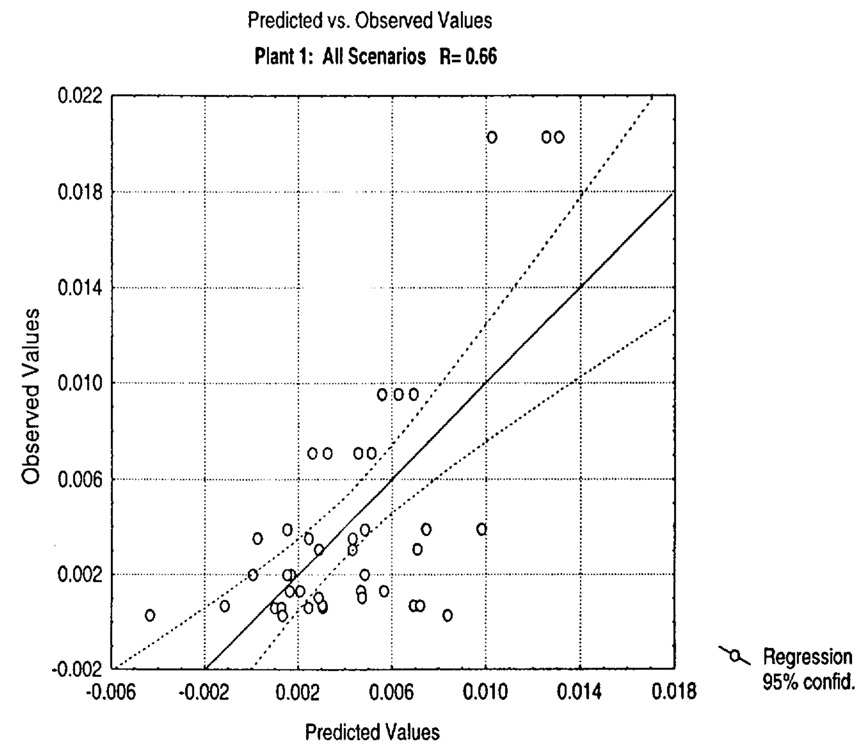
Results

- *Linear model with combined PSF weightings*
- $Y = a + b_1 x_1 + b_2 x_2 + \ldots + b_n x_n$
  - *Where Y= critical task mitigation performance*
- *Sensitive to scenario differences*
- *Sensitive to plant differences*
- *Demonstrated predictive ability (critical task performance)*

Plant-specific predictive power



All crews, all plants, all scenarios



Plant 1, all scenarios

Summary – Embedded Study

- *Demonstrated link between performance shaping factors and operator performance.*
- *Model, technique show promise for explaining variability in task performance*
  - *Limited to situations in which the defined set of PSFs are, in fact, influencing performance*
- *Potential use for data collection using plant-specific simulators*
  - *Time and training demands are small*
- *No assumptions about strength of relationship between PSFs and performance: empirically established in each data collection trial.*
- *Potential for reducing uncertainty in HRA.*

Summary

- Studies have already been conducted, and data collected that can be used to support HRA.
- New studies can be aimed specifically at HRA needs.
- Simulator studies can provide useful data for HRA.