Official Transcript of Proceedings

NUCLEAR REGULATORY COMMISSION

Title:	Advisory Committee on Reactor Safeguards Reliability and PRA Subcommittee - Open
Docket Number:	(n/a)
Location:	Rockville, Maryland
Date:	Tuesday, May 8, 2012

Work Order No.: NRC-1608 Pages 1-216

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

UNITED STATES OF AMERICA

NUCLEAR REGULATORY COMMISSION

+ + + + +

ADVISORY COMMITTEE ON REACTOR SAFEGUARDS

(ACRS)

+ + + + +

RELIABILITY AND PRA SUBCOMMITTEE

OPEN SESSION

+ + + + +

TUESDAY

MAY 8, 2012

+ + + + +

ROCKVILLE, MARYLAND

+ + + + +

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

SUBCOMMITTEE MEMBERS PRESENT:

DENNIS C. BLEY	Chairman
MICHAEL T. RYAN	Member
STEPHEN P. SCHULTZ	Member
JOHN W. STETKAR	Member

NRC STAFF PRESENT:

DEREK WIDMAYER, Designated Federal Official
VALERIE BARNES, RES/DRA
ANDY CAMPBELL, OE
RANI FRANOVICH, NRR
STEPHANIE MORROW, RES/DRA/HFRB
SEAN PETERS, RES
MARIA SCHWARTZ, OE-CRB
UNDINE SHOUP, NRR/DRA/AHPB
DAVE SOLORIO, OE

ALSO PRESENT:

G. KENNETH KOVES, INPO

I N D E X

P R O C E E D I N G S

Time: 8:30 a.m.

CHAIRMAN BLEY: The meeting will now come to order. This is a meeting of the Advisory Committee on Reactor Safeguards, Subcommittee on Reliability and Probabilistic Risk Assessment.

I am Dennis Bley, Chairman of the Subcommittee -- the sub-subcommittee. ACRS members in attendance include Mike Ryan, Steve Schultz, and John Stetkar, and if possible, we may have some others wander in and out during the day from the meeting next-door.

The purpose of this meeting is to discuss the Office of Research White Paper, Independent Evaluation of INPO's Nuclear Safety Culture Survey and Construct Validation Study. The White Paper was prepared to address concerns expressed by ACRS members and comments attached to the Committee's letter report of December 15, 2010, on the agency's final Safety Culture Policy Statement.

The meeting this morning is open except for portions of the INPO presentations which will discuss proprietary information, and when we get there, let us know.

The Subcommittee will gather information, analyze relevant issues and facts, and will formulate proposed positions of actions as appropriate. The Subcommittee will report its findings at an upcoming full Committee meeting, but at this time does not plan on issuing a letter report on this matter.

Derek Widmayer is the Designated Federal Official for this meeting. A transcript for the meeting is being kept, and will be made available on the web. It is requested that speakers first identify themselves, and speak with sufficient clarity and volume so that they can be readily heard. Thank you.

We will now proceed with the meeting, and I call upon Dave Solorio, Branch Chief in the Office of Enforcement, to open the proceedings.

MR. SOLORIO: Thank you very much.

My name is Dave Solorio, Branch Chief, Concerns Resolution Branch, Office of Enforcement. I am here to kick off the presentation to you all today.

OE was asked to take the lead for the development of the Safety Culture Policy Statement, and the majority of the staff who did the work for me within the Office of Enforcement. That is why you find me up here kicking this off.

While OE was asked to lead the effort to develop the Safety Culture Policy Statement by the

EDO, that was probably because of the longstanding expertise within the Office in the area of safety culture.

Prior policy statements had come out of the office. However, the Safety Culture Policy Statement is truly a superb team effort among a number of offices to work with our external stakeholders, to ensure we considered and utilized their input to develop the policy statement, and it has been recognized as a model effort by the Commission, FSME, NMSS, NRR, Research and the Regional Offices have all worked very collaboratively with OE to put together a policy that would speak to a wider spectrum of licensee environments.

What you will hear later from the nuclear power industry representative here today is that they have confirmed with their own independent efforts for their environment that to be the case.

We were here back last in November and December of 2010 presenting the work done to develop the Safety Culture Policy Statement as well as the industry's efforts to quantify the safety culture in their environment, and we appreciate the ACRS's endorsement of the work in their letter, and we are here today because of the Subcommittee's desire to have a more in depth discussion regarding some work done by the industry that attempted to quantify how safety culture might affect plant performance.

In addition, when preparing for this presentation, Research prepared the report that provided their evaluation of the industry's efforts in this area. The report was previously provided to the ACRS prior to today's meeting.

I want to also thank the Subcommittee Chairman and their staff for their efforts to help us prepare for today's presentation.

The research report was prepared in part as a result of the user need provided to Research to look at the industry's work. We saw the work that the industry as a chance to look at some empirical work and consider its results in the efforts undertaken to develop the Safety Culture Policy Statement.

So as my Director, Mr. Zimmerman, Mr. Royce Zimmerman, said last December when he was last in front of the Subcommittee, we used the user need as a vehicle to inform the development of the Safety Culture Policy Statement.

The Safety Culture Policy Statement was issued formally in the Federal Register last June, and the Commission's expectations are that all licensees and agreement state licensees strive to follow a positive safety culture as another way to contribute to enhancing safety of facilities regulated by the NRC and the agreement states.

Since that time, OE and program offices have been conducting a number of outreaches to the regulated community to foster the application of the Safety Culture Policy Statement under direction from the Commission, to ensure they have the necessary support to effectively employ the policy statement.

Feedback thus far is that the regulated community is embracing the Safety Culture Policy Statement and sees the value for working to emulate a positive safety culture. Many of them told us that they were already doing that.

I would like to just mention a little of our involvement in the area of safety internationally, which has really picked up over the last year. OE's efforts in the international community have included responding to several requests for presentations related to the development of our policy statement. They included requests from Russia, UK, and several South African nations and several European nations through the IAEA.

We have worked on a consultancy team to

aid in the development of a technical document for the regulatory oversight of safety culture, and have been successful in including our Safety Culture Policy Statement as an Appendix in that document, which will serve as an example.

We have also recently worked on a consultancy team at IAEA for the purpose of beginning the process of common language for safety culture attributes internationally. This is just the beginning of the process. We expect it will take a number of years to complete.

The group that made this request used the NRC Safety Culture Policy Statement traits as examples, along with some IAEA safety culture characteristics and the language which Japan uses to begin the dialogue. We continue to see a high level of interest in the Safety Culture Policy Statement internationally .

Also OE recently was requested to provide a presentation to the G8 National Security Group by the State Department. The group was very interested in the efforts the NRC has engaged in with respect to the development and publication of the Safety Culture Policy Statement.

The G8 National Safety and Security Group recognizes the importance of safety culture, and will be recommending endorsement of the IAEA Safety Culture Action Plan which was developed in response to Fukushima. Our presentation on the Safety Culture Policy Statement, as well as the information from the private industry, was instrumental in providing this group with the information they needed to make this decision.

Now I want to turn to what I consider the enjoyable part of my presentation, and introduce those here to my right who are going to speak to you today to help explain the research in this area and the results of industry study and our evaluation of it.

Dr. Val Barnes is a Senior Technical Advisor, to my right, for human factors in the Office of Nuclear Regulatory Research, Division of Risk Analysis. She has a Master's degree in organizational psychology and a Ph.D. in social psychology from the University of Washington.

Dr. Barnes has over 30 years of international experience researching human and organizational factors in the NRC. When I was reviewing the user need again recently, I noticed -- I reminded myself that Val was instrumental in helping us craft the Safety Culture Policy traits in the ROP initially, and helped revise them.

Dr. Stephanie Morrow, to her right, is a Human Factors Analyst with the Human Factors and Reliability Branch in the Office of Nuclear Regulatory Research, Division of Risk Analysis. She has a Ph.D. in industrial/organizational psychology from the University of Connecticut, and nine years of experience with survey development and safety culture research. Dr. Morrow has been with the NRC for just a year, and previously worked for the Department of Transportation.

To Stephanie's right is Dr. Ken Koves with the Institute of Nuclear Power Operations.

Dr. Koves has a Ph.D. in industrial/organizational psychology from Georgia Tech and has been working in the area of organizational culture for over 20 years in the telecommunications and nuclear power industries.

Ken first approached me about three years ago when we got started on working on the policy statement as a result of a meeting that the EDO had with INPO in Atlanta to talk about trying to align terminology between the reactor oversight program and NRC's Safety Culture Policy traits that lived in ROP, and he has been very instrumental in helping us get information and move forward in this effort. I want to thank him for that.

So without further ado, i will turn it over to our next presenter.

DR. BARNES: I wanted to start out today by giving a little bit of an update on what has been going on in the research related to safety culture from the standpoint of an organizational psychologist who likes to measure things as opposed to the anthropological approach or a sociological approach that focuses on organizational literature that has been looking at safety culture, has been on attempts to measure it, and evaluate the relationship between safety culture and safety performance.

The basis of that kind of research, of course, comes from an assertion and underlying belief that organizational factors are important to organizational safety performance. Of course, the question has always been which organizational factors, how do you assess them, how strong the relationship is, let alone how could an organization like the NRC possibly incorporate organizational factors into a regulatory oversight regime.

Well, the concept of safety culture hasn't answered all of these questions, but it has given us some new ways of thinking about organizational factors' relationship with safety, and it has helped improve the focus of the research that has been going on.

The initial work that was done on the relationship between "safety culture" and safety performance was actually -- The initial study was actually performed in 1980 by an Israeli researcher before Chernobyl occurred, and the term came into common use in our industry.

This individual came up with some exciting results which, I know, are less than thrilling in the engineering environment, but he was able to find a relationship, a correlation of about .2 between his measures of safety culture, although this area of research is safety climate research within the literature, but we are not going to go there. He found a relationship of about .2, which for psychologists was quite exciting and it was provocative, between his measure of safety culture and the safety performance measures that he had available.

I say it was exciting to us, because a low correlation, even at the .2 level or even lower as long as it is statistically significant and reliable, can actually be useful.

Last time we got together and talked about this, we talked about a few examples, but one of them that came to mind in preparing for the presentation was the relationship between high school grade point average and first year college grades. That correlation runs also at about the .21 to .23 level. The correlation between SAT scores and first year college grades is only slightly higher.

So, you know, colleges use high school grade point and SAT scores in making their admissions decisions, even though those relationships are not in the .9 type of level that we would want to see for doing the engineering work to build a bridge, and so on.

CHAIRMAN BLEY: Although we might mention that the number of prominent colleges have backed away from using the SATs in recent years, because they don't find it predictive at all.

DR. BARNES: Right.

MR. SOLORIO: I definitely want to find out which ones for my sons.

DR. BARNES: So in the 30 years since that initial study was done, there actually has been some progress made in understanding the relationship between safety culture and safety performance, and so

I sam going to talk a little bit about that.

For the folks who like to try to measure these concepts, there is a set of interrelated hypotheses that underlie the research. First of all, there is the hypothesis that the attitudes and beliefs of members of an organization provide insights into the organization's safety culture.

Now the attitudes and beliefs of the members of an organization don't reflect the totality of an organization's safety culture, but it is a way to obtain some insights into what is going on in the culture.

The idea behind this hypothesis, of course, is that culture is a social phenomenon. It arises as a result of the social interactions between members of a group. Members of a group learn the culture through those interactions. They experience it in their day to day work environment.

In response to those social interactions, they develop attitudes and beliefs about what is important in the organization, what the rules of behavior are in the organization, and those attitudes and beliefs can be measured, and that is where we get this plethora of questionnaire survey research studies that are being done, because those surveys are designed to help measure organizations' attitudes and beliefs about what is going on inside their organization's safety culture.

Now a second hypothesis here is that the employees' views, or the organization's members' views about what is important around here, the organization's values, and how we do things around here, the organization's norms, have an influence on their workplace behavior.

In our case where we are talking about safety, it is the assumption that their attitudes and beliefs with respect to the safety culture have an impact on their safety related behaviors.

Then a third underlying assumption or actually testable hypothesis here is that individuals' workplace decisions and actions, their behavior, affects the organization's safety performance.

So that is the framework for the research that has been going on in safety culture over the last 30 years.

Now in the beginning when this research was just getting started, it wasn't very sophisticated. The model essentially was: To measure safety culture, let's measure attitudes and beliefs, and we will see what the correlation is to measures of safety performance that are relevant for the environment that we are working in. So probably the first 10-15 years of the work that went on in this area was focused on, okay, which attitudes and beliefs do we need to measure to be able to identify that relationship to safety performance.

That is where the initial research and the continuing research has come from with respect to identifying the dimensions or traits of safety culture that are important for safety performance.

Then a second piece of this, which is what INPO did in this study during this first decade or so of safety culture research, has been, okay, Zohar back there in 1980 found a relationship between safety culture and safety performance in the industry that he was studying. Can we find a similar kind of relationship in our industry?

So there have been a lot of studies where researchers have gone out and looked at the safety culture/safety performance relationship in manufacturing settings, hospitals, construction settings. They have been asking, you know, does the same relationship hold internationally?

Essentially, the research has continued to show that, yep, the relationship generally holds, although again it is at the relatively weak level, at the .21 maybe to .40 level in those studies.

Now within the last few years -- I would probably say the last five to 10 years -- the relationship -- the research has gotten a little bit more sophisticated as our methods have improved and our theoretical -- you know, and what we have learned about safety culture has expanded.

So that at this point in time the kind of research that is going on is more focused on understanding what are the intervening variables between safety culture and safety performance that have an impact on how strong that relationship is? A number of intervening variables have been identified, and more are on the way.

The research has been looking in particular and borrowing from another area of research in the area of measurement psychology, which is the area of the attitude/behavior relationship. You can see that in the middle of the slide here where we talk about attitudes, norms, personal agency over to the little symbol of behavior.

This area of research is more than 70 years old. It really took off in World War II, and the focus of the research has been on understanding and predicting behavior based on the attitudes and beliefs that people carry around, with the eventual goal of being able to influence behavior.

So the results of this research -- and literally, there have been thousands of studies on the attitude/behavior relationship by now -- has shown again a moderate size relationship between attitudes that you measure with a questionnaire and people's behavior in the end, but they have learned a lot of about what moderates that relationship.

The slide shows a number of the factors that moderate the relationship between attitudes and behavior, including the specificity of how you measure the attitude and the behavior, social influences from important others. The personal characteristics of individuals has an impact on the relationship between their attitudes and particular kinds of behavior, as well as characteristics of the situation in which they might perform the behavior.

So for example, with respect to the specificity of the attitude and the behavior when it comes to measuring and predicting behavior from attitude, if you want to find out, for example, if somebody is going to go for a bike ride at 4:30 this afternoon, you want to ask them, you know, do you intend or are you favorably disposed to going for a bike ride at 4;30 this afternoon, rather than asking them what their attitude is overall toward exercise or bike riding. Right? The relationship is much stronger the more specific you are about the behavior -- the attitude toward the behavior that you want to measure.

In fact, depending on how well attitudes, norms, and these other components of the attitude/behavior relationship are measured, we see correlations as high as .7 or .8, which is very high in the social sciences. This matters for certain research topics like voting behavior, of course, purchasing decisions, marketing research. You want to be able to influence and predict people's behavior. That has been the impetus for a lot of this research.

So by doing a good job at measuring these components, you can get a stronger relationship. Of course, attitudes: We have been talking about someone's degree of favorability or unfavorability toward a particular behavior.

You would want to measure the norms that the person is operating under in terms of what they believe that important others do with respect to that behavior, and how important they perceive that performing that behavior is to important others; and then personal agency refers to whether the person believes that they actually can perform the behavior and the extent to which they have the wherewithal or ability to overcome any barriers that might exist between their attitude, their favorable attitude toward performing the behavior, and actually being able to do it.

So attitudes, norms, and personal agency together determine someone's intent to perform a behavior and how strong their intent is. But even if someone has an extremely strong motivation to perform a behavior and intend to do it, real life can interfere.

For example, you know, I might have a strong desire and all the motivation in the world to go bike riding at 4:30 this afternoon, but if I get out there and I have a flat tire on my bike, I am either going to give up for that day or be delayed in doing it.

Similarly, you might have all the motivation and intent in the world to go vote, but you don't have a ride or the bus doesn't run, and you miss being able to get to the polling place.

So the relationship then between attitudes and behavior is moderated by these kinds of factors.

Now what does the attitude/behavior relationship have to do with safety culture? Well, it is useful for understanding the influences on individuals' workplace safety behaviors.

So the question arises, where do people's attitudes about safety, the norms about safety, their sense of personal agency about safety behavior come from? The concept of safety culture helps provide some answers to those questions.

MEMBER RYAN: I've got a quick question. I am thinking in my own experience of a job that doesn't really take a lot of sophisticated protective clothing, a mechanical job of some kind. Now I get that same job where I in a full PC. I have seen two very different reactions to the work environment or the success of the work environment based on those extra requirements: This is uncomfortable. It is hot, all those kinds of things.

DR. BARNES: Right.

MEMBER RYAN: So where does that fit in on your --

DR. BARNES: That would be the situational barriers.

MEMBER RYAN: The situational barriers?

DR. BARNES: Yes. You know, and

motivation, of course, as well.

MEMBER RYAN: Some people feel claustrophobic in protective clothing or a mask, and others don't. So you tend to pick people who don't mind that environment.

DR. BARNES: Right.

MEMBER RYAN: Even though they might not have the same skill level in the activity itself, the pump replacement or the mechanical or electrical work, but they are kind of getting judged on their overall ability to perform it, not just the one characteristic. Does that make sense?

DR. BARNES: Yes. Go ahead.

DR. KOVES: Plus, I just think another variable that comes in there very often that I have seen is their kind of risk perception. Is it a really risky thing, either for me personally or for the plant overall? And if it is high risk they are more willing to do all that stuff, and if it is low risk, then they are less.

MEMBER RYAN: So that is the perception of value.

DR. KOVES: Right. Just another variable that comes in and hits that behavior.

MEMBER RYAN: Thank you.

DR. BARNES: So you want really strong norms, and you also want strong consequences for failure to comply, to try to encourage compliance behavior, despite the discomfort and claustrophobia and dislike. That is where leaders, supervisors, and peers come in.

MEMBER RYAN: And that is all -- Risk level awareness is what I call that, because you really have to make people aware of the risk level that they are going into; and even though it looks like you have a job which has a lower risk level, it is really not, for these reasons and so forth.

DR. BARNES: Right.

DR. KOVES: Or this one really has more risk associated with it than you think it does.

MEMBER RYAN: Yes, and here is why, and here is what why we need you to do it this way, and so on.

DR. KOVES: Right. Exactly.

MEMBER RYAN: And try and get buy-in for whatever the task is. Thank you.

DR. BARNES: Yes. So with respect to how safety culture has an impact on the behaviors that we care about, the research has continued and has been showing, not surprisingly, that people in different roles in the organizations have differential kinds of impacts on these intervening factors between safety culture and actual behavior.

Leaders, of course, set the stage in the organization for the organization's attitude toward safety as a whole. They establish the policies, of course. They play a very important indirect role in terms of allocating resources. It is leaders that decide how much money we are going to spend on training, for example, how much money we are going to spend on benchmarking, how much money we are going to spend on safety equipment.

Then leaders also, to the extent that they are visible to the workforce anyway, model behaviors. In our industry, one that we frequently hear about is -- They model the behaviors. The example that we see in our industry of senior leadership often modeling how important safety is in their organization would be extending an outage to make a repair to a system that they could delay, but they want to make the point that safety is important. So they add a couple of days or maybe even up to a week to an outage to get that repair done. So that would be an example of modeling the importance of safety in the organization.

Leaders, the research is showing, have

actually, like I said, more of an indirect relationship, even though their role is critical; whereas, in the day to day workplace it is supervisors that have more direct and strong impact on worker behavior because, of course, they are the individuals that administer rewards and sanctions for following -- procedure compliance or violating them. They are the individuals that actually interpret the leaders' policies and bring it down to what happens in the day to day work group.

Now peers also have a role to play in terms of reinforcing norms. Peers also control a lot of information that you need in terms of being able to do your job in many different types of jobs. Our very own Dr. Morrow here did a study that was published not too long ago demonstrating that co-worker -- that perceptions of co-workers's views about the importance of safety are probably less important than we might have thought about.

So we are learning more about some of these intervening variables and what is important about safety culture.

CHAIRMAN BLEY: You just mentioned that as an aside, but was that done here or was that --

DR. MORROW: No, before I came here, and

it was basically that kind of the individual's perception of the tension between productivity and safety and the management's commitment to safety, that perceptions of that were more important than just the co-workers' and peers' perceptions.

MEMBER RYAN: So at the end of the day, that little equation proves who is the real leader of defining safety along that peers group, working group, and supervisors group.

DR. MORROW: Of those three things, what we expected was that it would be the managers, but actually, it was kind of more general than that, and that it was that tension between productivity and safety which is partially informed by the managers, but it is broader than just your direct supervisor or your supervisor's supervisor. It is about how the entire organization is running and kind of the organizational commitment to safety.

CHAIRMAN BLEY: I would be interested in seeing your paper.

DR. MORROW: Sure. Most definitely.

MEMBER RYAN: My own experience is -- and maybe Dennis might be the same -- is that typically a younger worker or a newer worker to the organization will tend to look at a more senior worker or somebody who has been there a long time as to what is the right answer. They have seen more variation among all those variables, and they could kind of say, well, here is where the -- not saying it exactly this way, but tell you what the mean behavior is to get through.

MEMBER SCHULTZ: But I would also suggest they would be influenced by their peers, too. So I would be interested in taking a look at your paper. I want to know if it varies by age of worker, as an example.

DR. MORROW: That is a good question to look at..

MR. WIDMAYER: Was that in the nuclear power industry, Stephanie, or is it a different environment?

DR. MORROW: No, it was in transportation, railroads.

MEMBER SCHULTZ: So you have a commitment. In the context of that type of study, are you considering workplace safety or are you considering the safety of the enterprise? I am trying to draw a parallel between nuclear safety and worker safety in the nuclear industry versus what you may have studied in transportation.

DR. MORROW: It was more worker safety,

you are right, in that more kind of the behaviors like wearing personal protective equipment, following kind of the safety rules, which are mostly for the benefit of the worker. There is not as much -- Particularly, it was mostly like maintenance organizations in a railroad environment.

So they are moving the heavy pieces of equipment within the railyard, not so much out on track going across country where you have that risk of hitting someone or something like that.

MEMBER STETKAR: You didn't look at correlations between those attitudes and axles falling off rolling stock, though?

DR. MORROW: No. No, we didn't have that data for this study.

DR. BARNES: This is actually great conversation to talk about at this point, because that was the next step, was to talk about relationship between behavior and organizational safety performance, which is another issue complicating the relationship between safety culture and safety performance. Those of you that work in the area of HRA know how complicated this can be.

CHAIRMAN BLEY: Yes. One thing I would just mention here, because everything you have talked about and everything I read in the papers is searching for these, I will call them, relationships.

When you turn the problem around, though, and you see bad events happening, if not 100 percent, I bet it is something like 80 percent or more of the cases, when you really dig in, you find there are some of the factors this project has defined as factors affecting safety culture that are negative and are partially responsible for those events.

So there you see a fairly strong relationship between negative factors being coupled with bad events, and I suspect it is a much stronger relationship than some of the ones we are seeing by trying to relate overall score in a safety culture sense against a particular facility's overall operations where you have so many other confounding factors going on and affecting things.

DR. BARNES: Absolutely.

CHAIRMAN BLEY: But you don't -- We are not pushing on that side of it very much in this work.

DR. BARNES: But we are, and that is what I want to talk a lot about. I want to talk for a few minutes here about what you just said, which is that the relationship between individual behavior or even patterns of behavior that evolve in an organization and organizational level safety consequences where you have to look at or are looking at counts of very rare events, airplane crashes, events in nuclear power plants, even OSHA reportable accidents. It is a complicated issue.

How do we measure safety performance? And it is not a one-to-one relationship at all between individual behavior and organizational level outcomes. You will recognize this from this concept of the Swiss cheese model from our friend, Jim Reason, who of course, has talked to us about the extent to which patterns of safe or unsafe acts or even momentary slips or mistakes will impact safety performance, and that depends on the kind of tasks that the person is doing as well as the extent to which we have got organizational barriers or engineering barriers and other methods to prevent the act or mitigate the consequences of unsafe behavior.

To talk a little bit about the design of jobs or the tasks that people perform and the relationship of behavior in those environments to safety performance, I wanted to talk about an example of nurses' behavior with respect to needle sticks or transmittal of infections in hospitals.

Nurses' unsafe acts can have immediate and important effects for not only their own occupational safety but also for patient safety overall. So this is an example of a job where individual behavior and important safety outcomes are closely linked.

You probably are all familiar with the research and theory that is talked about, tightly coupled behavior and safety performance. So in those kinds of environments where an individual behavior can have immediate and direct impacts on some kind of important safety outcome where they are tightly coupled, we are going to see a stronger relationship between behavior and safety performance than we will in other kinds of task environments where there is not that close relationship and direct relationship in time.

For example, a finance organization -- you know, a finance work group within an organization, if they are sloppy or do a bad job, may eventually have some impact on the overall safety and functioning of the organization, but a pattern of unsafe or undesirable acts within a finance group is going to have a lot lower relationship with eventual safety performance.

So a big question in this research, as you guys have brought up, is what are the right measures for assessing safety performance, if you are trying to look at the relationship between safety culture and safety performance?

Across the research there have been different categories of measures as well as different levels of analysis. As we have been mentioning, a lot of the studies have been done on occupational safety measures, the OSHA accident and injury rates, certainly patient safety.

There has been a move in the last five years or so to focus more on measuring safety performance by looking at near-misses or what is being called micro-accidents.

Then, of course, there is looking at human error rates, if you can identify them, patterns of behavior with unsafe acts, and then self-reports of injuries, errors, near-misses, because self-reports often show higher rates of these kinds of phenomena than your OSHA reportables or report of your more objective measures.

So the research has been done using these kinds of safety performance measures, as well as looking at individual behavior, not only self-reports about I always wear my earplugs or I always follow the procedure, but also work has been done sending researchers or others into the field with the workers and having them make observations about the frequency with which workers engage in either safe or unsafe behavior, however that is measured.

Work has been done looking at developing safety performance measures for work groups. Then, of course, we have got safety performance measures like we have mostly in the NRC kinds of performance indicators and measures that look at whole organizations, whole sites. Other work here has been looking at whole hospitals or entire companies.

In general, the strength of the relationship that we see in the research literature between safety culture and safety performance depends on what you are looking at and at what level of analysis.

If you are looking at individuals, the correlations between individual attitudes and beliefs about safety culture and safety behavior are medium to large in social science terms. In terms of organizational level or in terms of safety performance outcome measures, it depends on, like I said, what kind of task the individual is performing, how strong that relationship is.

If you look at the behaviors of work

groups and you look at their safety behavior as your outcome measure, again we get medium to large correlations; and if you look at aggregate measures of safety performance outcomes -- for example, accident/injury rates for work groups -- you also get medium to large correlations.

When you start looking at organizations as a whole, the relationship between safety culture at the entire organizational level is fairly small. We are down at that .2 level again, and in terms of individual behaviors and in terms of overall organizational level safety performance outcomes, you get small to medium results. This is across industries, across measures.

So it is not -- So it requires some care in how you approach the measures that you are working with and some thought in evaluating the strength of the correlations that you see in any of the research literature that has been being performed up until now.

So all of the foregoing was essentially just to talk about how the relationship between safety culture and safety performance is more complex than the bubble and triangle model would initially suggest, and that there is a lot of intervening factors, of course, that moderate the strength of that relationship, in addition to which it is very important to understand how both safety culture and safety performance are being measured when you are evaluating research studies in these areas.

So all of the foregoing is academically interesting, but what actually matters is: Is this research turning out to be useful? In fact, a lot of organizations are using safety culture type survey research for several different purposes.

One, of course, is to do internal and external benchmarking. More importantly, a lot of organizations are using these safety culture surveys to identify areas in need of improvement. They are using the information that they are collecting plus, of course, additional information to help design interventions, organizational interventions, to help improve safety performance. Then they are using these surveys at mid-points and down the road as part of their assessment about whether their interventions have been successful or not.

Here are some examples of some of the outcomes of the safety interventions that have been undertaken to improve safety performance across a variety of industries. These are pretty exciting results. You know, 30 percent decrease in unsafe acts is a pretty valuable outcome.

The 70 percent decrease in the ventilator-associated pneumonia rates across Michigan hospitals is especially exciting. This study just was published in February, and some of the estimates said, you know, it is valuable not to put patients through pneumonia when they are exposed to having to be on a ventilator, in the first place, but they are also estimating that they are saving about 1,000 lives a year as a result of the safety culture interventions that they have implemented.

CHAIRMAN BLEY: In these examples you have put up here, was the definition of safety culture within these facilities and industries essentially the same as what came out of our work? It is not real clear what they were trying to change. What did they change to get these kind of results? Can you talk about that at all?

MEMBER SCHULTZ: Let me add a question. Was the intervention due to the culture of the organization or are there basically engineering barriers that were applied to have caused this effect?

DR. BARNES: Both. Both. The culture survey identified -- helped them identify areas where there were cultural type problems going on as well as areas that could be improved by the application of, for example, engineering barriers, the introduction of checklists.

In fact, you know, I would be very surprised in our industry if we were able -- Ken can speak to this, will speak to it, I think, better. But the 70 percent decrease in the pneumonia rates -- a significant portion of that had to do with improving communications among physicians, nurses, and the team members in changing the flavor or tone of those communications, the cockpit resource management type change of attitude where the physician decision makers listened more to the ICU nurses. But they also got big bang for the buck by implementing checklists.

CHAIRMAN BLEY: People in that field are getting famous now writing books about checklists.

DR. BARNES: Yes. I mean, we have used checklists and procedures forever in our industry. So, you know, it is hard to imagine a behavior that we would want to change where introducing a checklist would have a big impact. Communications? Probably we could have some significant impacts on safe behaviors and overall safety performance by interventions directed at that, but the problems -- The weaknesses were identified initially using the Safety Culture

Survey as a screening tool to help identify weaknesses.

Then, of course, they had to go in and do interviews and evaluate the work environment and get a more detailed root cause level understanding of what was going on, and then use that information to design and implement the interventions. They measured progress as they went along, with additional administrations of the survey.

As the safety culture scores on the survey improved at the work unit level, they started seeing these improvements in pneumonia rates, you know, fewer infections in their ICUs and so on and so forth. So the research is showing that this can provide some useful results.

MR. WIDMAYER: I have another question for you. Because these can be just as enlightening, is there a bullet that you could put here where there was no increase or decrease in what they were trying to measure?

DR. BARNES: Tons, yes.

MR. WIDMAYER: It gives you a chance to look at the trait and say, okay, maybe this trait isn't as important as we had initially thought it was, and we don't need to measure it anymore, or something.

DR. KOVES: I think you can get to that point in certain ways, but if you have incidents where they say, hey, we need to work on this, and they don't succeed in changing anything, the place I would look first is, okay, how did you try and change it.

I think that is where those implementations fail most of the time, at least in my experience, is in the implementation and how they try to do it. However, I think we do have some data you might see later where one of these traits in certain areas is predictive, and most of the time it is not really very predictive or very useful, actually, so it is so consistent across the industry and across the station. There's no variant cells, not very good.

CHAIRMAN BLEY: My only problem with this, as you said -- and I think all four of these cases -- we have been doing the kinds of things that helped them for a long time. So we are looking at some other kinds of perhaps marginal gains. I don't know if there are any big gains out there to be had, but seeing these doesn't help us know we will do well.

DR. BARNES: Well, they are not in nuclear settings. So until we have got research that demonstrates this or looks at this in nuclear settings, we can't say for sure that these kinds of results generalize or how this kind of research and organizational intervention approach would benefit or not.

Now that is the kind of work that Ken does at INPO, but he is not publishing it. So we are not looking at --

DR. KOVES: Give me some time.

MR. WIDMAYER: For a small fee.

MEMBER STETKAR: To follow up a little bit on what Dennis, I think, was talking about, have there been any studies done in other industries that profess to have a high level of safety culture where errors can cause fairly severe consequences? I am thinking, in particular, aerospace, airline industries. Has anybody done any work in those areas and published any results like this?

DR. BARNES: Yes. They have been doing similar kinds of work, but I can't say that I can point out -- I can't bring to mind a study. Can you?

DR. MORROW: Well, one -- you know, I will speak to the interventions -- is coming from working with the railroads. One study that they did was --

MEMBER STETKAR: I didn't say railroads. In particular, I didn't say railroads, because I said that profess to have a very high level of safety culture, aerospace and airline, in particular. I don't want to pull us to things that you know about. I want to find out how far you have looked into industries that are often compared with the nuclear power industry as having a very high attention to safety where errors can indeed result in very, very significant consequences.

Railroad industry, regardless of what you look at, running your train off the track isn't very severe, usually.

DR. BARNES: Two answers. One is I haven't seen published studies of, you know, flying related. They are doing work on this in maintenance.

MEMBER STETKAR: Well, maintenance is relevant.

DR. BARNES: Yes, aviation maintenance. So this same kind of work is being done there, the measurement of safety climate, and then linking it to different kinds of indices of maintenance effectiveness, not necessarily linked to safety performance, because in aviation maintenance effectiveness --

MEMBER STETKAR: Sorry. Occasionally, flaps don't work, and very often they trace things to maintenance performance on aircraft.

DR. BARNES: But are like us from the standpoint that, you know, their procedures and testing protocols are set up to, hopefully, detect and mitigate a problem --

MEMBER STETKAR: That is exactly why I asked the question.

DR. BARNES: -- before it results in a crash.

MEMBER STETKAR: Because a slide like this tells me that perhaps there may be benefits in other industries that perhaps didn't really have a very strong safety culture or a very high attentiveness to errors that could result in severe consequences, regardless of what you might say about the health care industry.

There are a couple that traditionally have for the last 34 years anyway, certainly. But if you searched and people haven't done that type of work, then there is not much --

DR. BARNES: Yes. It is not published. I mean, I would be shocked if it -- I mean, of this kind of empirical research. Of course, there is accident investigations that always point back to safety culture factors, organizational factors as contributing.

MR. SOLORIO: I want to ask a question of Dennis. I don't mean to get myself in more trouble here, but I think what Val was suggesting, for instance, with the last bullet here, the 70 percent decrease, is: In the study, they focused on -- One of the many things they did was focus on communicating better between individuals, and that contributed to reducing the pneumonia rates.

In the Safety Culture Policy Statement, we have a trait, communication. So I thought what Val was trying to do was draw an analogy that in another environment they focused on something like what we are focusing on in the Safety Culture Policy Statement, and it reduced rates.

So while, you are right, we are perhaps much more advanced in our way of being safe than other industries, we are still focusing on something that is really important, just as in other places, to reduce accident rates.

CHAIRMAN BLEY: I was thinking -- I have to go back and look at the Policy Statement again, but I was thinking that was a more general communication. We have, in fact, been beefing on communication within the control room for many years, and adding requirements and double checks and now triple checks on that kind of communication.

These guys weren't doing that. There was the king, and everybody did what he said and was afraid to say anything. So there was a big difference in the way they operated.

DR. BARNES: One of the -- The one study that I am aware of that has compared nuclear power related safety culture to other industries was in the UK, had nuclear power off-scale high in terms of their safety culture scores compared to some 400 other industries that they were looking at.

CHAIRMAN BLEY: Did they do that or was it --

DR. BARNES: University of Aberdeen for the government, and that is anecdotal gossip, not -- I can't point to a publication that documents that.

So there is a ceiling effect in nuclear power which, thank heavens, we are glad exists, but it also argues against there being, like you guys are saying, a lot of room for safety culture to have a big impact on safety performance.

DR. KOVES: The other side is we still see incidents where all of these things crop up.

DR. BARNES: That's right.

DR. KOVES: When it is done right, no

matter what we try to do, what we have tried to do so far to deal with that.

DR. BARNES: Yes. But for these kinds of reasons, when INPO stepped forward and said that they were going to volunteer to do this construct validation study, I was standing back and going, okay, well, this is good to extend the research that has been going on in the other industries to see what happens in nuclear power, but we are going to be looking at the organizational level of safety performance, organizational levels of measures of safety culture, and going into this in advance, I am going, this will be fun to do, but I don't know -- you know, I am not very confident that the outcomes will be -- that we are going to find anything, actually, that there will be a sufficient variability that the correlations will show much of anything.

So, in fact, when the study was done, when we started getting the results in, I was kind of taken aback and a bit shocked, because compared to a lot of the other work that has been going on, we apparently do have sufficient variability between sites in the nuclear power industry that we got -- as Ken and Stephanie will review next, we actually got correlations that are generally stronger than what has been coming out of similar kinds of research in other industries.

CHAIRMAN BLEY: Go ahead.

DR. BARNES: That was it.

DR. KOVES: Oh, am I next?

DR. BARNES: You are next, unless you wanted to discuss anymore of that.

CHAIRMAN BLEY: We will back. There are pieces of it.

DR. KOVES: While they are switching the slides and presentation over, I just kind of wanted to jump back to something at the very beginning that Dave talked about, and he had thanked me for some of the help early on in some of the policy statement stuff.

I just wanted to commend you, Dave, for kind of trying something different in terms of some of the activities and all and just -- I think Dave just did a really good job in helping move the policy statement forward and making it very successful.

So I just wanted to say that. I told him that in private, but this is my only opportunity to say it in public.

CHAIRMAN BLEY: Go ahead. Then I will ask one.

MEMBER STETKAR: No, I was going to ask

you need to close the meeting yet.

CHAIRMAN BLEY: No. I figure he will tell me when he wants to close the meeting.

DR. KOVES: That is correct. Actually, the plan is to go through all of the presentation and then I am going to hand out a couple of things at the end, and that is -- we will go ahead and close the meeting at that point.

CHAIRMAN BLEY: Okay. All of the slides are okay?

DR. KOVES: Yes, all the slides are okay. It is just my handout and the handouts that I brought, and I will hand those out at the end, and we will kind of discuss them, and these are overall results and more detailed results.

CHAIRMAN BLEY: I think the question to put on the table -- and I can understand you might tell me you can't talk about it, but if you can we would be interested.

We have had some meetings in the past with other folks from INPO who are associated with inspections and drills and that sort of thing. What I am interested in is are you on a path such that your results get mixed with what those kind of folks are doing so that, from the industry side, we are seeing this research end up in a place where it has practical impact through other programs that are going on?

MEMBER SCHULTZ: Where it connects to other evidence that INPO collects also from licensees. Is that what you are looking for?

CHAIRMAN BLEY: Well, partially, but also the fact INPO goes out and inspects facilities and oversees or at least observes drills, if they don't run themselves.

DR. KOVES: We don't run them.

CHAIRMAN BLEY: Okay. So you observe them, but you have some impact on what people do there. What I am wondering is there anywhere you can talk about how this work feeds into -- mixes with other sources of information and feeds into other activities would be of great interest.

DR. KOVES: Well, I am pretty sure that I can go ahead and talk about that definitely during the private session.

CHAIRMAN BLEY: Okay, that would be great.

MR. WIDMAYER: Excuse me. As far as I can tell, I think we are okay. Everybody here is an NRC employee.

CHAIRMAN BLEY: Is that right? Okay.

MR. WIDMAYER: Yes. So as far as -- It is okay if you want to go ahead.

CHAIRMAN BLEY: Well, we would have to do it officially.

MR. WIDMAYER: Yes, we do.

CHAIRMAN BLEY: But we can wait until later.

DR. KOVES: Yes. Let's go ahead and do it officially, but also I don't completely -- at this point, completely understand the question. So why don't we go ahead and go through this.

CHAIRMAN BLEY: That's fine.

DR. KOVES: Then when we get into that session, go ahead and ask the questions, and maybe it will be a little bit more detailed, and I will do the best I can. Remember that I have never been an evaluator.

CHAIRMAN BLEY: I can't remember, but I can be informed.

DR. KOVES: I have never been an evaluator. So speaking from that perspective, I am modestly familiar with it, but I have never been one.

CHAIRMAN BLEY: Okay. But you interact with those folks?

DR. KOVES: Yes, quite regularly.

CHAIRMAN BLEY: That is what I was getting to, and we will talk more about it later.

DR. KOVES: Yes. Okay, great.

Me? We have already talked about me a little bit. yes, I am a Ramblin' Wreck.

What is the purpose? Why did we do this? First of all, to validate the traits in the Safety Culture Policy Statement, as we have talked about, but also then to determine a relationship between INPO's safety culture survey and the concurrent measures of safety performance. Is there a relationship there? Then ideally, because this ended up getting -- going over time, we were able to look at the relationship between safety culture survey results and then also the same measures of safety performance one year later.

CHAIRMAN BLEY: Now you had a safety culture program before this work that came up with the policy statement. That's true?

DR. KOVES: I will -- Yes. I am not sure exactly how to answer that question because of the word program. However, INPO has had a lot of emphasis on safety culture on the concept without that language prior to Davis-Besse, and then on that concept after Davis-Besse ever since then.

So the Principles for a Strong INPO Safety

Culture came out in 2004, and I am not sure how many years, but for a number of years now, INPO has had an organization effectiveness part of their evaluation, and then a subset of that is safety culture. So they have been looking specifically at safety culture for many years.

CHAIRMAN BLEY: I'm sorry, just two more questions on development of the Policy Statement?

DR. KOVES: Yes.

CHAIRMAN BLEY: And by the time that evolved and came out, it matched reasonably well with what was already there, or have you had to change things?

DR. KOVES: There was a lot of alignment. This whole language thing gets a little murky, because of -- The example that I use is that it is like a cherry pie. I mean, how many different ways can you divide a cherry pie, you know?

So the NRC, the INPO, IAEA, JANTI, VTT -- you know, they all kind of divide a little differently, but really, we all pretty much agree that it is a cherry pie, and we agree what the cherries are, to a large degree. So anyway, that has -- We are working on that, and also we are in the process of getting an aligned language with the NRC at the more detailed level. We made some great progress in December, and it looks like in July we are going to be kind of pushing that on and probably coming to -- my expectation is, coming to a conclusion probably in July.

CHAIRMAN BLEY: Of this year?

DR. KOVES: With NRR. So the Commission's original vision was to have a high level description of safety culture that applied to all the different stakeholders, but then each stakeholder group or each -- what is it, department? I don't know what they call them in the NRC -- department would work with their stakeholder group to then define that at a lower and more detailed level for that group; because they admit that a power plant is a little different than maybe Joe Radiographer who has a source in the back of his van, and that has been moving very well.

Okay. Anyway, that was the purpose. Just for your information, what were some of the roles? The INPO did survey strategy planning and industry indicators. A vendor administered the survey. NEI paid for the vendor, and then the NRC reviewed the work, and then also analyzed the NRC data. There should be another little bullet point underneath there, because they have certain proprietary data that we don't have access to, and vice versa.

So it was a collaborative effort in that process. Let me move on. I am giving you this slide as just a bit of background, because it is going to come into play in the next slide. What this is, is these are some of the current existing models of safety culture.

We have the IAEA on the left, INPO principles and the ROP components on the right. The point here is to show that, even though there are differences and at a high level they may look -- sometimes they look very different at a high level. Once again, when you get down to a lower level, they are really very similar. So this is just kind of a mapping, a very quick mapping, across the different models to show the similarity.

Now where does that come into play? In developing the survey or the questions, what we did was we started with the Utility Service Alliance Safety Culture Survey, and that was very much -- It was built off of the principles document, and the person who put that together, his goal was to align very strongly with the wording in that document.

Unfortunately, that document was never intended to be a survey. It was intended to be a guidance document, and that will come back into play in just a minute. But we started there, first of all, because I pointed out, particularly in nuclear power space the models are really similar when you get at a certain level.

Also, there are really only two nuclear models with any type of existing questions, first of all, that I was aware of anyway. First of all was the Alliance Survey, and then secondly is IAEA in their SCART, Safety Culture Assessment Review Team, process which is a two-week long analysis of safety culture -- they have a lot of questions. They are not survey questions, but they are more like interview questions.

So those were really the only two sources of anything in existence in nuclear space that I was aware of, and we also were under somewhat of a time crunch. The OA was working forward on the policy statement, and they were going to make this thing happen, and we needed to develop the survey. We needed to administer the survey.

We needed to do the analysis in kind of a limited time frame. So we were fishing on all of this to get it done. So, therefore, I just started with the USA survey, and then went from there.

What did we do with it? Here you see some of the editing. Took the original survey, and then what I did was I did a bunch of editing of the original 73 items. You kind of see a little bit of the feel for what that initial editing went like.

As I said, it was built upon the principles. The principles were never intended to be a survey, and so there was a substantial amount of editing to just get the questions that were into a more appropriate survey format; after that, then set it aside for a little bit, and then came back to it with fresh eyes, did more editing on it.

Then also as a part of that, did a comparison between what was in the survey, kind of mapped it over to the traits, and looking for, okay, were there any traits that I didn't feel were covered really adequately by the existing questions. There were a couple of traits that I -- you know, I only had maybe like three or four questions. So I ended up adding another 10 items to the survey.

After that, I sent it off to NRC Research, and they then did editing, and they can talk more about it, but I know that there were a number of people in their group and other groups that looked at it, plus also what they did was they did the same process of comparing it to the traits, but they also

-- my understanding is they compared it to the IAEA framework. They also compared it to the ROP framework, and then also compared it to the research. As you see, there is a lot of the research that they are familiar with, and ended up adding another 32 items to it.

MEMBER STETKAR: You mentioned IAEA a couple of times, and they do have some type of evaluation process. Do you know how frequently they have actually one those evaluations?

I am not familiar with that. I do some work with the IAEA, but not in that area.

DR. KOVES: Yes. I actually know quite a bit about it.

MEMBER STETKAR: Oh, good.

DR. KOVES: The very first -- and it is the SCART. Their long process is the SCART, and their very first session, the very first time they did it was in South Africa a few years ago, but that was not at a nuclear plant. That was at a vendor.

Then their first time at a nuclear plant was Santa Maria de Garona like about a year or so after that. Their third administration of it was at Laguna Verde, and I was actually part of that team. So that is kind of why I know a little bit about it.

Well, I don't know how much -- I could babble about this for a while, but I am not sure if I should or not.

MEMBER STETKAR: No, no. What I was more curious about was the number of times that they have actually performed these assessments. For example, I know their IPSART process is fairly mature, and it took a while for them to work the kinks out of that.

They have some other review activities that take a while to review, to work the kinks out. So I am trying to understand at least how far they are in the learning process.

DR. KOVES: They have done three administrations, and the short story is the reports that came back from that, the people inside IAEA felt that they were just like OSART's. So they have now gone down a path of trying to do what -- similar to what INPO does, and that is assess safety culture as a part of the OSART.

They have done two of those, one in Brazil and the other one in South Africa.

MEMBER STETKAR: So it is now becoming part of the OSART -- a subset of the OSART?

DR. KOVES: My understanding is that those didn't work out quite as well as they have hoped, and so, therefore, my understanding is the debate as to what they are going to do is not settled.

MEMBER STETKAR: I will let you get back on track, and thanks.

MEMBER SCHULTZ: And also, in summary, their activity associated with survey and results has been developmental. They have done it three times, but not developed something and applied it the same three times.

DR. KOVES: So far with the SCART, they have not used a survey at all. If the station has done something related, they will use that as part of their input, but they have not -- It is not part of the -- So far, it is not part of the SCART process to use a survey. However, going forward, they were talking about doing that, and as of right now -- I mean, I am talking with them about actually developing a global safety culture -- doing global administration of a safety culture survey to try and get global norms.

MEMBER SCHULTZ: So they were using an interviewing technique?

DR. KOVES: Yes. It is all interviewing. Interviewing and observation is what they were doing.

NRC then added these additional items.

What happened then is then they sent it back, and we had to do some more, and then sent it back; and then they had to do some more, and then finally we ended up with a survey of 110 questions, which was 51 percent more items than what we started out with.

CHAIRMAN BLEY: Let me just ask a picky question. You have 110 items, but when I go to some of your tables where you count things up, how many items were put into different labels, I turn up with like 106 out of 110. Were there some that didn't fit or is there a miscounting?

DR. KOVES: You will see where some of those went.

CHAIRMAN BLEY: During your talk? Okay.

DR. KOVES: Later on.

CHAIRMAN BLEY: So I am not just a bad counter. They did go somewhere. Okay.

DR. KOVES: Well, yes. Oh, it's four difference. Perfect. Perfect, thank you. Thank you for pointing that out. I appreciate it. It's amazing when a plan comes together, let me tell you.

All right. So that was the survey development. In terms of lousy administration, what do we do? It was an online survey. It was administered by the vendor, as I mentioned. They had a seven-point scale, from strongly disagree to strongly agree, also had a "don't know" option, which comes into play, and then also randomly selected a sample of 100 personnel from each site.

MR. WIDMAYER: But the choice of this scale and "don't know" option and everything, that is important. I mean there is a reason that you chose the --

DR. KOVES: Yes. If you look in the literature -- I mean, from a psychometric perspective the sweet spot for scales is five to seven points, and over years using both Sprint, I just finally settled on the seven-point as, really, I think, a little preferable.

Basically, it gives you the option for a little more variance, and it is -- Variance is king in this analysis. But other than that, it is pretty typical Likert scale.

DR. BARNES: And this Likert scale of strongly agree, 90 percent of the studies that I was talking about earlier use that same format -- item format, questionnaire item format. It made these results comparable to the literature from other domains.

DR. KOVES: I thought you guys might ding me on the 100 people from each site, but if not, I will move on.

MEMBER STETKAR: I was going to ask you about that. The 100 people from each site randomly selected, were they clerical people up through licensed operators and supervisors?

DR. KOVES: It was all permanent employees at the site, including what I call permanent contractors, you know, those contractors who had been there for years. It was a dataset of that, and then just used Excel to randomly pull out 100 people for each one of them.

MEMBER SCHULTZ: So you did all the work groups at the site?

DR. KOVES: Yes, and what they are going to talk about later is they kind of analyzed it between the results and your typical -- I guess, typical pattern that you have in a plant, and it lined up pretty well. So it seemed very representative.

DR. BARNES: I strongly encouraged INPO to include -- not just limit it to those who have safety related responsibilities, but to include everyone from the site, because our measures of performance were site-wide as well, and research literature suggests that, even though people in finance have a less direct relationship, they share in the culture, and so wanted everybody in the organization who has been there long enough to experience the culture and participate in it have an opportunity to be selected for the study.

MEMBER SCHULTZ: I guess, in terms of the selecting 100 as the metric for the survey, it would get into what was discussed earlier, that 100 might be adequate to look at the whole organization, but the importance of being able to go down to the work group level and draw meaningful conclusions from that may require a much larger sample size.

DR. BARNES: Absolutely.

DR. KOVES: Right.

DR. BARNES: If you were going to use this survey to see which work groups might be having problems compared to others, you would need a much larger sample, and then design interventions to solve the problem. But that wasn't what we were doing here. We were looking at the question, is there a relationship between safety culture the way we have defined it out of the workshop and safety performance at the very broad level.

So we needed -- INPO needed enough people in the sample to give us a reasonable estimate of perceptions of safety culture at the site as a whole.

DR. KOVES: Station-wide.

DR. BARNES: Station-wide, and that is where our safety performance measures were.

You know, we didn't have access to individual work group measures. So this design was appropriate for the purposes of this study, but wouldn't at all be appropriate for solving -- identifying and solving an organizational problem.

DR. KOVES: Plus, we were asking the stations to take this survey out of t he goodness of their hearts, and so we wanted to reduce the burden on the stations. So this was a way of being able to get at that station level data, but also reducing the burden on the stations.

MEMBER SCHULTZ: Did any station volunteer to take a larger sample or did they all go along on the survey?

DR. KOVES: None of them asked for a larger sample, and we could have done that if they wanted, but no one really asked for it.

MEMBER SCHULTZ: It is about a two to three percent sample of the -- I don't know how many people.

DR. KOVES: Well, typically, it depends on the number of units, but I would say for a one-unit site --

MEMBER SCHULTZ: Okay, this is site.

DR. KOVES: -- typically, about 500 people there. So this would have been almost 20 -- asking 20 percent. Now we didn't get 100 percent response. Is that a fair statement?

DR. MORROW: Yes, that is about right.

MEMBER SCHULTZ: But it included all work groups, security included?

DR. KOVES: Yes, included security. One of the questions was -- that had been asked was, well, is there going to be a difference between what we will call the core functions like ops and maintenance and that type of thing, and those other groups. So we asked a demographics question very early on. We did a split-out between the two of them, and there wasn't really much difference. So we didn't pay any attention to it after that and just used the entire -- all the groups together.

So what was our response? We had 63 of the reactor sites participating, 97 percent of the industry. One station asked early on to not participate, because of some other things that were going on. One station was -- due to a clerical error, was inadvertently left out. We thought we were administering them. We were not. Anyway, that is where those two stations went to.

Had an average of 46 individuals. So we had a 46 percent response rate from each site. Had a total of almost 300 individuals. It was over 300 who gave responses and responded. However, a number of them were missing a lot of data, and so we came to about this 2876 that we felt provided enough valid responses to the majority of items.

DR. BARNES: He meant about 3000.

DR. KOVES: Three thousand, sorry. What did I say?

DR. BARNES: Three hundred.

DR. KOVES: Oh, 3000.

CHAIRMAN BLEY: Ken, we are running a little behind. We are not going to be done by the time of our scheduled break, and it looks like on the next slide we start getting into the stuff that gets complicated. So I think I am going to take a break now for 15 minutes, and we will come back after that.

DR. KOVES: Sounds great. Thanks.

CHAIRMAN BLEY: We are recessed until 10:15.

(Whereupon, the foregoing matter went off the record at 10:00 a.m., and went back on the record at 10:15 a.m.)

CHAIRMAN BLEY: We are back in session.

MEMBER SCHULTZ: Ken, before you move forward, I have a question on the previous slide. I was with you with selecting 100 individuals from each site.

DR. KOVES: Correct.

MEMBER SCHULTZ: It is a small license sample, but it is okay. I was surprised that it was voluntary and that the response rate was only 46 out of 100 on average, and by that I am wondering whether the evaluators would be somewhat suspicious about who did not participate in the survey and who did participate in the survey, given that the survey is focused on safety culture.

DR. KOVES: I have done a lot of surveys in the past, and people often talk about, well, you have got the extremes, what about the extremes, and all this other stuff. It really just -- When you use the means and look at the means, the means wash out a lot of that.

Also, my experience is that the response rate is very strongly driven by executive and leadership communication. So if the leadership communicates frequently that this is important, we need you to do this, then response rates go up. If they give it to their communication person to send out some email to everyone, then the response rates aren't quite as high.

MEMBER SCHULTZ: I don't want to drive too far in that direction, but that is part of my point, which is when the survey is about safety culture and the response rate is 46 percent, I am a bit surprised by that.

DR. KOVES: Well, also this is in addition to all the other things that they do relative to that. You know, pretty much every station has some type of safety culture survey. There is a survey as part of the INPO eval. Typically, they do -- At the mid-cycle, they will do another one. Very often, they will do employee satisfaction surveys.

So there is -- I am not sure how valid it is, but there is always this concern about survey fatigue. So I am not sure. I realize -- Plus, this also was billed as a research survey. It is not like something that is going to be directly impacting their plant.

CHAIRMAN BLEY: Can I ask it a slightly different way? There are two pieces to what I want to ask.

Were there some plants where you got nearly everybody participating and others where you had almost nobody? How is that distribution? The other question was, did you have any idea -- Were you able to discern whether it was a certain class of employee that decided not to participate?

DR. KOVES: Well, going to the first one, we had already said that we were not going to give up on any station until we had like -- I think it was 35 responses, and that is based on central limit theorem.

So once they got over that threshold of like 35, then we stopped sending reminders and that type of thing. So there was that type of an attrition, too. If they were a laggard, then we kept on them until they crossed that criterion.

Based on looking at the demographics, we did not really see any major deviations from kind of your typical plant demographics. So I would say that there did not appear to be any difference in them.

DR. BARNES: Generally speaking, based on what we have seen at other sites with other kinds of surveys, we often see security personnel have a much lower response rate, in general, which is kind of an interesting finding.

DR. KOVES: Plus, historically, they have some of the lowest scores, too.

DR. BARNES: Yes.

DR. KOVES: I mean, I know that we did have security. Security participated in this.

DR. BARNES: Ideally, of course, it would have been great to have 80 percent or more from the entire industry from every site, but as Ken mentioned previously, this was in part schedule driven. So the focus was on getting enough to have a level of comfort that the results were likely to represent each of the sites that were in the population that we were looking at. So it wasn't perfect, but --

DR. KOVES: And if you remind me, I will comment more on that during the closed session.

CHAIRMAN BLEY: Oh, okay.

MEMBER SCHULTZ: And your comment that this as presented as a research survey to understand relationships.

DR. MORROW: When we were evaluating it, we looked at, okay, what is the typical response rate in research based surveys, surveys that would be comparable. That response rate is like on average 30 percent, looking across web based surveys that are for research purposes voluntary. So 48 percent was a reasonable response rate for the purpose of this study, kind of in the context of looking at a cross-section.

We didn't have specific information on the nonrespondents, but we looked at the distribution of work groups. It seemed like, even within sites, that you had a good selection of people from different work groups.

We also did some analyses looking at kind of the stability of the group mean, because what we end up using is the group mean to look at the correlations, and that group mean was very stable in almost every single case.

So adding a few more respondents and tracking a few more respondents wouldn't have really changed that. I will talk about that a little bit more in the next presentation as well.

CHAIRMAN BLEY: Thank you.

DR. KOVES: In terms of real difference between, there was only one station that had accidently gotten administered twice rather than got left out. They had a little higher number. Other than that, they were all fairly within that 35 to 55 range for the n.

CHAIRMAN BLEY: I have got another question here, because both of you have focused on the mean a lot, but I wondered, thinking from the other side where we see events and we see people who didn't do some of the things that the Safety Culture Policy Statement supports, regardless of the mean, if you have a substantial number who devalue certain of these goals, we don't have any way to know. But it could be that those are the ones who get involved in cases where they actually do something that is troublesome, and we blame the safety culture.

That is fine, but if you've got wide variability, I wonder if we are more vulnerable to the thing not working right. I suspect that has to be true. Have you thought about that?

DR. KOVES: Yes. Also, we look at the standard deviation and look at some of the variability within the plants, and that variability of plant performance was important for the validation and the correlations with the key performance indicators.

You don't need a group of individuals. You really only need one who doesn't value it to contribute to -- I won't say cause, but I will say contribute to an event.

CHAIRMAN BLEY: Well, you kind of need a class of them for the event, the susceptible event, to match up in the Swiss cheese with that person. So if it is one, we are pretty unlikely that they are going to be the ones that catch the hole in the Swiss cheese, but if there are more of them, if you've got really wide variability and the same mean, you are more likely to line up.

DR. KOVES: What happens is, if you have more of them that are in that category, then it pulls the mean down, and that is where we ended up seeing it in the results.

CHAIRMAN BLEY: That is why I asked about variability. If you have got some extremely high and some extremely low, then that is not true. But if it is just mostly clustered near the mean and you have got a few outliers on the low side, then --

DR. KOVES: Yes.

CHAIRMAN BLEY: And I don't have any kind of a clue about what you saw in that regard. A standard deviation doesn't quite tell you the same thing as --

DR. KOVES: Well, as looking at the frequencies.

CHAIRMAN BLEY: That is the thing I was just curious about.

DR. MORROW: I was curious about that, too, maybe in a little bit different way, but other ways.

For the most part, safety culture, when you are looking at kind of grouping individual responses to the organizational level, then they are using the mean, because we just don't have enough information to know what is more important at this time. But another way that they have looked at it is just to look at what percentage of the sample indicated that there seems to be a problem with the safety culture.

So they had scores of strongly disagree, disagree, that kind of thing. We did look at that with the INPO survey, and the correlation between the mean score and the percent of problematic responses was like .93. So very, very similar there, and we weren't really getting much new information by using that.

CHAIRMAN BLEY: Well, that is encouraging.

DR. MORROW: Yes. It is kind of -- You know, it is not done as much in the literature yet. So we are more comfortable with the mean, because that means that we can continue to compare these results to what else is out there right now.

CHAIRMAN BLEY: That's fine. I just remember, one of the first PRAs that was done. This isn't people. This is equipment, and we had grand mean failure rate from the whole industry, and our plant aligned all right on that, but when you looked at distributions, we found that we were out there on the 90th percentile on a fair number of things. So we got to do something facility by facility, rather than just an industry mean, because there was variability.

DR. MORROW: I think it is the next step in the safety culture research, is to try to look at are there ways we can kind of tweak this rather than just using the generic mean.

CHAIRMAN BLEY: Go ahead. Now we come to the fun stuff.

DR. KOVES: Oh, fun stuff, yes.

Principal components analysis: Why? Why is it kind of the standard in psychological research. Principal components or factor analysis -- factor analysis is kind of a more generic term, and why? To find patterns. What is the underlying structure in multi-dimensional data is what we are trying to find, so that we can then use that for data reduction.

As I said, this is a research survey, and with this we throw lots of items in, so that we find out which ones are the bad ones, which ones are the good ones, and then can reduce the size of the survey, not just to eliminate bad items but also -- I am a really kind practical, applied kind of guy, and surveys cost plants, organizations, money to administer, and I don't mean just cash outlay. I mean the time that is involved.

So the better instrument we have is both a function of its quality and then also of its length, too. If you have it good and short, that's great. Also, I view it to then build better scales, which is a whole part of the process.

So very often PCA is used synonymously with factor analysis. There are some technical differences.

What I am going to do now is I am going to talk about kind of a conceptual or visual idea of what PCA is. I am not going into going to the map explanation. A couple of reasons for that: One is we don't have time, and also, secondly, because there are just gobs of resources out there. I mean besides on the Internet, books and the Internet, also YouTube videos explaining both, coming at it from a linear algebra as well as a matrix algebra approach.

So lots more examples, but if I took all day, there is a lot more out there than I could cover. But what I want to do is I want to give you kind of a view of, at least in my head, how it works from a more visual perspective.

Think of the data as an -- and this is the data of the individuals, so the almost 3000 people in the data -- as points in multi-dimensional space or think of them as clouds in multi-dimensional space. You know what I am talking about, multi-dimensional. We have 110 items. So we have a possibility of 110 dimensions.

What is a PCA effect analysis does is, first of all, of these clouds of data it allows you to look at different perspectives of these clouds of data to find out which perspective gives you the most information about what is in the data.

CHAIRMAN BLEY: So each point in your multi-dimensional space is one individual score on one of the 110 items?

DR. KOVES: It is the -- Yes.

CHAIRMAN BLEY: Okay.

DR. KOVES: And so these clouds of data -- So first of all, it allows you to look in different dimensions to see what perspective will give you the most information about this. Secondly, what it also does it also uses the covariance matrix and, by reducing the variance, finds a point in the middle of the cloud to best represent that particular cloud of data.

You will see this in a minute. It gives you what are the points -- or actually, it is the items, gives you the items that are associated with that cloud, and then also what amounts to the correlation between that particular item and that theoretical point in the middle of the cloud.

I found this example. This is actually cancer data. What they are trying to do is they are trying to represent this idea. What they are showing here is, in three dimensions, they are showing kind of the plane that best represents this data, how it is folding through three-dimensional space. But if you look at this data from one direction, this is what you get, and this is what you see.

Out of that, you might say, well, okay, it looks like maybe there's two things that are going on here. If you look, you might say that, well, all right, so we have something going on over here maybe and something that is maybe going on over here. However, if you then look at it from a different direction, you get a different picture.

So you are seeing more and better differentiation between what is going on over here and, obviously, what is going on over here, and then you've got something going on out here and over here a little bit.

So when I saw this, I go, wow, that really does look like a factor analysis. I thought it was kind of interesting. It actually never had -- I had kind of that in my head, but I had never seen it on paper before.

So if you were to then say, okay, how many clouds do we have here, well, obviously, we have two large ones. If you were to say we have three, then probably something over here might be the third one. If you said we had four, I could easily see this one being the fourth one.

Part of this is determining -- you know, telling the software how many of these to look at and what does it look like from these different dimensions.

I do want to point out this one over here, and I do want to point out this one over here, because we are going to see some examples -- in my opinion, examples of something like that later. So that is kind of just a visual of what PCA looks like.

What I want to take my time and focus more on is exactly what are the mechanics of actually going and doing this.

CHAIRMAN BLEY: I am not a person who has done much factor analysis outside of school, but when I read this sentence, I have a question about it: "PCA attempts to account for all the variants in the data by creating as many factors as there are data points."

These factors -- The original 110 items are something that make physical sense to me or real sense or something like that.

DR. KOVES: Right.

CHAIRMAN BLEY: I understand it. I understand the scales. These principal factors -- do they have any meaning to them or are they mathematical constructs?

DR. KOVES: Those are mathematical constructs, and the idea is to then narrow that down to the ones that do have meaning.

CHAIRMAN BLEY: Are you going to talk about that some as you go on, how you do that?

DR. KOVES: Yes, and that is exactly where we are going. Really, the thing that you have to do is that you -- I mean, that is where I am going.

CHAIRMAN BLEY: Go ahead.

DR. KOVES: So exactly how do you go about doing this? Well, you enter the data into some --

CHAIRMAN BLEY: This was the least

satisfying part of your paper. You go get a computer, put stuff in, and out comes the answer.

DR. KOVES: Actually, no, it is much more art. There is more art to it than that, and I will explain part of that, and we will actually have the opportunity to -- You will have the opportunity to experience this.

You put it in, you know, whatever it is that you want to use. Ideally, it is one that you are familiar with, because as my old stats teacher used to say, getting answers out is not the hard part, even though when you are first using these, you think that is the hard part. The hard part is getting the right answers out.

Then you determine the number of components, and here is where there are a number of different methods that have been used over time. I kind of like looking at all of them to get a feel for it. Eigenvalue greater than one was Kattell -- anyway, I think Kattell, a number of years ago, and back when they were doing these by hand.

Basically, the theory is that everything -- all the Eigenvalues lower than, smaller than one are just noise. That is really the break point where you really start getting meaningful Eigenvalues.

Then also the screen plot -- it is a graph of the Eigenvalues, and you are looking for where the break is between where they are coming down and where they start flattening out.

The last one that I have also been using that also uses kind of a coherence approach is the results coming out, does it make any sense? Basically, what I was using -- I mean, they were looking at the data also -- is trying to use all three of these to make sure that they converged.

Then what you do is -- What you see is you see a number of items that are grouped together. At that point, you have to decide what is the meaning.

CHAIRMAN BLEY: Just to keep language straight, here when you say items, you mean one of the 110 items?

DR. KOVES: Questions, yes. I'm sorry, one of the 110 questions.

CHAIRMAN BLEY: I just want to be sure. It is not a particular score from one particular person. This is a question?

DR. KOVES: Yes, and the reason is -- I got that from my Chair -- the overlap between surveys and then like SAT doing testing is so similar, they just use the term items.

CHAIRMAN BLEY: That's fine. Just wanted to make sure you were using it the same way.

DR. KOVES: What you get is you get these, and you will see these in just a minute -- You get groups of questions, and you see what they are or at least abbreviated versions, and from that you understand these 3000 people, what was guiding them in terms of their responses to some of these items.

Then you label the groups of the items, and this really ends up being somewhat of an iterative process to make sure that either you have got -- you know, you haven't left something valuable on the table and that you are not including junk.

CHAIRMAN BLEY: In simple language, you are saying we know there is some redundancy and overlap in our 110 items; is there a subset that captures all the key information.

DR. KOVES: And that is exactly what we are looking for. We want the smallest subset that is complete.

So here, hopefully, you can read this.

CHAIRMAN BLEY: We can, over here.

DR. KOVES: You can see, this is literally the output, and you are looking at one of the factors. If you go over to the fourth column here, you see where it starts with .7. If we read down through these items, we see that -- what do they say? They say, my supervisor gives me useful feedback. My supervisor observes me working. My supervisor is usually available when I have a question or problem. My supervisor has personally recognized me. My supervisor responds to questions in an open, honest manner. My supervisor discusses safety before I start work on a job. When I need his decision, my management is usually available. Supervisors are responsive to employee questions. Supervisors are visible in the plant, and my supervisor supports senior management policies.

Now if you had this list in front of you, what might you -- how might you label this factor? I'm asking.

MEMBER SCHULTZ: Supervisory support.

DR. KOVES: Supervisory support, supervisory responsibility, something like that. I is obviously very obvious that what all these items are talking about is the supervisor and what they are doing.

CHAIRMAN BLEY: And, of course, it would have been this one without the factor analysis.

DR. KOVES: Actually, when you get into

these, if you look at some of them -- See, the problem is that very often the items -- even though you want to try and avoid double-barreled items, which are kind of asking about two different things, there are -- Because of the construct of safety culture, all these dimensions are so overlapping, very often items will kind of cross different dimensions.

So if I ask all of you to look at all the items and group them together, there would be a lot of similarity, but some of you would pick on certain elements of that and then move it over to a different factor.

So, basically, what we have done by using this approach is, rather than just asking the six of you or 0 of the people in the room or whatever to do that and then looking at those results and trying to come to a conclusion, here we have asked 3000 people, and we are looking at the results from almost 3000 people.

CHAIRMAN BLEY: I wonder how much the ordering of these among the 110 affects this. You didn't do any shuffling as you went plant to plant or something like that to see if --

DR. KOVES: No, there was no shuffling over effects. We did have a question about order effects with one particular factor, and I can go ahead and get into that when we talk about that a little later. But pretty much, these were scattered. Well, you can see.

CHAIRMAN BLEY: Well, they have the numbers, yes.

DR. KOVES: They have numbers. It is 26, 8, 108,24, 104, 57, 82, 85, 41 and 35, were the different items in the order. So they were spread out all over. It was the prioritizing safety that we saw that, and we go, uh.

MEMBER SCHULTZ: Ken, what are we looking at here in terms of the categorization? You directed us to look at column 4.

DR. KOVES: Okay. Well, you see, what these are, are the -- not loadings component scores -- trying to use the right terminology -- and of how this item -- and of how this item, number 26, correlates with all of the 11 factors. So, basically, this is very close to a correlation.

So, basically, you can think of this as this factor -- or this item, number 26, has a correlation of .72 with this theoretical data point called a component. Okay? What you see these other numbers are, these are the correlations with all of the other components.

The ones that are missing have a score that is less than .10. That is why there is nothing there. It is just an option to delete those out, so it cleans things up a little.

CHAIRMAN BLEY: So the numbers, again, in that column is for the fourth principal component.

DR. KOVES: The fourth component, correct.

CHAIRMAN BLEY: Given that component, this is the r, the correlation factor, square root of the r-squared.

DR. KOVES: Okay. With that particular theoretical point of the component.

CHAIRMAN BLEY: Over all the respondents.

DR. KOVES: Over all the respondents. Correct. Dave? Now don't ask me a tough stats question. It's been a long time.

MR. SOLORIO: I am not going to. Sorry. I just want to remind myself, because I probably missed it. The 11 components or theoretical components are what again? What were they?

DR. KOVES: Well, I am answering the question of how we got to those, and then I will answer what those were in just a minute.

MR. SOLORIO: Thank you.

DR. KOVES: And as you see, these come down in descending order, a nice sort feature, and you see down here at the bottom, number 35, "my supervisor supports senior management policies." You see in here this is the lowest loading. However, it is loaded on this factor, because it has the highest loading there. But you see a number of strong cross-loadings, particularly with 6, also then with 1 and 2.

This is telling you that this item probably is not really as good at picking this out as these ones here up at the top with the .7 correlation.

Just from experience, as a general rule usually when they print out like this, it is real easy to name your factor based upon one of those first three items. Okay?

Here is a different -- another example, and I want to start down here with 75: "It is my responsibility to raise nuclear safety concerns. It is my responsibility to report security concerns. I am personally responsible. I always use human error prevention techniques. If a procedure is not correct, I report the problem, and security is just as important as safety." So this would probably be about personal responsibility for safety.

You see the factor loadings over here that correlate with that. You also see this one here at the very bottom, "Security is just as important as safety." Once again, not very strong on that factor and a lot of cross-loadings.

Now just so that you don't think all of these are real easy all the time, I have put this other one up here above it: "Co-workers hold one another to high standards. Workers usually follow procedures. Personnel do not proceed in the face of uncertainty. Workers maintain questioning attitude. Workers follow procedures and make conservative decisions. Personal conflicts are not allowed to interfere."

I ended up labeling this one questioning attitude. Even though questioning attitude falls a little bit lower, it is what a lot of people talk about. It is also -- when you do subsequent analysis, it actually becomes a little more clear, but they are not always -- I did start out with a couple that are really easy, but actually, to be perfectly honest, most of the time, 80 percent of the time, it is really fairly obvious as to what the items are talking about.

Okay. Here is my last one.

CHAIRMAN BLEY: You just defined these. You do re-sort this table with results on every principal component, so you can see the ones grouped. Is that what you did?

DR. KOVES: You choose the option to sort them, and it prints them out as sorted. So the ones that have the highest loadings in whatever factor, and so it loads them on that factor. It prints them out that way, and then orders them that way. It makes it a lot easier than trying to pick through it.

Okay. So here we get to the end, and here we have training: "My managers assure high quality training. Training helps understand how I contribute to safety. Continuous learning is expected of everyone. Training at this station reinforces safe working behavior. New personnel know the difference working in a nuclear site."

So this takes us out, and you see that we are out to factor or component nine, and you are also seeing the loadings are starting to -- normally, the highest loadings are going to be on the first factors. Also, the most variance is going to be accounted for by the first factors. This is just a function of the way PCA works. However, you get down here, and you see factor number 10. it has three items in it: "I know how to enter an issuing cap. Management oversight is provided for safety significant t asks, and safety culture assessment -- we have had a safety culture assessment in the past two years."

To me, there was no real coherence or theme in this particular factor. So here is where we get into that coherence idea. It is like, okay, that is junk. Then this very last one: "This station has a knowledgeable and experienced workforce." You see that it only has three loadings in the whole survey that are even registering here, and it is off on its own.

If you recall, we had that graphic, and I said, you see these little ones out here in the fringe that are off on their own? I thought that was a really nice graphic of what these kind of look like.

So this is how we ended up with nine factors and going forward with nine factors.

Let's see. Where am I? Basically, what we then did was, because we had so many items, obviously, you are going from 110 down to seven factors or seven components, a number of them had a lot of items in those.

So what we did was we then took those items -- so, for example, in leadership -- and we then did a factor analysis or a PCA on just that set of items on that subset of items, looking to see if there were any patterns within that subset of items, and that is how we got to the subfactors.

So what did we come up with? The very first -- and these are in order in which they came out, which meant that they are accounting for less and less variance as we go through.

The first one was management responsibility, and I have seen it time and time again, and you read the literature, and it kind of lines up with the literature where everyone says, hey, you know, it is management that takes the lead on driving the culture. Well, ironically, it actually works out that way statistically, too. So, convenient.

What were some of the subfactors underneath there? Respect for work environment; continuous improvement; performance indicators; resources and rewards. Once again, these are labels that I put on it based on -- through that process that you saw where you look at the items and then you draw a conclusion as to what it is that they are talking about, and put a label on it.

MEMBER SCHULTZ: Is the subgroup a group of items or are there any prioritization?

DR. KOVES: Once again, these come out in the order of how much variance they account for. So it is also listed that way, too.

MEMBER SCHULTZ: Thank you.

DR. KOVES: Also, you know, we haven't gone through it, but you will see that usually the first factor -- not does it account for the most variance. It typically also has the most items in it also.

Going back to that one graphic, you saw the one group that was the largest. That is why I said, wow, this really looks like a factor analysis.

The next factor that came out was willingness to raise concerns. Factor analyzed that, and came out with kind of two subfactors about the willingness to raise concerns, both informally and then formally.

In terms of decision making, which was the next factor, factor analyzed that, and there was nothing that made sense there in terms of subfactors. So this is just a definition. Decisions are conservative, timely, safely focused, and engender confidence.

The next one, which is supervisor responsibility, had four subfactors that came out: Communication; presence or availability in the field;

coaching ability; and then one item that was talking about alignment with management, how aligned were they with management.

CHAIRMAN BLEY: That one is not in your tables in the report, the fourth one.

DR. MORROW: Because it is one item, yes. One thing we did, especially when we were looking at the correlations, is because when you get down to the subfactor level they tend to be just two or three items, especially -- which Ken will talk about just briefly -- is once you go from like a 110 item survey down to this final 60 item survey.

So those factors, the subfactors, are no longer very stable. So to continue to conduct statistical analyses on these -- it doesn't hold up as well.

DR. KOVES: The reliability of one item is about .2 in theory, and for our scales we want reliability at least of .8 or more. So the reliability really goes down.

DR. MORROW: And best practice to kind of triangulate a factor is to have at least three items. So when you get down to that subfactor level, there is just not as many items. They do a really good job of kind of informing what the factor is, but in terms of analyses, we kept everything to that main factor level, so the nine factors that Ken is talking about.

CHAIRMAN BLEY: Okay.

DR. KOVES: In terms of questioning attitude: Situation/problem awareness; process use; plant knowledge.

Then for, once again, communication, it was not a -- there were no real subfactors inside of there, but it talked about communication that is broad, includes plant level communication, job related communication, worker level communication, equipment labeling, operating experience, and documentation. These are all the types of things that were in there.

Lastly, personal responsibility: It is my responsibility to report concerns and practice nuclear safety. You saw that up there.

Then this one, prioritizing safety: Nuclear safety is a priority that is seen in meetings, expectations, coaching, and decisions. This was probably the most -- at least from a philosophical perspective, the most problematic of all the factors -- or components.

CHAIRMAN BLEY: Why it explains so little? Is that what you mean?

DR. KOVES: No, because first of all, it

happened to be -- inconveniently, it happened to be the first six items of the survey. So we had a real serious question about order effects.

What made it more problematic was that, as you cut back your number of factors and reduced it down, it would not go away. It stuck together like glue. The other factors would blow apart, you know, and distribute themselves -- The items would blow apart and distribute themselves across other factors. This one just kept sticking together and sticking together.

So from a philosophical perspective, those are the things that were kind of driving us crazy. So we are not -- Because of the order effects question, in the 60-item survey we have three items representing this, and they are spread across. So we are going to see if this hangs together in the future or not.

CHAIRMAN BLEY: Yes. I was kind of surprised in that first one we looked at that they were spread so widely. I assumed you did that on purpose, because you had seen the linkage among them.

DR. KOVES: No.

CHAIRMAN BLEY: That isn't true? It just turned out that way?

DR. KOVES: It just turned out that way.

CHAIRMAN BLEY: That is curious. Okay.

MEMBER SCHULTZ: Ken, we are down to the eighth component here.

DR. KOVES: Safety, yes. Then the ninth one was training quality. Very narrowly focused: Training is high quality; supported by management; encourages nuclear safety.

As you see, most of these were really pretty easy to define and see what it was they were talking about. That was the other thing about prioritizing safety. It was really kind of -- It was more difficult to really say what exactly is this talking about. So for a number of reasons, that was our problem child.

CHAIRMAN BLEY: On the other hand, Steve, factors two through eight are of the same order in contribution and reducing the variance.

MEMBER SCHULTZ: That's right.

CHAIRMAN BLEY: Even though it is eighth, it is not that far away from two and three.

MEMBER SCHULTZ: Yes.

DR. KOVES: So then what was the next step? The next step is reducing, trying to get down from 110. My target was around 60 items. Basically, the way I write items and the way these items go, people will respond to about two or three a minute. So then you can -- Sixty items, they can do it in about 20 to 30 minutes, plus 60 items, if you have got seven factors, really -- or no, if you have nine factors, about 30 items is probably about all you need to capture those factors and have a good survey.

So it is kind of giving more information than the bare minimum, but also reducing the number of items.

Here is where this probably most becomes an art, because what you do is you take a number of different facts about these items, and then make a judgment call as to which ones you are going to keep and which ones you are going to drop out.

So some of the things that we looked at were, first of all, the number of missing data points. There were some items that were missing lots of data, and it made sense when you looked at them, because the bulk of our sample is coming from craft employees, front line, that type --

MEMBER STETKAR: When you say missing data, people didn't score that item.

DR. KOVES: Didn't score that item, right. They put N/A, the eighth option. A lot of them, if you looked at them, these were things that would probably be more known by management and not by craft. So for the purposes of this survey, we didn't feel that those items that a lot of people didn't feel like they could answer would be good moving forward.

Also looking at the reliability of the scale with the item when you remove it -- However, this didn't come into play much, because -- I mean, the reliabilities were so high with so many items, and the high inter-correlations.

Then you look at the inter-item correlations. If you have two items that are highly correlated, really, one of them is redundant. One of them is not giving you a lot of additional information. So you kind of then also look at item content.

Then also what came into a strong play on this one was correlations with the key performance indicators. If there was an item that had a strong correlation with a particular indicator that maybe was not in the other items, it was more a candidate to keep.

So it is looking at these, all these together, and making a judgment call as to which ones to keep and which ones to drop out.

MEMBER STETKAR: Can you go back? I am

trying to get my head around the rationale for deleting items, because specific elements of the demographic did not feel they were relevant, let's say, which is, I think, what I heard you say.

DR. KOVES: Right. Yes.

MEMBER STETKAR: On the other hand, if the elements of the demographic that felt they were relevant and did respond were elements of the demographic that indeed strongly affect nuclear power plant safety, are you screening things out that you ought not to be?

DR. KOVES: Now that is a very good question. It kind of goes back to the purpose --

MEMBER STETKAR: I understand in a holistic organizational sense you might rationalize that, but --

DR. KOVES: And what I would recommend then is I would recommend -- because you are right. I mean, you have got a good point, but I would recommend that you then have a management level survey that you would administer to management.

First of all, it is more targeted. You can ask better questions, and then also you are not having a lot of -- you are not spending a lot of utility time on people reading things that they then say, I have no idea.

So that is a valid point, but for the purpose of the survey, which is to be very broad, that doesn't mean that we got rid of all of them. Actually, because of some of the validation, we kept some of them, but it was one of the criteria that was mixed in with the whole thing.

We did keep a few of them, but I would say, if that is a concern, then let's put together a management survey and really do that one right, and really get what we want to get at.

DR. BARNES: And that is where some of the research is going, is developing targeted surveys that focus on different levels in the organization, and it makes sense, you know.

DR. KOVES: But, good point.

Okay. So if we could go ahead and close this session now, I would appreciate it.

(Whereupon, the foregoing matter entered into closed session at 11:02 a.m. and went back on the record at 11:20 a.m.)

CHAIRMAN BLEY: This meeting is officially open again.

DR. BARNES: Process question: Do you want to take a couple of minutes and look through the rest of the survey items before we gather them up again?

CHAIRMAN BLEY: I don't think we have the time, really, to do that, actually.

DR. BARNES: You have seen enough.

MR. SOLORIO: If you want to look at them.

CHAIRMAN BLEY: It will be in the record. We can look at them later. We are an hour behind.

DR. KOVES: And Stephanie is glad, too. She was thanking me.

CHAIRMAN BLEY: At least, the Subcommittee, I think, doesn't have a problem if we go beyond 3:30. I hope you don't either.

DR. KOVES: I think my flight is at six. Should be leaving about now. So that means I would kind of like -- I need to leave by four.

CHAIRMAN BLEY: Okay. If we aren't finished by then, we will stop just before then and see if anybody has anything else for you or you have anything more for us. You have one more talk. So, Stephanie, go ahead. We will break for lunch at some natural place sometime Noon or a little later.

DR. MORROW: Sure.

CHAIRMAN BLEY: Whenever it feels right.

DR. MORROW: Okay. My presentation today

is really focusing on the White Paper that you reviewed, and is the Office of Research's evaluation of the INPO survey and the construct validation study that they conducted.

CHAIRMAN BLEY: Now Ken is going to talk more later. So I guess you will do everything up to that validation part. Then Ken will go, and then you will come back.

DR. MORROW: Right. So my morning presentation is really up to where Ken left off, which is with the six-item survey, so looking at the factor analysis and kind of the reliability of the survey. Then in the afternoon, we will show the correlations between the INPO metrics and the NRC metrics.

CHAIRMAN BLEY: Okay. go ahead.

DR. MORROW: So again, as kind of outlined in the White Paper, when we were looking at doing an evaluation of the INPO study, we were looking at a couple of primary global questions. One, is the survey valid and reliable? Do we have confidence in the survey in the context of this study?

Also, does the survey show support for the Safety Culture Policy Statement traits? Again, we were doing this as a user need from OE, and this was kind of happening at the same time as the Policy

Statement traits were rolling out.

Also, is there a relationship between the safety culture survey results and safety performance metrics? That is what we will talk about this afternoon.

I am going to start right off with kind of the key findings from the White Paper, which I will talk about in the remainder of this presentation.

One is that we found that the INPO Safety Culture survey did demonstrate evidence of construct validity and reliability within the context of the study that was conducted.

Also, although there wasn't one-to-one alignment with the Safety Culture Survey factors and the NRC's Safety Culture Policy Statement traits, they did share many commonalities. So there were some consistent themes that emerge from the survey factors and the Policy Statement traits, and each trait was represented by at least one survey factor. So that was kind of promising, from our perspective.

First, kind of our evaluation approach: How are we evaluating validity and reliability? We look at it from the perspective of construct validity, and we use the term construct to describe an IDS. In this case it is safety culture, but it can also be things like intelligence or personality.

When we think about kind of the validity of the construct, what we are asking is: Is the survey measuring what it purports to be measuring? We can look at that in a number of different ways, but the ways that we focused on within this evaluation is the content validity of the survey -- so, really, does the survey cover kind of the breadth of the safety culture construct -- and also the criterion related validity. This was where we look later at the correlations: Does the survey demonstrate a relationship with outcomes that it should theoretically be related to.

So kind of the overlying theory and what Val talked about earlier today is that we believe safety culture to be related to safety performance. We have some research literature showing that there are about small to medium size relationships. So we want to see if those relationships appear using our data and using this INPO survey.

Also we were looking at reliability. So kind of a prerequisite for validity, is it measuring what it is supposed to measure, but does it measure it consistently?

We looked at two types of reliability in this evaluation, one internal consistency. This was specifically looking at the items, and do they produce similar results? So we just saw about how there are items that are grouped within factors, and do those items kind of hold together in a reliable way? It is almost an additional check on the factor analysis to see if there is some internal consistency there.

Also within-group reliability --

CHAIRMAN BLEY: So that is a qualitative look?

DR. MORROW: The internal consistency?

CHAIRMAN BLEY: Yes.

DR. MORROW: That is statistical. That is a quantitative look.

CHAIRMAN BLEY: But separate from the factor analysis?

DR. MORROW: Separate from the factor analysis. It is a different way of looking at it.

Also within-group reliability: This is really about -- I kind of mentioned earlier the stability of the group mean. So one thing that we are doing with safety culture -- and I will speak to safety culture specifically, because we are conceptualizing this as an organizational level phenomenon, but we are measuring it with a survey of individuals' perceptions.

So somehow we have to go from that individual level to the organizational level. It is really important to somehow reflect whether or not the survey results show that this is a concept that is shared among members of the same organization. So that is kind of what the within-group reliability is tapping into. Is this a shared idea?

In addition, kind of just some more holistic evaluation questions that we thought about as we were reviewing the study is, first, were the data collection procedures and the resulting sample appropriate for the research questions?

Is the data analysis approach consistent with good practice in social sciences? And then given the data, would an independent researcher produce the same results and arrive at similar conclusions?

So here we wanted to go above and beyond just looking at what INPO did, running the exact same, identical analyses, and saying, oh, yeah, we got the same results, we used the same analyses, but going deeper and looking at, okay, they used this analysis; was this analysis appropriate. Was this sample appropriate? Was the response rate appropriate, given kind of the style of the study? So we were really looking at all of those questions when we looked at the study.

So I will start by talking about the content validity, and this is more of a qualitative subjective view. If we had some perfect measure of safety culture out there, if there was one that was well accepted, it existed, then the ideal would be to compare this survey against that true score on safety culture, but that doesn't exist.

So what we do is we look at, in some ways, a more subjective look so that we can establish does this survey seem to get at the same things that other surveys have touched on. Does the content seem appropriate? Does the content cover all of the possibilities?

Ken presented quite a bit on kind of how the survey was developed, how the items were constructed. So when we look at kind of where the items came from, they were drawn from multiple sources that are relevant to the nuclear industry, INPO, NRC ROP, IAEA, and also surveys from non-nuclear research literature. So there was a good spectrum of kind of sources where these items came from.

We also look at how the survey was developed. Is it following good practice in the social sciences? We talked a little bit about the rating scale that was used, the 7 point rating scale from strongly disagree to strongly agree, and that is one of the most common used in the social sciences.

It is also important to kind of go forward with this more standardized rating scale, because it allows us to take what is data about employee perceptions, my perception of the culture, and make it quantitative.

We are taking someone's saying that they strongly disagree with this statement, and we are assigning some number to it. So you need the standardized rating scale, so that you have perceived equal distances between the potential options, response options.

The other thing we looked at was how the items were written, and Ken talked about kind of the original USA survey, and that those items were not written to be survey questions. You really want kind of a specific type of item for a survey question. You want it to be simply written.

You want it address a single topic, and to the extent possible, you want to avoid double barreled statements, something like safety and security -- that would be confusing for someone taking the survey.

Well, what is more important, safety or security? Do I see those as conflicting things?

DR. KOVES: Safety is good. Security is bad. You know, then how do you answer?

DR. MORROW: Yes. So you really want simple statements. What we saw is, for the most part, that there were kind of simply written statements. What we have on the slide there is kind of some examples of the good practice items in the survey.

MR. WIDMAYER: I had a question.

DR. MORROW: Sure.

MR. WIDMAYER: Is there something that is typically done to look at content validity that either you didn't do here for some reason or that you didn't think it did a good job or did it pass muster in all of your content validity examinations?

DR. KOVES: That is why we added the 51 percent more items, to make sure that it did cover.

MR. WIDMAYER: Okay.

DR. KOVES: Does that answer your question?

MR. WIDMAYER: No.

DR. KOVES: I'm sorry. I will show you.

MR. WIDMAYER: You guys are looking at this independently for validity.

DR. MORROW: Right, yes.

MR. WIDMAYER: Okay. You mentioned three things where you said --

DR. MORROW: It passed muster.

MR. WIDMAYER: Yes. What did it not pass muster on, something that is typically looked at for content validity or did it do well in every area?

DR. MORROW: IN terms of content validity, from what is in the literature now --

MR. WIDMAYER: Typically?

DR. MORROW: Yes -- it covers the breadth of kind of safety culture, and we also see that when we look at the alignment between the survey and the Policy Statement traits, the fact that -- What I will talk about later is that, for each trait, there is at least one factor or subfactor that kind of covers the idea that is in that trait. So we see some correspondence there that kind of also bolsters that content validity argument.

MEMBER SCHULTZ: And just for clarification, were we talking about the 110 question survey here or the 60 question resultant set?

CHAIRMAN BLEY: The traits are with respect to the NRC Policy Statement.

DR. MORROW: Yes, the traits are with the

NRC Policy Statement. In terms of evaluating the content validity, that was originally with the 110, but we also looked to see what items didn't make the cut, and did they seem reasonable.

For the most part, it was items that didn't seem to fit as well within a factor. So we look at the factor analysis results and, if they were items that kind of loaded all over the place, they didn't load cleanly on one factor, then those were some that kind of were included.

So what kind of you ended up with, with the 60-item survey, is a cleaner survey. You are tapping into all of the main factors that came out of the factor analysis without some of the noise that was with the additional items.

MEMBER SCHULTZ: Understood. Thank you.

CHAIRMAN BLEY: I don't have the tables in front of me again, what you just said, I think, says that if a subfactor under a principal factor also had effects on other principal factors, you somehow got rid of that cross-effect? How would you do that?

DR. MORROW: A single item.

DR. KOVES: A single item.

DR. MORROW: A single item.

CHAIRMAN BLEY: A single question?

DR. MORROW: Yes.

DR. KOVES: If you go back to the tables that I had, and you saw that usually the items at the bottom, there were cross-overs. Those are really less desirable, because --

CHAIRMAN BLEY: And are those the ones that ended up in the four down at the bottom on the new -- You retained all the 110 somewhere, right?

DR. KOVES: Well, they were all on that list. The 106 were in factors, and it was those last four that we just basically said --

CHAIRMAN BLEY: And those are the ones you are talking about, that those four didn't align well with any one factor?

DR. MORROW: Right. I think one of the first things that happened when they went down to a reduced item survey, which is --

CHAIRMAN BLEY: Sixty item survey.

DR. MORROW: Yes, the 60 item survey, which is standard practice, when you are developing a survey.

CHAIRMAN BLEY: So between the 60 and 110, some of the ones that disappeared were ones that affected multiple?

DR. KOVES: Correct.

DR. MORROW: Right. They didn't load cleanly on one factor.

CHAIRMAN BLEY: Okay. So we haven't actually seen what those 40 are.

DR. KOVES: That went away?

CHAIRMAN BLEY: Yes.

DR. KOVES: No. What you had was -- This was what was left.

CHAIRMAN BLEY: That is what is left.

DR. KOVES: Yes. You don't see that. You can do the analysis later.

CHAIRMAN BLEY: So I expect, when I do the analysis later, I will find that among those 40 are the ones that affected multiple principal components fairly strongly.

DR. MORROW: Right.

CHAIRMAN BLEY: And that is part of the rationale.

DR. MORROW: Yes, and Ken talked about that a little bit, that they used different criteria for looking at which items to eliminate, and still there was a good kind of breadth of the literature covered, and items were still taken for multiple sources.

CHAIRMAN BLEY: I think I am envisioning

a problem that doesn't exist. I am just thinking that if one of them affected many, but really affected one of them quite strongly, and you take it out, we are losing something important, but there was enough near-redundancy in the subfactors that that is almost surely not a problem.

DR. MORROW: Right, exactly.

DR. KOVES: And even in the subfactors, the ones that came out were the ones that had more cross-relevance as a general rule, but some of them might have stayed for other reasons on that list of criteria.

DR. MORROW: Okay. So next we looked at the data collection and kind of the characteristics of the sample that was used for the study. First just kind of pointing out, they started with a web based survey. That is a pretty big assumption, you know, just as this is the collection methodology that we are going to use.

When we look at what is good practice in the social sciences, surveys are the most appropriate means if you are looking at trying to capture employee attitudes, perceptions, values, things that aren't directly observable, particularly by outsiders.

So just I myself going into an

organization and trying to say what is the safety culture here is something where it is actually more efficient, less expensive to do a web based survey of the workforce, and you get much better date; because what I see as the safety culture around here, I might be misinterpreting what I am looking at. As an outsider, I wouldn't have as good of an idea of what I am seeing.

You can get a lot more rich information from the employees. So in this case, a web based survey was an appropriate means to collect the data for this study.

Also they looked at a cross-section of the nuclear power industry. So in the context of this study where it was really a construct validation, the cross-section of the industry was most appropriate. Again, Ken had talked about this 97 percent of the operating nuclear power plants. They used a ratified random sample to select 100 individuals from each site, and they had a response rate of about 48 percent.

That kind of cross-section of the industry, getting a 48 percent response rate might not be appropriate in other circumstances, but when we are looking at this voluntary research based survey, we kind of just want to get a feel for what does safety culture look like across the entire nuclear power industry. Then these are adequate for the questions being asked.

Also, the survey touched on different occupational groups, and here on the slide you can see some of the percentages of the final sample of the survey. Seventeen percent was maintenance, 16 percent operations, 10 percent security, so on and so forth. This is just some examples of the percentages, and also long term contractors were included in the final sample. The sample was seven percent contractors.

CHAIRMAN BLEY: Is this a percentage of the final respondents or of the 100 selected?

DR. MORROW: The final respondents. So the final sample of 2,876. These are the percentages from different work groups.

MEMBER STETKAR: If I add quickly, that is 61 percent of your respondents. Where were the other 39 percent, just randomly scattered through?

DR. KOVES: You know, there is HR, etcetera, basically balance of plant.

DR. BARNES: Housekeeping.

MEMBER RYAN: There were 2876 respondents across 63 sites. What was the distribution for per site? What is the mean standard deviation for that?

DR. KOVES: For each site?

MEMBER RYAN: Yes. What was the variability site to site.

CHAIRMAN BLEY: You got at least 35, you told us.

MEMBER RYAN: Did they average about the same?

DR. BARNES: You mean the number of respondents per site who gave us usable data?

DR. KOVES: Okay. It ranged -- I don't know the exact number, but it ranged from about 35 to about 55, except for the one station where we accidently double surveyed, but other than that, it was pretty much in that range, although I don't know --

DR. MORROW: It was pretty consistent. I don't want to misrepresent what the actual mean and standard deviation were, but it was fairly consistent.

MEMBER RYAN: Sixty-five and 35 is the range of respondents per site, something like that?

DR. KOVES: Yes, and I wouldn't even say 65 was. I would say more about like 35 to 55, because once they got over that 45 limit, we kind of ignored them, and they would trickle in some more. But it was a modestly tight range.

MEMBER RYAN: Okay.

DR. KOVES: That is the technical term.

MEMBER RYAN: Very good. Thank you very much.

DR. MORROW: Okay. Ken spent a lot of time on this. So I only have one slide devoted to the principal components analysis.

Again, principal components analysis is a type of factor analysis. Some of the unique characteristics of principal components analysis is that it tries to create unique factors until all of the variants in the items is accounted for.

So again, you can have 110 factors to account for the 110 items, and where the Eigenvalue greater than one comes from is that they looked at what would be a reasonable number of factors to retain, and the Eigenvalue being greater than one is saying that this factor contributes more than a single individual item would if just had 110 factors. So you look at keeping the factors that are at least contributing more than one item would on its own, and that is kind of where that comes from.

The table I have on this slide is just kind of showing you the variance accounted for by each of the factors that came out of the analysis, and in this case, because 110 were near 100 items, Eigenvalue equals greater than one is about greater than one percent of the variance.

So in this case, you are starting to get down to, with factor 9, that 2.7 percent of the variance.

CHAIRMAN BLEY: And by this time, you've got about 60 percent.

DR. MORROW: Yes, exactly.

CHAIRMAN BLEY: So the remaining 40 is just scattered over all sorts of little --

DR. MORROW: Right.

DR. KOVES: The rule of thumb that I learned -- I don't know about you guys -- was, you know, the goal is that you want to capture 60 percent of the variance with 30 percent of the variables or less.

DR. MORROW: And 60 percent is what I learned as well. So on something we looked at, okay, we are at 60 percent. So this is reasonable.

MEMBER STETKAR: One day you can explain to me the strong basis for 60 percent, but go ahead.

DR. MORROW: So when we looked at the principal components analysis, we were really interested in, okay, so this is what you get if you run it this way; you know, what if we run different kinds of factor analyses? What if we make different decisions about keeping items or removing items using some of the same criteria that ken talked about when it got down to the 60 item survey.

Then again, okay, now we have a 60-item survey. If we run the factor analysis, do we still see the same factors coming out when we just look at those 60 items?

CHAIRMAN BLEY: You haven't given the 60 factor to anybody yet, have you?

DR. MORROW: It is the same. The conclusion is it is the same factors, that you still have the same factors.

DR. KOVES: Well, he is asking about using new data. Actually, we have used it with one vendor, but that is all that we have used it with.

CHAIRMAN BLEY: One vendor? That means all the power plants that used that vendor or people at that vendor?

DR. KOVES: The people at that vendor. So that is not exactly a power plant either.

CHAIRMAN BLEY: Right.

DR. KOVES: And I have the data, but I

haven't factored it. So I am not even sure. However, we did use the 110 with a different vendor, and the factor structure stayed pretty stable. It wasn't exactly the same, but it was pretty similar.

CHAIRMAN BLEY: I am just curious. When do you expect to have some experience using the new survey with some power plants?

DR. KOVES: As the opportunities become available.

CHAIRMAN BLEY: I have no idea what that means. In the next two years, are we likely to see one or 50? Any idea?

DR. KOVES: Can we talk about this in the closed session?

CHAIRMAN BLEY: Absolutely.

DR. KOVES; I will be glad to answer it. It depends. Six years of graduate school it took me to learn that. It is the right answer.

CHAIRMAN BLEY: Go ahead.

DR. MORROW: All right. So it was basically a sensitivity analysis that we did. So you can conduct different kinds of factor analyses with this data to see if we would get the same factors.

CHAIRMAN BLEY: Tell me a little bit about what you did. The 60 items were among the 110 items?

DR. MORROW: Right. Yes.

CHAIRMAN BLEY: So you went back to the original surveys and just looked at those 60 items then to see?

DR. MORROW: Yes.

CHAIRMAN BLEY: But we already know that they are going to work, because they came out of the factor analysis.

DR. MORROW: When you take away items -- When you are doing the factor analysis, part of that kind of identifying dimensions is based on all of the variance in the items. So if you take away 40 items, then you could see different dimensions start to emerge where some of those items were kind of contributing to, for example, a safety communication factor.

Once you take away some of those items, suddenly there is not enough items within the remaining -- that factor, and they kind of move apart. They attach to other factors.

MEMBER STETKAR: But didn't the winnowing process, by definition, remove that source of variability? As I understood it, you said, if you saw a specific item that had fairly broad applicability, you tossed it out, because it didn't reinforce the factor. So why --

DR. MORROW: It removed some of that variability, but what can also happen is, because that item was in there and it was loading on different factors, it can force other items to kind of --

CHAIRMAN BLEY: Agree with some of the support from one of the factors. Some rocks may come up, but it is hard to imagine that it won't still -- won't align with the original factors.

DR. KOVES: Well, actually, on the big ones, yes, it is usually on the ones that are accounting for a lot of variance. Those are the ones that usually end up being more fragile.

CHAIRMAN BLEY: Yes.

DR. MORROW: Right. That is consistent with what we found, is that the most stable factor was that management responsibility. Those items tended to stay together, no matter what other items were included in the analysis.

CHAIRMAN BLEY: We would be looking at methodology problems, if that didn't come out.

DR. MORROW: Right. That is an inherent part of the principal components analysis, is the first factor always accounts for the most variance. You see that, no matter what.

Actually, the least stable factor that we found was the safety communication factor, which is sort of interesting, because it is not the one that accounts for the least amount of variance. It is kind of in the middle there. it is number six, but those items were kind of -- Depending on the other items that were included in the factor analysis, those items may be loaded on the management responsibility factor, the decision making factor, I think also the supervisor responsibility factor.

When we look at those items and kind of look at it more from a theoretical standpoint, it made sense, because communication is inherent to multiple aspects of kind of even just plant operations, but sometimes it is about your communication of your supervisor or how management communicates to the rest of the workforce. So those items didn't always kind of stay together in a single factor.

DR. KOVES: Can I throw in what I thought was an interesting tidbit?

DR. MORROW: Sure.

DR. KOVES: If you look at the principles for a strong nuclear safety culture, there is eight of them, and there is not one for communication. Yet when you look in the individual items, you see communication scattered throughout all the eight principles.

Well, what I found just totally fascinating was, yeah, if you manipulated this a little bit, that is exactly what happened to communication. it just kind of splattered across everything, and wasn't separate.

MEMBER SCHULTZ: It is an interesting observation.

DR. KOVES: Yes.

MEMBER SCHULTZ: Very telling.

DR. KOVES: And "splatter" is the technical term.

DR. MORROW: So the next, once you have the nine factors that came out of the principal components analysis, Ken talked about the factor labeling that happened, and he actually went into that pretty extensively.

So what I have on this slide is just kind of when we looked at duties, are these factor labels appropriate? Do they seem to make sense, because this is a very subjective -- it is more of an art. It is the art part of the factor analysis.

What we have here is just an example item for each factor that kind of illustrates why we thought the factor was labeled as it was labeled. So we found support for the factor labeling. It did seem like these were reasonable labels for each of the factors that emerge from the analysis.

MEMBER SCHULTZ: Was there some validation of this part of the process? It seems very qualitative.

DR. MORROW: It is very qualitative.

MEMBER SCHULTZ: Did you do a peer check in terms of this selection in the examples?

DR. KOVES: We didn't.

DR. MORROW: A peer check of these examples? No. This was the peer check, in a way. This is us looking at, okay, INPO labeled these factors in such a way, does it make sense? If we look at these items, would we label it the same way?

MEMBER SCHULTZ: How broad was "we"? That is part of my question.

DR. BARNES: There were three or four staff at Idaho National Lab that did a check on the PCA and the labeling at the beginning, myself, staff. I think that was the extent of it that I know did theirs independently.

MEMBER SCHULTZ: But a healthy group of experts who looked at the elements that comprise this result. That's fine. Thank you.

DR. BARNES: Yes.

DR. MORROW: The next thing we looked at was the reliability of the survey factors. So for this, this was a quantitative analysis. We used Cronbach's Alpha, which is a measure of the internal consistency of the items. The kind of rule of thumb cutoff value is .70.

Again, we stayed at the main factor level, because I talked a little bit that, when you get down to the few items that are at the subfactor level, they aren't as internally consistent. You want at least three items per factor, really.

What you can see here is that all of these values are above the cutoff of .7, which indicates the good reliability of the factors. This is about whether the items in the factors seem to be measuring kind of the same thing. So you can see that there is a very high value for the safety culture overall.

CHAIRMAN BLEY: I didn't chase this Cronbach's Alpha thing. I don't know it. Tell us a bit about what it is actually measuring, and why .7 is good reliability and what you mean by that. You have told us what you mean by that, but why is that good, and what actually is it measuring?

DR. KOVES: It is measuring the average inter-item correlations, and it is a combination of the inter-item correlations and the number of items. So that is part of the reason --

CHAIRMAN BLEY: Correlations through the data from the test?

DR. KOVES: Right. The inter-item correlations of those particular items. So for training quality, we got three items. What are the inter-item correlations of these three items? So the number or the actual Alpha ends up really being a function of the number of items and the correlations between them.

So if they are not very correlated, then you don't have a lot of confidence that you are getting at the same thing.

Point-seven? You know, honestly I am not sure if that is one of those things like .05 where it is just like this is what everyone uses.
So there may be a rationale beyond that, but I don't know what it is, if there is.

CHAIRMAN BLEY: You call it reliability.

DR. MORROW: Right.

CHAIRMAN BLEY: And if we just take one of them, the management responsibility one, this is then going through all of the items that said management responsibility --

DR. MORROW: Correct.

CHAIRMAN BLEY: -- and seeing if all of those are matching on their scores, kind of that sort of thing?

DR. MORROW: Yes.

CHAIRMAN BLEY: The error that it is looking at is -- it is not -- you know, your overall correlation coefficients measuring -- I am still not completely sure what this is measuring.

DR. MORROW: It goes through a process of pairwise correlations. So it takes like the 20 items that are in management responsibility and goes through, and for each pair of items that looks at the correlation, and what you get is kind of that average of all of the pairwise correlations.

CHAIRMAN BLEY: Of all of the pairwise?

DR. MORROW: Right.

CHAIRMAN BLEY: So it is pairwise. If all of them align, this is going to get a one up there.

DR. MORROW: Right. Exactly. if they are all getting the exact same response, then you will hit a one. So what this is, is kind of an overall average.

CHAIRMAN BLEY: It is just like an algebraic average of all of those pairwise correlations?

DR. MORROW: I am; not certain on that point.

DR. KOVES: I don't remember the exact formula. I remember the primary elements are the inter-item correlations and the number of items.

DR. MORROW: And the number of items, yes.

MEMBER RYAN: I got it for you. It is K over K minutes 1 times 1 minus the sum over K of sigma y squared divided by sigma x squared.

MR. WIDMAYER: That is what he said.

DR. KOVES: Thanks for reminding me.

MEMBER RYAN: There you go.

DR. MORROW: But it is an alternative to like a test/retest reliability where it is like, if we took this item and only used this item to measure management responsibility, would we get the same result as if we used any of these other items? So it looks at the consistency of the items in measuring that same thing.

So we want there to be at least some reasonable amount of consistency, so that we can say that these items are all measuring kind of the same underlying thing.

CHAIRMAN BLEY: And personal responsibility for safety is getting close.

DR. MORROW: Yes. The other thing is, it is a function of the number of items. So this is inflated when you have many more items. You will notice like the .98 with 60 items. That is another reason why we are looking for at least three items, so that we have at least the possibility of having some internal consistency.

CHAIRMAN BLEY: Okay.

DR. MORROW: The next thing we looked at was the within-group reliability of the survey items -- or I'm sorry -- of the survey respondents. Thank you.

What we did for this was to look at two types of intra-class correlations. So again, all of this is kind of based on correlations, looking at the relationships.

For this, what we are looking at is the relationships between how people from the same site responded to the survey, and the first ICC is looking at the extent to which individuals at a site had the same responses. This is very similar to doing like a test of intra-rater reliability.

For these values, you can see, just looking at the table that the range of values is lower than the second type of ICC, and this is expected. We are not expecting to have extremely high values when we are looking at whether individuals have the exact same responses, because we expect some variability in how they respond to the survey.

What we are looking for is that there be a statistically significant value for this ICC(1). So that tells us that there is some degree of sharedness among how respondents are answering the survey. Again, this is particular to safety culture, because it is an organizational level construct, and we expect that there will be some degree of sharedness when people respond.

The ICC(2) is about the internal consistency of the mean score for the site. So this is really looking at -- It is similar, actually, to the Cronbach's Alpha where we want to see, if we took out one of those respondents, would we have the same mean score? If we added in a few more, would that mean score be affected?

So this is really important if we are going to have any faith in using that mean score to look at correlations. Is this a shared concept? Is this a score or something that seems to be stable? Again, we are looking for kind of a rule of thumb cutoff of .7, and this is a range of values, because we calculated these ICCs per site.

CHAIRMAN BLEY: Personal responsibility for safety on the range had the widest range.

DR. MORROW: Had the widest range, yes.

CHAIRMAN BLEY: So does that mean there is less certainty in how people think about it individually versus the collective?

DR. MORROW: Yes, I think that variable overall kind of acted differently from some of the other factors in the survey.

DR. KOVES: Communication acts a little funny for one reason. Prioritization -- This one is very interesting, because, basically, when you look at the scores, they are pretty much consistently high across all the stations, and the correlations don't do very well, because they don't vary across the stations.

So this is another one that is like, you know, this is behaving -- kind of interesting behavior by this one. So I am not sure if that helped any.

MEMBER RYAN; Not a lot. I am trying to figure out why the numbers are large, that the range is very tight, like number 1.

DR. MORROW: Well, this one was interesting, because this is showing the range of the values. So we have got the minimum and the maximum from all the sites. There are actually only a couple of sites who have those low values below .7.

MEMBER RYAN: I see. So you really have to dive into the full dataset to understand that.

DR. MORROW: Yes, and it kind of behaved weirdly, because for the most part, there wasn't a lot of variability on that factor, and that is what Ken was mentioning.

MEMBER RYAN; Okay. Thank you. That helps.

DR. MORROW: Again, I think I have talked about this a bit, but why is within-group reliability important? It is something that we don't necessarily look at with all surveys, but again what we are doing is taking these individual perceptions of safety culture and aggregating them up to the organizational level, because kind of the next step is to look at whether the organizational level score is related to organizational level safety outcomes.

So we are looking at safety performance metrics that are collected at the site level. We have to get those on the same level, so that we can make comparisons.

The next question that we were looking at is whether the survey factors that came out of this factor analysis show support for the Safety Culture Policy Statement traits. So are the factors identified similar to the traits included in the Safety Culture Policy Statement?

What we did for this was really just kind of a review of the factor labels and the items within each factor, and then compared those factors and items to the definitions of the policy statement traits.

This was really kind of a qualitative look at are we seeing the same themes coming out of the factors and the policy statement traits, and it was really important to dive down into those items, because you will recall these factor labels were also kind of just produced from looking at the items. So we wanted to make sure and go back and see are there items that seem to be similar to the definitions of the traits.

This is just a review of the traits that were included in the Safety Culture Policy Statement. They are: Leadership safety values and actions; problem identification and resolution; personal accountability; work processes; continuous learning; environment for raising concerns; effective safety communication; respectful work environment; and questioning attitude. Conveniently, there are nine of them.

Then this next slide is the crosswalk that we developed, kind of looking at how do the traits relate to the INPO survey factors. Of course, kind of the first observation that you can take from this crosswalk, just looking at it, is that for each trait there was at least one factor, in some cases part of a factor, from the survey that was related to that policy statement trait.

In some cases, it seemed to be pretty good one-to-one alignment. So there was a questioning attitude trait. There is a questioning attitude factor, and those items seem to support the definition of the questioning attitude traits. Same thing, environment for raising concerns/willingness to raise concerns. Personal accountability was similar to personal responsibility.

Then there were some cases where it seemed like the trait was supported by more than one factor. So leadership safety values and actions really seemed to have aspects of management responsibility, decision making, and supervisor responsibility.

Also cases where, because that management responsibility factor from the survey was so large, there were areas that were called out as traits in the policy statement that were supported by the survey, but by one of the subfactors under management responsibility. So work processes, for example, seemed to have elements of the subfactors, procedure communication and resources.

MEMBER RYAN: Why didn't safety communication have a management aspect?

DR. KOVES: Well, very often, like I said, if it didn't hang together, part of it went to management; part of it went to decision making, which is part of -- as you see in the policy statement, in the leadership category. But also there are other elements of communication, too: Peer to peer communication and even some of the items in there, as I recall, were even more like plant labeling and that type of thing that were part of the overall communication.

MEMBER RYAN: My question stands, though. Effective safety communication means everybody has the same view of safety, based on what they see, what they are trying to do, what they hear, what they observe --

to me. I like to say communication is about -- you know, evreybody is on the same safety page at the end of the day.

DR. KOVES: But you are talking about the results. I think the factor or the trait is more about the doing. It is about the communicating and not the result. What you are talking about is how I would measure the effectiveness of the communication.

MEMBER RYAN: Maybe that is right. Nonetheless, I would still think that there is a management aspect to safety.

DR. KOVES: And there is. We could take all of these, and we could roll them all up in the leadership, if we wanted. But that doesn't do us any good.

CHAIRMAN BLEY: I think what you were doing here, you are saying, given the Commission's policy statement, then effective safety communication as one of them. So it is a one-to-one match, even though you could have spread it up in different ways. I think that is what you are doing.

DR. MORROW: Right. The key point is that in the traits leadership was called out, and safety communication was called out, and they were different in the factors as well.

MEMBER RYAN: I got you. Thank you. I was thinking more about implementation and other aspects. Well said. Thank you.

DR. MORROW: Yes. All of these things are very interrelated, and I think the kind of global thing to keep in mind is that these are all aspects of safety culture. So overall, we have one big global construct.

MEMBER RYAN: Thank you.

CHAIRMAN BLEY: I am just wondering -- and I don't expect an answer right now, but if somebody wants to -- But as you begin to try to pull the policy statement into the ROP, these places where the factor analysis told us some reason things were clumped, I wonder if that will affect how you think about -- That is an ill formed question, but if you get the idea, management responsibility and decision making get distributed, and in the factor analysis they clumped kind of nicely.

I wonder if that has any practical implications to what we do with this stuff later on. I just don't have any idea.

DR. KOVES: I think probably the most -- kind of the biggest thing in terms of practical application was that in the factor analysis it really kind of put together PI&R and also organizational learning, and it is kind of like people were saying, listen to us.

You know, if you are learning, you are going to implement it. You are going to change. You are going to improve. And if you are identifying issues, you are learning.

CHAIRMAN BLEY: That was a connection you found in the data for the analysis, yes.

DR. KOVES; Yes, that was the connection that found in the data. To me, I think that is the one biggest thing, if I understand what you are talking about.

CHAIRMAN BLEY: You must, because it makes sense to me. Go ahead, Stephanie. I think you do better than I do.

DR. MORROW: Okay. So just some key observations from looking at kind of the crosswalk we developed, and overall looking at the factor analysis of the survey, we saw that management responsibility/commitment to safety accounted for most items and the most variance of the survey results. It was also similar to multiple straits in the policy statement.

An interesting thing that came out was

that supervisor responsibility for safety was a separate factor in the INPO survey, and it really seemed to pull out kind of the idea that a supervisor's commitment to safety was slightly different from the overall management level.

That is kind of rolled into leadership safety values and actions, but it is something to kind of keep in mind as we think about safety culture, because there are -- when we look at employee perceptions, they differ whether they are thinking about their immediate supervisor, who they have the most interaction with, versus kind of the overall organization and kind of management as this nebulous being.

Also decision making came out in the factor analysis with the survey, and it seemed to have elements of the leadership trait and problem identification and resolution.

Kind of in summary, I am going back to where I started, and it seems like good timing as well. The design of the construct validation study was appropriate for the research questions, and we did see evidence of reliability and validity in the study.

Then there were also many common themes between the factors that emerged from the survey and the traits in the Safety Culture Policy Statement. We didn't see that one-to-one alignment, but it gives us confidence to say that the survey supports the traits and that there were these in-common themes, and that the traits were represented by at least one of the factors.

CHAIRMAN BLEY: Okay. Committee, anything? Thank you.

MEMBER SCHULTZ: I had one question on the listing that you had on slide 35. Ken alluded to this before. I may have missed it in your presentation, but prioritizing safety did not show up under the INPO survey factors that match up with the --

DR. MORROW: That is correct.

MEMBER SCHULTZ: That is not on the list?

DR. MORROW: Actually, I apologize for not mentioning that, but I ended up leaving it off this crosswalk. It seemed like it could fit under the leadership trait or under personal accountability. There were some aspects to it, but when you look at those items -- again, Ken mentioned that it kind of behaved weirdly and that they were the first few items in the survey, but the content of the items were also kind of tapping into more global perceptions.

So it was kind of like, overall, does it seem like my organization supports safety. That may be getting into something different. It may be getting -- or something kind of more global, that at the higher level the safety culture construct as a whole rather than a specific factor within safety culture.

So it is not specific to management. It is not specific to decision making or communication. It is just kind of like what is my global perception of safety culture here. So it didn't fit very well. I will say that.

MEMBER SCHULTZ: I am just trying to get a personal feeling for it, but it all appeared at the beginning. It is certainly, within the industry, a mantra at the site. Safety is a priority. Safety is a priority. So does that have an effect?

DR. KOVES: Well, that was our question: Was there order effects? I think that safety conscious work environment and that idea is very highlighted in the industry. So I think that that is one reason that it has come out as a factor, but the items that Steph is talking about - - Nuclear safety is routinely emphasized as a priority at meetings; The station ensures the contractors/vendors understand our expectations to performing work; I am always informed of current safety concerns or issues that affect my job; Peers coach each other on behaviors that promote nuclear safety; I have an influence on decisions involving nuclear safety that elate to my job; and At this station nuclear safety takes priority over production goals --

The items weren't really as nice and tight and as pointed as you saw in some of the other factors, and like she says, it is kind of they are more global, plus as you point out, it is kind of the mantra, plus they happen to be the first six items. So that is why this is -- You know, we see it as kind of a problematic factor that we are going to see in the future.

We did keep -- As you see, we kept three items that represented that to see how it behaves in the future, but we will see.

MEMBER SCHULTZ: Thank you.

DR. MORROW: Thank you.

CHAIRMAN BLEY: Okay, thanks very much. We will recess for lunch, and start at quarter after one.

(Whereupon, the foregoing matter went off the record at 12:15 p.m.)

- - -

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

Time: 1:15 p.m.

CHAIRMAN BLEY: We are back on the record. It is your turn again.

DR. KOVES: Why not? I've got good news for you. A whole bunch of these slides, they have already covered. So we are going to kind of go over them modestly -- at least the beginning, modestly quickly.

What is the central question of this part of the validation? It is does the measure actually measure what it purports to measure, it says it measures, and in this situation are the results related to other measures of safety?

We have talked a little bit about reliability. We have mentioned the words a lot of times, a number of times. What is reliability? It is the consistency of a measure. Does the measure produce consistent results under consistent conditions.

Whereas, validity is about the accuracy of the measure. You know, does it really measure what it says it will?

I have got an illustration that I have been waiting about 10 or 15 years to use. So I am going to go ahead and use it, because this is my first opportunity since I have thought of it.

Anyway, in terms of reliability and validity, I have a scale at home. It is an old spring scale. This thing is very reliable. I mean, if I gained a couple of pounds, you can see it; if I lose a couple of pounds, you see it right on there.

Unfortunately, it is not valid, because it under -- The weight that it gives you is about 10 or 15 pounds less than you actually weigh. So that is an example of reliability without validity. Anyway, thank you for indulging me with that. I appreciate it.

Stephanie already talked about criterion validity. Is the measure related to a criterion in the real world, concurrent and predictive? The concurrent is how well do the results relate to the current criteria, and predictive is how well do the results relate to future results of the criteria?

Correlation: I put this in here just in case that there were some people in the audience who aren't familiar with correlation. I will go over it very quickly, assuming that the committee is very familiar with it.

It is the degree to which two or more

variables show a tendency to vary together. The most familiar measure you know as the Pearson correlation coefficient obtained by dividing the covariants of the two variables by the product of the standard deviations.

Examples of correlations --

CHAIRMAN BLEY: That is a general statement, and before you leave it, there is just one thing that I want to whine at you.

DR. KOVES: Would you like some cheese with that?

CHAIRMAN BLEY: When you use a -- you apply that to a model, be it a regression model or some other model, the two things you talked about, which were squares, are the -- you are essentially measuring how far you come from the predicted model, how far the data come from that, and the square of your correlation -- I just want to read this, because you know this, and you say this in the report, or you say this in the report -- Was it a joint report? It's their report.

You say this in the report. But I am going to read this, just because I want to make the emphasis on it. Whereas, the R-square, which is at least in the regression called the coefficient of deviation, indicates proportional reduction in the variability attained by the use of the information from the X value, as he says, the square root, the correlation, does not have any such clearcut operational meaning.

Nevertheless, there is a tendency to use this in most all reports, but R, the correlation's coefficient, may give the impression of a closer relationship between X and Y, if you are just talking two variables, than does the corresponding R-squared. So if you have a .3 correlation coefficient, you are only explaining .3 squared, 10 percent of the data.

DR. KOVES: Right.

CHAIRMAN BLEY: I am afraid, for most people seeing the report and seeing the correlation coefficients, they don't get that. They get -- they read it, they are getting much more confidence in how much is being explained by the model than is really there.

I think we are putting it in terms that kind of trick people. Back to the stuff I talked about a long time ago, and you have been emphasizing social sciences this time, the .9 -- I was talking about something quite different. Now I have thought more about it since then.

In seismic, if there is correlation in the time histories of the parameters, like acceleration and frequency, that correlation has to be nearly perfect before the two things that are being affected act in a dependent fashion. Otherwise, even if there is reasonably high correlation but not nearly perfect, they actually respond as if they were independent.

That wasn't good analogy that I brought up at the last Subcommittee meeting, but still, what was bothering me is this idea that .3 and .5, which according to this one fellow you quote, is considered medium or high correlation really is only explaining 10 percent or 25 percent of the data, which means the model isn't telling you everything you want to know, and especially if you are down at the low end, .2, .3, it is not explaining very much of the data.

I am jumping the gun just a little, because I want you to think about this along the way. I think part of the problem and part of the reason we see such low correlations, which you are going to get to -- to me, are low -- they aren't explaining much of the data, and because the things we have picked, the unplanned scrams, safety system actuations, forced outage hours and equipment outages aren't highly correlated to bad safety events, things that approach core melt kind of problems.

DR. BARNES: True.

CHAIRMAN BLEY: In things like that, which some real bad events were, people have made errors in the plant that has come close. They haven't gone to melt, but they have still been pretty severe events. There, I think we see these factors very strongly involved.

I am just thinking, if somehow you could sometime along the way look at things that are closer to real events that affect the safety of the public, I really believe we would see much stronger correlations, and you would have a stronger case.

What has bothered me is I see it. When I look at the events, I see that. When I look at the analysis, I see very low correlation, in my opinion, correlations that aren't explaining much of the data, and I think we are looking at the wrong things to make the case.

The reason we want to do this is to really protect the public, and those kind of things aren't real high on the list. If you look at risk assessments, those kind of events don't contribute very much at all. They aren't the big actors.

I think, if you were looking at some big actors -- we don't have a lot of them. You don't have as much data to look at.

DR. KOVES: That is a problem.

CHAIRMAN BLEY: On the other hand, it is what is important, and the correlations, I think, are going to be strong as can be in there. There is a lot of significant events around, and not hundreds.

DR. KOVES: Your correlations would be higher, but actually it will be harder to get statistical significance, because the ends are not reaching one.

CHAIRMAN BLEY: Maybe you can change the kind of statistics you are using where you speak in terms of the probability as being important, rather than what you do to look at large samples.

DR. KOVES: Yes. Let me --

CHAIRMAN BLEY: Those sorts of things perhaps would get you there, because I think it is going to be strong.

DR. BARNES: Near misses. Like I was saying earlier, there's a lot of different ways to measure safety performance, and you know, safety is a construct, just the same way that safety culture is, and there is lots of different ways to measure it.

CHAIRMAN BLEY: Yes. But what is driving

us to have a safety culture policy is trying to protect the public from bad things, not -- I mean, we do want to protect the workers from minor injuries and even significant injuries, but really, what is driving the agency to be concerned and want to force the people that regulate to be concerned is worrying about the things that could harm significant numbers of people.

I just -- I think there's probably many ways to go at trying to link this up. I just don't think this case is all that convincing. I mean, it is there, but it is -- You explained a little bit of the data. Yes, that is good. You can, but it is hard to justify going out and having people do lots and lots of things to putting up a little bit of the data that is not affecting the real safety issues.

Now it probably does affect them, but we can't see it does. Anyway, that is my concern. My other concern was that one about I think people don't know the correlation coefficient, and they believe they seem stronger.

DR. KOVES: Well, yes, but it --

CHAIRMAN BLEY: Stronger association than is really in the results.

DR. KOVES: Yes. The variance explained,

it was only the square root.

CHAIRMAN BLEY: Right. I'm sorry, I'm done. You had something?

MEMBER STETKAR: I was going to give my example, but it's just as well.

CHAIRMAN BLEY: Go ahead, Ken.

DR. KOVES: All right. By the way, if you come up with that study and you want, let me know, will you? I would be more than happy to try and pursue it

CHAIRMAN BLEY: Well, if I am not doing this anymore, I might talk to you about that.

DR. KOVES: All right. Correlations run from positive one to negative one, you know, perfect correlation. Basically, if you have it, it is a straight line, diagonal line. Anyway, what you are seeing in the upper lefthand corner is data with a correlation of .9, positive correlation. Upper right -- excuse me, that was lefthand corner.

Upper righthand corner is a negative correlation of .9, and then in the lower left you are seeing zero correlation, and on the right you see a correlation of about .4. Basically, what that is saying in general as one is moving up, the other variables are also going to be increasing, too, but not in a stepwise fashion.

In terms of the application to the safety culture data and most of the variables that you are looking at, what we do is we have a high score on a safety culture survey, and then we have these other safety indicators that are usually good, if they are low. We want very few of these to happen. So we end up with a negative correlation. Negative correlation of about negative one there, and a correlation of about negative .4, and looks something like that.

We've beat on this one quite a bit already. So I am not going to spend time there.

Here is actually an example, and you will see -- I mean, we really threw this up. This is not confirmatory but very -- not experimental -- exploratory. Thank you. So we threw the survey in on the survey results up against a whole lot of variables, and you will be seeing more of them during the closed session.

These were some of the ones that came out with the best correlations, and this is an example. So you see on the lefthand side we have got nuclear power plants, and you see that they are rank ordered by where they fell relative to the survey score, and then you are seeing -- and this is all actual data --

a number of unplanned scrams, unplanned automatic scrams, emergency power unavailability, personnel safety index, the chemistry index, and human performance error rate.

CHAIRMAN BLEY: I have forgotten. Remind me how you calculate the mean survey score for the power plant.

DR. KOVES: The mean? It is the mean of all the individuals and all the -- I mean, it is the mean of the means for all the factors.

CHAIRMAN BLEY: And the mean -- you are taking this out of the correlations of the score itself. Score goes zero, one, two, three, four up to the seven.

DR. KOVES: Yes. So you are seeing plant one for the entire plant, their mean score on the safety culture survey was 6.12.

CHAIRMAN BLEY: And seven was "agree very strongly this is good," something like that?

DR. KOVES: Yes, correct. And actually, what I did was I condensed them all. So the bottom of the range was 5.52. This is an example of what real data looks like. So you end up with correlations of about .3, very close to that.

MEMBER STETKAR: Average -- Then an

average mean would be three and a half?

DR. MORROW: Four. It is one to seven.

MEMBER STETKAR: It is one to seven.

DR. KOVES: One to seven.

DR. MORROW: So four is the --

MEMBER STETKAR: So everybody is a little above average?

DR. KOVES: Well, but that is --

MEMBER STETKAR: We are still looking for the one that keeps the average, average.

DR. KOVES: Virtually, every survey that was administered and looked at, the distribution is skewed, and it is skewed to the positive, and this data is really no different than that. So, yeah, you are going to have --

CHAIRMAN BLEY: Is there something wrong with that? The things we are looking at in the survey are things that plants have been working on for years.

MEMBER STETKAR: That is true, but there is still some average. I mean, you are looking at objective numerical things you can count.

CHAIRMAN BLEY: Yes, but if you are at the worst power plant, and you evaluate your communications, you are probably not going to say they are lousy. You know, if they are bad, they are probably down in the middle of that scale, and I forget what the words are associated with that.

DR. MORROW: From "strongly disagree" to "strongly agree." And closest to the midpoint is like "somewhat disagree," "somewhat agree." I think that is what the terminology is.

DR. KOVES: And then neutral is in the middle.

DR. MORROW: You do see people like, ah, somewhat disagree, sometimes this doesn't happen, and that might be effectually a red flag indicator, like ah, this isn't happening. So at that point, they are just using the three on the scale.

CHAIRMAN BLEY: Sorry, I won't chip in.

DR. KOVES: That's okay. So anyway, this is just an example of the -- you see this in live data. You will see here that on this first row we've got mean score for the nuclear power plant, and you see once again the same correlations we had on the previous page except we have added the ROP in here, and this is where they stand on the ROP -- or they were in the ROP at the time. But also you see some of the other factors, and them correlated to these.

You are seeing more variance here in terms of how some of the factors relate better to some safety indicators than others, and then also in the brackets what you are seeing are some of the subfactors and those correlations with those particular indicators, and you will see a lot of that here in just a minute.

Just some select ones, but once again showing how some of the factors and subfactors can do a better job of zeroing in on certain things than others.

CHAIRMAN BLEY: What is the score on the ROP?

DR. KOVES: Well, they are in column one, two, three or four.

CHAIRMAN BLEY: Okay. I am still not sure how you calculated.

DR. KOVES: No, it is just what -- It is what the NRC reports out.

CHAIRMAN BLEY: They report out -.26?

DR. MORROW: That's the correlation.

DR. KOVES: No, that is the correlation. That is not the score. So what that is, is we have that -- Going back to the previous page, rather than here in unplanned critical scrams for this particular plant, there would be what column they are in.

CHAIRMAN BLEY: One, two or three?

DR. KOVES: Yes, one, two or three, and then going on down, you would see that. So that is where the correlation comes from with the ROP, is where that station is with the safety culture and/or relative to the ROP. Then that gives you that correlation of -- what is it, .26, you said?

CHAIRMAN BLEY: Yes.

DR. KOVES: All right. Moving on, so this is what I use as a sample of concurrent. We will dive into more detail here in just a minute, and then we also tried to do the predictive validation.

One would imagine or one would assume that, even if a survey or this construct is related to current performance, part of the theory around safety culture is that it impacts plant performance and drives some of that. And if that is the case, then you would expect that there might be -- The correlations might at least stay the same, if not increase in the future.

Whereas, typically if the culture or the survey is only representing what is going on at that given time, then over time you would see the correlations decrease.

So what we did was we also then ran -- Because it was a year later, we pulled the results again for all the stations and did the correlations, and most of the correlations either did what you would expect, and that is stay the same, or decrease. However, we did have a few that actually increased.

One was forced loss rate, as you see, from 2010 to 2011, and these were the correlations with the mean score that increased and then also with some of the factors. That is industrial safety accident rate for 2010 compared to 2011, and then also a total -- you expect these to be fairly correlated -- total industrial accident rate.

What was interesting that I didn't mention earlier and was a little disappointing when we did the concurrent validation, was the lack of relationship between industrial safety accidents and safety culture, and which I was like, boy, that seems very odd to me.

Well, what is interesting is that now you really see a much stronger relationship between them after one year than we did in the concurrent information.

MEMBER STETKAR: After one year, what is the normal variability in the parameters that you are measuring there? You get to look at 30 year and see what a trend is, for example?

DR. KOVES: Well, I am going to get to a problem, a fundamental problem, with these results in one second, but I think we will address that.

MEMBER STETKAR: I guess, just looking at two snapshots in history and trying to draw conclusions from those two snapshots -- people have done that an awful lot with a lot of other things like losses of offsite power, and tried to draw conclusions out every year better and better and better until you get a worse year, and then it is, oh, my god, we have to redo our statistics.

So looking at two individual snapshots out of the universe is something you put on a slide, but it doesn't seem to make much sense.

DR. KOVES: Well, and there is another confound in this data and with all of these, is that these numbers of 2010 and 2011 for these here are not totally independent, because what INPO does is, for a lot of these indicators, it will take a score and average generally between outage cycles. So they go from outage to outage.

So you cannot say that the 2010 and 2011 scores are --

MEMBER STETKAR: Are really a calendar year ago.

DR. KOVES: -- are really totally independent exactly. Now some of the other data -- Some of the data is separate, but this is not. So that kind of throws a little confound into the whole analysis piece.

We will close the session here and get to the details in a second, the "what does it all mean?" You know, safety culture appears to be solidly related to other measures of safety at the plant, and then also it may be a predictor of some indicators of safety and safety culture, but we obviously need more research into the clarifying of things about this.

Shall we go into the private session?

CHAIRMAN BLEY: I guess I don't need the hammer for that one. We are now in closed session.

(Whereupon, the foregoing matter went off the record at 1:39 p.m. and went back on the record at 2:40 p.m.)

CHAIRMAN BLEY: We are back in session. Welcome back. I think you are up.

DR. MORROW: Yes. So our final presentation of the day is looking at similar information from what Ken just presented, we are looking at the results of the INPO survey with NRC performance metrics.

So just some of our key findings from this analysis: Like Ken just said, based on accepted standards in the social sciences from other studies, they do measure safety performance.

CHAIRMAN BLEY: I know where this came from, but we would communicate better, I think, if instead of trust us, we are social scientists, we explain -- I know that is what it means to you, but it probably doesn't to some other people.

DR. MORROW: Yes, and actually I have a slide that shows the correlation.

So we saw moderate, statistically significant correlations between the safety culture survey results and some of the NRC performance metrics that we looked at. We also saw some moderate correlation between the survey result and performance metrics that were measured one year after the survey was administered. These were similar to what Ken found with the INPO data. These are kind of more of the broad based performance metrics.

So we used, which was a variable, looking at whether a plant was in an elevated oversight position in the ROP Action Matrix, so basically anything other than the baseline column 1, and also counts of allegations from the licensee personnel.

We also, again, just want to clarify our acknowledged limitations of this research. One, this study was a single study, of course. So these results would need to be replicated to show reproducibility.

We have talked about this is a snapshot in time, and we are looking at a snapshot of performance metrics. So this can kind of give us some information that can be used for more focused exploration in the future, but it is still just a single study.

We also were just looking at correlations. So you can't establish a causal relationship with just a correlation. All this says is that these survey results were in some way related to or associated with the performance metrics that we looked at.

Also this study in particular was cross-sectional, looked at a high level across the entire nuclear industry. It only made comparisons between sites rather than with a single site over time. So to do more research where we look at what is the safety culture score this year, what is the performance, what is the safety culture score the following year, kind of looking at within a site, would provide us with some different but also valuable information.

I will mention, because this site was particularly kind of related to the Safety Culture

Policy Statement, keep in mind this is only nuclear power plants. The Policy Statement was written to apply to all of the NRC's regulated communities.

CHAIRMAN BLEY: I guess the only thing that bothers me -- you briefly talked about it in the report -- is although, as you say, you can't establish causal effects, we are introducing it to begin to justify putting requirements on people to do things in accordance with this.

DR. BARNES: What requirements?

CHAIRMAN BLEY: Now that we have a Safety Culture Policy Statement, it is hard to imagine we won't have requirements to go along with that sometime.

DR. BARNES: Will we want to address that?

CHAIRMAN BLEY: We could.

MR. SOLORIO: Currently, there is no move afoot to take the Safety Culture Policy Statement and turn it into a rule, for example. Right now, we are waiting for various industries to work with it. I think the nuclear power industry is showing very strong initiative for doing so, but in the other areas that we regulate, the materials area, there is actually quite a lot of -- the best word to use -- passion for trying to implement it.

We have been working very closely with the agreement states, and they are trying to get the word out the best they can. We have actually taken our Safety Culture Policy Statement brochure, and they have actually used the electronic copy and modified it for the agreement states, and in many cases their own document.

So, really, we are in a situation now, I think, where we want to give it some time to see what it can do based on it being a policy statement. So I wouldn't think -- Well, right now there is clearly no move afoot to take what the industry has done and use it to justify a rulemaking.

MEMBER RYAN: A quick follow-up. The agreement states are aiming at, I'm sure, the larger licensees.

MR. SOLORIO: They are aiming at all licensees.

MEMBER RYAN: All licensees?

MR. SOLORIO: All licensees. We were just in uranium -- excuse me -- in Colorado last week -- Maria was. She is sitting over there, Maria Schwartz -- presenting to uranium miners -- what do you call them, uranium --

MS. SCHWARTZ: Uranium Mining Association.

We have like a workshop every year.

MEMBER STETKAR: Identify yourself, so you are on the record.

CHAIRMAN BLEY: Use your microphone.

MS. SCHWARTZ: Maria Schwartz, and I work in the Office of Enforcement, and I made a presentation on safety culture at the Uranium Mining Workshop, Recovery Workshop, sponsored jointly by the NRC and the National Mining Association.

They received it very well. I got a lot of compliments on the presentation. They actually said they were very interested, and there wasn't enough discussion about safety. They were more focus on regulations and requirements, but this sort of larger view toward safety, safety first focus, really seemed an important part of their activities. So they were very interested.

MR. SOLORIO: This week, Josie Piccone is in Glasgow, which is in England -- Scotland -- if I ever get to Europe.

MR. WIDMAYER: Not now.

MR. SOLORIO: IRPA.

MEMBER RYAN: Yes, International Radiation Protection Association.

MR. SOLORIO: There you go. They have a

-- It is a world conference going on right now, and she is speaking about the Safety Culture Policy Statement.

MEMBER RYAN: That is two weeks from now.

MR. SOLORIO: And one of the things she is going to speak on is actually an IRPA document that was just produced, I want to say, in the last few months by a committee. Basically, it is called Radiation Protection Safety Culture, kind of the name of the document. Basically, IRPA is working on inculcating their membership with safety culture.

The document actually contains a table with the Safety Culture Policy Statement traits. So it is being --

MEMBER RYAN: For those who don't know, I might just add real quickly, IRPA is the health physics societies from all of the world countries in one international organization for radiation protection.

MR. SOLORIO: Right. So I would say a lot of -- I could name others. I could give you a table after the meeting that kind of shows you all the different places we have outreached in the materials area, and continue to do that.

Cindy Flannery was here in the morning,

she is from the Office FSME. She is not here now. There's other outreaches I think she is doing this month.

In the materials area, also the upcoming conference that NMSS has annually. It is called the Annual Fuels meeting they have every year.

MEMBER RYAN: Actually, I think it would be helpful we could get that list from you. That would be very helpful.

MR. SOLORIO: Yes. So we are on the agenda for that. So there's a lot of -- I would say all our licensee organizations are trying to get it into their meetings. We want to see how that works first, and then if there is a lead, then we might consider down the road, but right now I think we are trying to get everyone to be educated about it and start to use, and they are showing us they are using it.

MEMBER RYAN: And I guess -- I don't want to put words in your mouth, but it sounds like you have an ongoing plan that you are going to be following, and learning, and then maybe adjusting as necessary, that kind of thing.

MR. SOLORIO: Right. Well, under the Commission's direction in the SRM they wrote for the

SECY that we put the Safety Culture Policy Statement in last year, they directed the staff to do education outreach and provide support to the various regulated communities to help them implement this Policy Statement. So that is what we are doing.

CHAIRMAN BLEY: Thank you.

MEMBER STETKAR: Let me follow up a little bit, though. Dennis asked the question immediately reactive with the word rulemaking. I think what I have been hearing today, though, it is not clear to me when or how some of these notions may be integrated into the reactor oversight process, which certainly is not rulemaking, but it certainly has an effect on how people are rated and how people do business.

So although it might not be rulemaking, it may very well affect how people are evaluated, which is the same thing, in my mind. So back to Dennis' question that, if we can't establish causal effects, if there is -- This is a nice research effort -- how does that affect our interactions with licensees from a regulatory -- a reactor oversight process, from a regulatory oversight process?

MS. SHOUP: Certainly. Thank you. This is Undine Shoup again.

If you look -- Actually, the Common

Language Initiative was started before the Policy Statement even came into being, actually. We deferred the Common Language Initiative until after the Policy Statement was complete, because re recognized that the Policy Statement had the potential to change the Common Language Initiative.

So that is actually an initiative that industry came and asked us to do well before the Policy Statement. So how we see this having an impact is that, as Maria had alluded to earlier, after we develop common language, and common language will include all the common language for power reactors -- I should say it is not being developed for all licensees, just power reactors. After we develop the common language, then we will be able to evaluate how to best incorporate it into the ROP, and I think Rani has more on that.

MS. FRANOVICH: Just to add to the response on the slippery slope question, after Davis-Besse, the GAO made it very clear that the NRC needs to look at safety culture, and I think it was in 2006 the Commission directed the staff to do that using cross-cutting areas of the ROP.

We also have safety culture assessments that we can do with the 95-002 and 95-003 applications. So I don't see the slippery slope happening as a function of this work. I think we are already there, and it is a matter of getting alignment between the policy statement, INPO terminology, and ROP terminology and how we do the ROP implementation, as I had mentioned, is really a future activity that we haven't started yet, but we are getting very close with the common language framing up.

CHAIRMAN BLEY: Interesting. Thanks.

MR. WIDMAYER: Well, frankly, your question kind of goes to the additional comments that the other members made to the letter: Are these really the right traits to be looking at? So the question sort of still stands.

CHAIRMAN BLEY: Thank you for pointing that out, and that is true.

DR. BARNES: Yes, thanks a lot, Derek.

MR. WIDMAYER: We don't write user newsletters.

CHAIRMAN BLEY: If we had a project.

MR. CAMPBELL: Just to jump -- Andy Campbell. I forgot. Sorry, sir. I am the Acting Director today for OE.

It is important to keep in mind, though, that when you say are these the right traits that the basis and how these were all derived isn't just that there is one set of traits that are absolutely perfect for all circumstances, but that these traits were developed through a process that got large unity of our licensees to buy into, and that they represent and are tied back to the history of the different aspects of safety culture characteristics.

So although there could always be different traits, there could always be additional traits, there could be less traits, but these, I think, represent the collective efforts of a fairly large cross-section of our community, and then there are things like the ROP to make them to be appropriate to what we are trying to do.

CHAIRMAN BLEY: Thanks. The origin of the question was, though, not that process, which we all appreciate, but was, now that we are there, will we actually improve safety if we push ahead in this area. It smells good, but do we have any real evidence that it will improve safety.

DR. BARNES: Correlational study can't answer that question, and that is why at the beginning of this I was talking about some of the interventional studies where they have actually used results like these to go in and identify what kind of interventions need to be made in a particular work group, in a particular organization, to address real problems.

Something like this, like a survey like that, it is a screening tool. Helps you identify where the problems are. In terms of whether there is value in doing those kinds of interventions in nuclear industry, that is what we see with corrective action plans, and at least we walk away from some of the plants where we have seen problems with greater confidence that, because of the interventions that they have made, that they have solved the real problem.

Now that is Ken's business. So, you know, other than developing survey tools, Ken spends the majority of his time working with sites on solving problems related to organizational effectiveness and safety culture, and he is still employed. So I think maybe he is achieving something. I don't know, Ken.

DR. KOVES: I would say most of my time, some of my time. That is one of the things that is most fun. It is actually making a difference, and I think at one station -- I was just talking with Steve about it -- recently, in certain areas they have made a real turnaround, and their goal actually was culture change.

They were a good plant before, and now in a number of these indicators, number of these areas, they have really stuffed it up and made a difference.

CHAIRMAN BLEY: Okay. I think we better forget ahead. We will be here until midnight. Go ahead, Stephanie.

DR. MORROW: All right. This is just kind of a breakdown of the performance metrics that we used when we are looking at correlations between the survey and various performance metrics that the NRC maintains.

What we tried to do was to kind of get a broad spectrum of sources when we looked at performance metrics, and also pick out ones that had some variability. So we are not going to see any correlations, if all of the plants are always at zero. So that was kind of our approach when we looked at these.

The next slide is just kind of another view of those performance metrics, looking at them from where they come from. So we have some performance indicators in there. We have data that comes from inspection reports. One is just kind of total count of inspection findings, but also looking at the cross-cutting aspects that are part of the ROP.

We broke that down into just looking at the total number of cross-cutting aspects that were tagged to inspection findings, but also where those aspects kind of sat within the framework. The aspects are within components which are within cross-cutting areas.

So you see there are two of the three cross-cutting areas of the ROP here, human performance, problem identification and resolution. There is a third one, safety-conscious work environment. That was not included in this analysis, because those aspects are not used very frequently. So there wasn't a lot of variability to that data.

We also looked at overall performance assessments in terms of total number of substantive cross-cutting issues, and this is for end of the year 2010, and then also end of year 2011; and as I mentioned earlier, the site's placement o the action matrix.

So we looked at -- Basically, we just divided this into two, whether they were in column or in another column. So the other column would be in an elevated oversight role. They are not in that baseline condition of the action matrix.

We also looked at allegations from

licensee personnel to the NRC. We looked at total counts and then also an exploratory variable that was more specifically related to allegations that were categorized as relating to the organization's safety conscious work environment.

Just another background slide of correlation. I think we've all got it now. Some of these are a bit duplicative. One thing is that the Pearson's correlation is sensitive to outliers.

So the additional analyses that we did were to also look at the Kendall tau correlations, which is a non-parametric test that basically just ranks the data. So what is your rank on the safety culture survey compared to all of the other sites, and what is your rank on this particular performance metric.

We saw the same patterns of results. That does kind of take away some of the variability in the data. So where we can, we report the Pearson's correlations.

The other thing about being sensitive to outliers is that makes it that much more important to look at the scatter plots. I think that was brought up the last time we talked about this in ACRS. So we do have some scatter plots to show you of the relationships between the survey results and the performance metrics.

Here is the percent variance that Dennis mentioned. Again, the effect size rules of thumb that we are using are a correlation of .10 is about a small effect size, and that is only accounting for about one percent of the variance. So it relates to how overlapping these variables are.

When we look at meta-analyses of previous safety culture research, these are large studies that capture all of the correlational results from individual independent safety culture studies and say what was the overall effect look like.

Those studies have found the correlation between safety culture surveys and accident and injury rates to be about a medium effect size, between -.22 to -.39. This actually goes back to what Val was talking about earlier, that when we look at, more specifically, individual attitudes and individual behaviors and look at just the survey results against self-reported safety behaviors, we do see larger effects.

That just illustrates that kind of when you look at something that is more focused, you are going to see larger effects. When we are talking about organizational level variables, those effects are smaller, because there are so many other intervening variables' moderators that can affect that relationship.

So to put it into variance terms, what we are talking about is these meta-analysis studies have shown that there is a four percent to a 37 percent degree of sharedness between safety culture and safety performance, just a very high level, and it really depends on how safety culture and safety performance are both defined and measured.

Now we get to the good stuff, the correlations. Just to orient you to the table, again the safety culture safety is listed in the first column going down. What we have here is the performance indicators also calculated oversight on the action matrix, total inspection findings and total substantive cross-cutting issues.

The cells that are highlighted are the statistically significant correlations, and you can kind of see the patterns of results, just looking at the shaded cells.

Some that I want to kind of just note is, if you look just down the first column, unplanned scrams, and see the overall safety culture survey, and most of the factors were negatively related to unplanned scrams. Same for total inspection findings.

One of the largest correlations on this table is between questioning attitude and total inspection findings -.41.

We don't have questions here.

This is a scatterplot showing the overall relationship between the safety culture survey score and unplanned scrams. So we are looking at this. Does there appear to be any kind of outliers that stand out that might be affecting this correlation.

There is one point that is a little farther out, but if you look at the scale, we are talking about from zero to five unplanned scrams. So there's quite a few sites that are clustered around zero, but there is also a slight trend that you can see, kind of follows that best fitting line.

CHAIRMAN BLEY: What you get then is how much is accounted? Did you try taking that last point out and seeing what it does to your--

DR. MORROW: Not in this situation.

CHAIRMAN BLEY: Because that is a nice technique in this kind of analysis, and it is everything disappears.

DR. BARNES: She did it in others.

DR. MORROW: Yes. when there were outliers, when they related in the outliers, we took it out to see if that effect held, and I will actually show you --

CHAIRMAN BLEY: Just staring at this one, you can't tell, but I wouldn't be surprised if it flattened ut a whole lot.

DR. KOVES: I would say it would flatten out a little bit, but if you look -- to me, looking at it, you are looking at the bottom there where you've got the one, two and three.

CHAIRMAN BLEY: It is a nice thing to do. Go ahead.

DR. MORROW: These are additional concurrent correlations looking at some of the survey results, the allegations, total count of ROP aspects tagged to findings, and then also the kind of different subcategories.

CHAIRMAN BLEY: Is this one in the report?

DR. MORROW: Yes, it corresponds to Table 8 in the paper.

CHAIRMAN BLEY: Oh, there it is. I skipped a page. Thank you.

DR. MORROW: Again, just some of the interesting patterns that we saw on this was total ROP aspects was correlated with the safety culture survey overall, and many of the factors. You will notice that willingness to raise concerns and personal responsibility for safety were not correlated with any of the variables on this table; and again, we talked about personal responsibility for safety. That really just did not seem to be related to any of the metrics that we looked at.

CHAIRMAN BLEY: Then you talked about why that is.

DR. MORROW: Right. Again, this is a scatterplot looking at the safety culture overall and total ROP aspects in 2010. Again, we do have some clustering. What I found interesting looking at this scatterplot is that kind of the top performers on the safety culture survey were all within under about 15 aspects tagged to findings, and then you've got quite a few that follow the line of best fit.

CHAIRMAN BLEY: Since you brought up Jim Reason, there is one thing he talks about in organizational accidents. I guess it refers to the Swiss cheese idea.

That is that he is trying to deal with people from, say, Dow Chemical and other places that really put a big focus on lowering the incident rate of industrial type accidents, minor accidents; and some argue that that makes you less susceptible to large accidents.

At this point, I kind of agree with it. I don't think there is anything to strongly support it except good intuition and studying some bad accidents. If you are bad on those everyday kind of things, it is real hard to be good on the serious risks. If you are good on those things, it is no assurance that you will be good in the big ones such as the plant John was talking about earlier.

You almost kind of see that,a because on this one you see there's a lot of people with low scores here, and they are not sticking out, but don't have anything related to the bad, really serious sorts of events here, at least not clearly. There might be some of the ROP aspects that get there. Go ahead.

DR. MORROW: Well, the alternative to that is also that those who do have good safety cultures -- how do we show that they didn't have an accident, when it didn't happen? How do we show the nonexistence of an accident?

CHAIRMAN BLEY: The bad accidents, we know.

DR. MORROW: Yes. But you can't look at the opposite.

CHAIRMAN BLEY: Well, that's true.

DR. MORROW: This is a situation where we look at the correlation here and the correlation was -.28. So we are talking about seven, eight percent of the variance. But when we look at the scatterplot, this is where the correlations can be misleading if you don't look at the actual data and go into the data.

So when we took away this outlier, there was not a significant correlation. So this was --

CHAIRMAN BLEY: It just went away.

DR. MORROW: Yes. Safety culture overall, and the allegations that were specific to safety conscious work environment, which sounded promising, but at least looking at this snapshot of data, we can't really establish it according to that we can be confident will be repeatable.

Our key observations looking at those concurrent correlations are that the overall safety culture survey was moderately correlated with unplanned scrams, total inspection findings, and also total ROP aspects. Specifically, it seemed to center around that problem identification and resolution cross-cutting area.

We found,similar to what the INPO correlations looked like. Questioning attitude seemed to kind of stand out as we saw higher correlations when we just looked at that questioning attitude factor compared to all the other factors. So that questioning attitude was correlated with total inspection findings, total substantive cross-cutting issues, allegations from licensee personnel. We know that SCWE related allegations, one, because we know there is a terrible outlier there, total ROP aspects, and also the human performance area and problem identification and resolution area, specifically.

We also saw -- and training quality is one we haven't talked about too much, but training quality was moderately correlated with some of those performance indicators, which was interesting in that it was correlated with the performance indicators which are more plant safety equipment performance and not so much the inspection findings or the broad indicators.

MEMBER STETKAR: We are back on the record, so I have to be careful about what I speak of. The example I used: The folks at that plant were very well trained for the things that they did every day. There was a high correlation.

They said they were very well trained, and indeed they didn't have many unplanned scrams. They didn't have many forced outage hours, and their equipment outages were very infrequent and of very short duration when they had to so something.

Yet they were completely untrained on things that could be of substantial safety significance. How do you rationalize those kind of examples, given this type of process? Let me call it a process.

DR. MORROW: It is almost a question of where would that show up, if we look at the factors of the safety culture survey.

MEMBER STETKAR: At what point?

DR. MORROW: It might not show up. Yes, well, but it might not show up in training quality, because they believe they are well trained, but it might show up in other areas like questioning attitude where we don't question things.

So it might be that they believe they are well trained. It is really an issue of that they are not trained for emergency procedures, that sort of thing. They are not trained in every way that, if we look at it from an outside position, they should be trained on this. They should have emergency operating procedures. They are not aware of that. That is a blind spot for them.

So it might not show up in the factor that we would expect it to, and that is why it would be necessary to dig deeper. That is why it is necessary to do more than just a survey.

MEMBER STETKAR: But what in this process -- and I will call it a process. What in this process gives me confidence that that flag will be raised, because I don't see anything here.

DR. MORROW: I am going to consult a friend.

MEMBER STETKAR: Because that is what we are trying to do ultimately, I think.

MS. FRANOVICH: Rani Franovich, NRR staff. My perspective: We have regulatory requirements in this country that ensure that plants have EOPs, and that they train and they practice on EOPs. I think that what staff is presenting here is a framework with that premise, that it is our regulatory environment we have here. Could we rely on some of these implications.

If you take that out of this context in which we are talking, all bets are off. In another country where they are not required to have EOPs, maybe safety culture is not a requirement.

MEMBER STETKAR: You know, I didn't say that -- and I have to be careful here -- that my specific example did not have EOPs. They had very, very detailed EOPs.

MS. FRANOVICH: I thought I heard you say that they did not.

MEMBER STETKAR: Not for this particular class of events.

MS. FRANOVICH: Okay. Okay.

MEMBER STETKAR: They had extremely detailed, and they were reviewed and trained on them, and their regulators looked very carefully at the EOPs that they had.

MS. FRANOVICH: That is a different understanding than I had earlier. That said, I think --

MEMBER STETKAR: They were not a developing Third World country that was doing this ad hoc.

MS. FRANOVICH: That said, I think it is probably appropriate to look at what the staff is presenting on the context of this country, these plants, the regulatory basis that override the things that we would expect them to do.

MEMBER STETKAR: And I don't want to belabor that specific example. I just bring it up as an example to kind of pique curiosity about this process, and trying to gain some assurance that somehow -- Now I said how do I gain confidence that, if something like that were to exist at any of the plants in our country, it would be flagged.

DR. KOVES: Do you mind if I jump in here? Okay. You know, when you say "this process," I would have a -- To me, you are asking the question, you know, will this survey always catch everything? And the answer is no.

MEMBER STETKAR: Obviously, not.

DR. KOVES: Obviously not. That is why, you know -- and as Steph had said, you need to do the follow-up. There are areas that you see that are weak or even that you don't see that are weak.

The assessment process, the nuclear safety culture assessment process, SCART, SCAV which is being done by the Wano Paris Center now - they all come back, and even on areas that are correlated strong, they will do some interviewing and questioning about that to see, okay, where do these people think they are? Where do I as an evaluator think they are relative to a national or international scale, not just what they think.

So there is a follow-up that way, plus more follow-up in the areas of where they are perceiving themselves to be weak.

So when I would say the overall process, I think that overall process, the survey with the interviewing and observation -- I think the two of them together will give you indications 95 -- this is my number, okay; I am just pulling it out of my ear -- 95-98 percent of the time, you are going to catch something.

It is still guaranteed to, you know, not, but I think the two together -- When you are just using the survey alone, well, you know, is there a greater chance that something is going to slip by? Yes, there is. But it also can be very useful to point out those hot spots where you need to dig in. Thanks.

CHAIRMAN BLEY: I find this slide interesting in a few key ways, and suggest it to others. There are some nice things here, but down at the bottom it points concerns about unplanned scrams and forced outage hours. Training quality seems to be a key, and it crops up across the board.

There are other ones up here that would be very useful in the plants, What isn't here are those links to serious events, events that could challenge offsite risks. I can think of a couple of ways to go at that.

One is flipping the process around and looking at real events and overlaying them with the survey and trying to -- not with the survey -- overlaying them with the issues, with the factors, and seeing how they align with what happened in those events.

It is not the same kind of analysis, but I think it could generate a similar chart that might relate to those. There is a project going on now in research with one, from rumors I have heard, maybe some other utilities in actually collecting data from simulator exercises for those simulator exercises that put people through really challenging events, and looking through that information, and playing it against these, maybe even in a similar way to the way you have done this might be able to tease that out.

Right now, what you are showing us says we have some confidence that the things we have tagged in safety culture will help us in a number of areas, but they are not saying they will help us in the big nasty area that we are really concerned with, and I think there are things that could be done that would go a long way to doing that.

I hope you can think about that some, and maybe we will think about it more here. But I think there are ways to go at it, and this chart shows you very nice things that are important and operations that this can help with. So we are on the edge of something that might really link to public safety if we can push it.

DR. KOVES: It is very interesting you mentioned simulators, because INPO this year has stepped up evaluation of simulators and people working in the simulators.

CHAIRMAN BLEY: And actually urging to get more complex scenarios.

DR. KOVES: Right.

CHAIRMAN BLEY: Like some of the real events that have happened in the last year or two.

DR.KOVES: Exactly. More complex events, and also then evaluating more people, because my understanding is before they would just like evaluate one team and naturally that is some A-team for that. But I am not part of that. So that is just my understanding.

Yes, if we were doing that, and we were

seeing more failures, that could be a very interesting research project, obviously more qualitative. You could kind of do a quantitative aspect, too. Coming from that direction, it would be more qualitative, but it could be very interesting.

CHAIRMAN BLEY: And I think there is a growing body of information like that. One of the big events was the Robinson fire a year or two ago, and there have been a number of drills run that tried to set up that kind of complication, and people don't do as well with things coming in multiples and hiding behind each other. It is no big surprise, but we haven't been working on that real hard, and now we are starting to. So there might be something there.

DR. MORROW: That also reminds me that OE has been doing a number of case studies, I think, both inside and external to the nuclear industry, and they have been doing, actually, similar to what you were talking about.

They have been viewing those studies as severe accidents and applying the Safety Culture Policy Statement traits and going through the documents from those accidents and looking at whether there seemed to be issues cropping up related to the Policy Statement traits.

MEMBER STETKAR: I don't even think it needs to be severe accidents. I was just reading an event report the other day where, because people didn't check and ask about things, something was not programmed correctly. This was a nuclear plant.

As a result of that, a plant lost not only offsite power but all AC power because of the plant configuration, and although it wasn't as severe as the Robinson event, it wasn't as severe as Davis-Besse, it is an indicator of -- I don't want to point -- organizational or individual or something that was not done that led to this process.

I think you can look at those things. It doesn't have to be big dramatic things.

DR. KOVES: But I think what you are -- If we go back to the Swiss cheese model and reason and some of the reading that I have been doing recently, I think you are going to define -- those big events are going to define as multiple barrier failures.

I think you are -- My hunch is that what you are going to see is that on those big events you are going to see failures in across multiple areas.

CHAIRMAN BLEY: But you will find latent problems that are linked to these. I am convinced of it.

DR. KOVES: Oh, yes.

MEMBER SCHULTZ: It goes back to what was on your slide, Val, of looking at micro- events or mini-events, however you want to classify them, but a database that would categorize -- collect and categorize that information would be extremely valuable to set up against this process.

DR. BARNES: The question that I have about doing those post-event analyses or post-event/post-accident is whose judgment are we going to rely on to determine what trait was active in the event, number one, and number two, how bad or good the organization was on some scale in terms of that trait?

It is not that those kinds of analyses are not very informative. It is simply that we lack a large number of people agreeing that this trait was important in this incidence or we are missing some inter-rat er reliability there, and often, you know, if you have been involved in accident investigations and are trying to go back and figure out why something occurred, there is often large amounts of disagreement, and it takes a lot of data gathering and arguing to come to agreement about what the causes and contributors were to an accident.

So I think that we would be at risk of --

unlike a study like this where we have got opinions from 30 or more people about the state of each trait sort of within their organization, and then go out and link it to whatever performance indicators we might be interested in, when we go back and do these detailed events analyses, we would have small counts of traits being contributors, because we -- and there's lots more room for disagreement about which trait was actually having what impact where.

Not that it is not worth doing. It is just that that approach also has some limitations in terms of being able to identify causal relationships and the strengths of causal relationships between a particular trait and one kind of outcome or another.

MR. PETERS: Dennis, this is Sean Peters. I did have a question about how you would envision collecting safety culture information from simulator trials. What types of ways do we see how we gather organizational performance or safety culture capabilities based upon basically made-up simulator accidents?

How we have been relating safety to culture is associated with a number of real events out there in industry.

CHAIRMAN BLEY: Real events of very little importance to reactor safety, actually.

MR. PETERS: Yes.

CHAIRMAN BLEY: So these would be simulated events of significant importance to reactor safety. So while there is a tradeoff there, I lean toward the second. It would be rather like post-event analysis of certain exercises that were of real significance, but also real events. And sure, there are problems, but for the more significant events there have been whole teams get together and analyze them, and eventually agreement between the regional inspectors and the utilities on what were the key factors involved.

i could envision in the future that the safety culture traits could be incorporated into reviews of events. They weren't in the past, but very often those events are tagged to things that you can link with at least as much confidence as we have given large labels to the primary factors as to their significance.

MR. PETERS: I guess my question is: The actual events out there, a lot of the initiators and other underlying defects are associated with some type of safety culture trait in the plants.

CHAIRMAN BLEY: Sometimes. Some of our

responses along the sequence of events that occurs after the initial, and some things that were preexisting latent problems that you didn't notice until you got in the middle of the event, and they cropped up and caused problems.

MR. PETERS: Sure. And here the existing latent problems would be pre-programmed by a trait or something along those lines. So I was a little worried that we wouldn't be evaluating those aspects of the safety culture; maybe you just evaluate operator performance under the assumption that, hey, you already know we have issues; how do we perform.

That seems to be more of an evaluation of like the training quality of certain sub-aspects but not the whole safety culture as a whole.

CHAIRMAN BLEY: I am not sure we linking the whole safety culture to real safety right now. Making this up on the spur of the moment, I can't completely testify, and we don't have the whole answer here. We have a good answer for some things. So I would see that as a research project to do it right.

MEMBER SCHULTZ: Trying to make a connection to it, one way would be simulator, and it is another set of data that could be used. I don't know if you have to attribute all kinds of human error problems and issues related to what events you might catalog on the simulator, but the information could certainly be useful, or at least we could find out how useful it is.

Another area to explore might be a station's Correction Action Program, and thinking of micro-events and so forth. Now I think it would be difficult to do, but perhaps there are categories within the Correct Action Program that you could investigate and identify certain trends that might be associated with elements here.

If Davis-Besse happened today, we have got a tool here we could look at immediately, use immediately to assess safety culture pre-event, because it would only be yesterday.

The problem we have today is what has been developed at INPO comes from the state safety culture and Davis-Besse safety culture, and so that is where that -- That is where the technology of safety culture has emerged from, and we are trying to identify how we would connect this into, I think, the micro- or mini-events that need to be identified to connect to.

What is being developed here is an important link. I don't think we have it, but it is certainly worth thinking about in a research sense.

MS. FRANOVICH: Rani Franovich, NRR staff. I just wanted to add that when there are events and reactive inspections for those events, to the extent that performance deficiencies are identified, they will get a cross-cutting aspect assigned, and those cross-cutting aspects, whether they are event related or performance deficiencies in routine operation, will be rolled up into the cross-cutting issue process, which was one of the things that the staff looked at for their correlations.

MEMBER SCHULTZ; And Ken mentioned that in his presentation as well, and I think that is a good approach. I think it is has the potential to produce information that can be useful for the licensee and for the regulator as well.

DR. MORROW: All right, we will move on. Now we are into the predictive correlations.

CHAIRMAN BLEY: I was just wondering if you are getting ready to leave. Are you?

DR. KOVES: Getting set up.

DR. MORROW: Take a little more time.

DR. KOVES: I don't have to run out yet.

CHAIRMAN BLEY: Okay. Well, but when you get to the point you have to run, if any members of the Subcommittee have anything for you, this would be a good time maybe to interrupt the flow of things. I guess not.

So we appreciate your being here, and whenever you have to slip out, we thank you very much for your presentation and discussion.

DR. KOVES: Well, thank you. Thank you for the opportunity to be here. I just want to say it has been great fun, like always. Hopefully, we can do it again in the future with new data and be able to narrow in on some of these things. So thank you.

CHAIRMAN BLEY: We will give it back to Stephanie now.

DR. MORROW: Well, this table is the first of two tables looking at the correlations between the survey results and the same performance metrics but in 2011. So before, you were looking at the performance metrics assessed at the same time period, and this is one year later.

First off, what you will notice right away is that there are fewer shaded cells. So there are many fewer significant correlations when we look at this information. Mostly what we see is the correlation between the Safety Culture survey results and the problem identification and resolution cross-cutting area, specifically the Corrective Action

Program component of the PI&R cross-cutting area.

CHAIRMAN BLEY: Now these with this -- No, never mind.

DR. MORROW: Okay.

CHAIRMAN BLEY: So where did they go?

MEMBER SCHULTZ: Is there a conclusion you have, Steph, about having looked at these two sets, 2010 and 2011? It is a dramatic change.

DR. MORROW: Right. This is not unexpected, because what we were looking at first is kind of everything happening at the same time period. So the employees' perceptions of their safety culture related to how the plant was performing during the same time period.

Now if we look at -- You know, as we get father away from that same time period, things change over time. So the safety culture may change as well, but these kind of plant safety metrics are going to -- during the next year, are not necessarily going to be as related to the safety culture in the previous year.

Now we would love to see these predictive correlations, and actually on the next table that I will show you -- I will skip forward just a second.

MEMBER STETKAR: Let me ask you something before you get to this. On the statement you just made, I understand how plant performance can vary from year to year. Get a bad year, a couple of times. Do you really think that plant safety culture is that variable over time year to year or did I misunderstand what you were saying?

DR. MORROW: I will correct my statement, because I misspoke.

MEMBER STETKAR: Okay.

DR. MORROW: What I think can vary from year to year is employee perceptions of the safety culture, so what the survey is tapping into. The underlying safety culture, it is much more stable, and it doesn't necessarily change from year to year unless there is some significant event that will change it. However, if there are smaller events that are occurring, such as they move in the action matrix and suddenly they are a lot more doing reactive inspections or something like that. The employees' perception of the safety culture may change as a result of that.

So I will distinguish between kind of the underlying overall safety culture and what the survey is tapping into, which is more of the surface level. So those perceptions are much more likely to change.

CHAIRMAN BLEY: What has to have happened in the intervening year is that all of these things that get measured have changed in a way that they no longer correlate as well as they did.

DR. MORROW: Right.

DR. KOVES: Except a few of them.

CHAIRMAN BLEY: A few, but even those are down a little.

DR. MORROW: Actually, if you look --

CHAIRMAN BLEY: I want you to swear it is not, but --

DR. MORROW: But still, if we are talking about 10 percent of the variability in these plant safety metrics, if that is a unique 10 percent, then it might be worthwhile to look at. What was interesting about the significant correlations, looking at the 2011 data -- and I actually have an extra slide which I may need to pull up in a minute, but the elevated action matrix, the totals of standard cross-cutting issues and allegations, looking at the 2010 results, they weren't significant correlations.s

I just have this up here in case it came up, which it seems to. This is looking at the 2010 versus the 2011 data. So you will see the overall safety culture results were not significantly correlated with this data in 2010, but it was in 2011.

You will also note that these three variables in particular are the more broadbased performance assessments: Whether a site is in one of the elevated oversight columns of the action matrix; whether they have outstanding substantive cross-cutting issues.

So this was a rather interesting finding in that it seems that, if we are looking at data in the future, the safety culture survey results seem to be more strongly related to the broadbased performance assessments the following year.

MR. WIDMAYER: You need to send me a slide.

DR. MORROW: I will send this slide to you. Actually, it is in the electronic file.

MR. WIDMAYER: Oh, it is?

DR. MORROW: Yes.

MEMBER STETKAR: Personal responsibility for safety is very dramatic, at least in the elevated source column -- in the first metric.

DR. KOVES: To me, just on a very simple level, particularly the allegations results -- I mean, to me it was like, if you have a poor safety culture this year, you are going to have more allegations next year, you know. Also, I mean in terms of the cross-cutting issues, next year you are going to have more of them. To me, it is logical.

DR. MORROW: Well, we could speculate that, if employees at a site feel like there isn't a strong questioning attitude, they feel like issues aren't being raised through the Corrective Action Program, then it might be kind of when they fail at that, that later on they start to go to the NRC and make allegations, or if they feel like there is not a respectful work environment.

That is kind of later on in that process that they will start to look outside of their organization for help with these issues. But it is one possible explanation for these correlations.

Okay, now we are going backwards. This, for example, is the scatterplot looking at the overall safety culture survey scores and allegations in 2011. So there is some dispersion of data around kind of that best fitting line, and it is not just a single outlier.

These are some of the key observations we noted from the correlations of the 2011 data.

Sorry. There we go. Okay. Most of the correlations when looking at the 2011 data between the survey results and the performance metrics that were from the sources of data from inspection reports were small and nonsignificant, looking at the 2011 data. That was that first table that I showed you, but the overall safety culture survey was moderately correlated with elevated oversight and action matrix and allegations from licensee personnel.

Kind of the strongest correlations that we saw looking at the 2011 data were with, again, questioning attitude was correlated with allegations, and the management responsibility and willingness to raise concerns factors. This is what you were asking about, I think, and the 2011 data was correlated with allegations

DR. KOVES: Yes. Right. It wasn't correlated well with the '10 data.

DR. MORROW: Right.

DR. KOVES: But it was correlated very well with the '11 data.

MEMBER SCHULTZ: Steph, I am trying to compare the two 2010 and '11 for allegations. One is titled SCWE-related allegations, and the other is allegations. Are they, in fact, the same?

DR. MORROW: The allegations is the more general variable. So that is the total count of all allegations from licensee personnel. The SCWE-related one is a variable that we had with the 2010 data. It is allegations that were specifically categorized as relating to a safety conscious work environment. We didn't have that specific categorization when we were looking at the 2011 data. So it just looks at the general allegations.

MEMBER SCHULTZ: All right. Thank you. And that explains why there are more in 2011, and it is a different data structure for the correlations.

DR. MORROW: Right.

MEMBER SCHULTZ: Do you have the 2010 for total allegations?

DR. MORROW: The correlation is there. I don't think I have the scatterplot.

MEMBER SCHULTZ: Okay. I will try to compare that. Thank you.

DR. MORROW: Sure. I will bring that last slide up again. these were the comparisons, 2010 to 2011, because you don't have that with you.

MEMBER SCHULTZ: Okay, thank you.

DR. MORROW: And that concludes my presentation. So if you have additional questions for us about -- while we still have Ken here as well for a few more minutes.

MEMBER STETKAR: See how quickly we get done when you are not here.

CHAIRMAN BLEY: Thank you. Anything more? I think we will go around and just make comments from committee members. Steve, I will start with you.

MEMBER SCHULTZ: Thank you. I am very impressed with the work that has been done and the presentations that have been made today. I certainly appreciate, Ken, you being here and providing the insight that you have developed and the methodology that you have been developing on behalf of industry, in collaboration with the NRC. I think overall that program has provided what appears to be a more promising product that has been aimed at developing a program that will be useful to the licensees as well as to the NRC, and time will tell.

As we have discussed today, I think it is important to continue to look for the connectivity between the questionnaire and information we can derive from safety culture to a real safety benefit, which we have discussed today. I still feel we are missing that connection, although, certainly, it would appear that progress has been made on that accord in two different areas.

I think we have a much better

understanding about how one is able to quantify these relationships and the difficulty -- at least some understanding of the difficulties related to that. But I think, from what we have seen today, there is certainly some promise, but I think we are still missing that connection to, as Dennis described earlier, what we are trying to connect to as an end product, which is public health and safety. But I do believe we are making some progress there, because as we have discussed safety, in and of itself, we know has a benefit, I believe, to public health and safety. It is a difficult one to quantify.

These last results, which we haven't had a chance to study, are interesting, looking at this information related to the survey results and its correlation to these measures, broad measures, both from INPO and from the NRC's program, are somewhat telling.

Again, clear connectivity to public health and safety has yet to be established, but it would appear that I am gaining some ground and understand certain input with regard to understanding. But thank you very much for the presentations. It was very high quality, and it has been very helpful in moving the technology, as it were, forward. Thank you.

CHAIRMAN BLEY: John.

MEMBER STETKAR: Thanks. I have said, I think, most of what I need to. I would echo Steve's.

I think my sense is that the survey, the questions, I think, is -- it seems to be asking the right things in the right format. I think that the process that you have gone through to winnow down the questions and to sort of systematically challenge yourself in terms of what is a reasonable set of information to derive is very, very good.

I will share Steve's concern. It is not -- pretty evidently to me anyway, not clear to me that correlating those responses with particular metrics that you are using in terms of either INPO or reactor oversight process is the right way to look at the problem.

I don't know if it isn't the right way. It is not clear that it is. So I will just leave it at that. Thank you.

CHAIRMAN BLEY: Thanks, John. Mike.

MEMBER RYAN: Thanks, Dennis. I guess I would second Steve's comments, too, and offer a few that it is fascinating data, particularly the way you have laid it out. I would second that idea. It is very insightful and helpful to somebody that hasn't been involved to learn quickly.

Two questions come to my mind. Okay, what do we do with this now? And it is obvious you've got a next phase planned, and that is underway, and you are going to collect some additional data.

Whenever folks talk about collecting data, I have a favorite question. When are you done? The answer is never, okay. What benchmarks are you going to have along the way to assess when you are done, and then when you are done, what do I do with it? I am a licensee.

So I think those are two big questions that come after the accolades that Steve talked about, and I agree with every one. Now is the time that you are kind of in the middle with the alligators swimming around, and stuff to see.

You know, where are you going to take all of this in terms of the guidance, regulations, requirements and things of that sort to memorialize what you are learning and how you are learning it and the way you would, I'm guessing, recommend that licensees use this information or gather this information and analyze this information in the way that you recommended.

So it is not too early to start thinking a little bit more about that, I think. So I just offer that to you as an observation. Thank you. And again, I think you did a great job. It has been very, very well prepared and well presented.

CHAIRMAN BLEY: I would like to thank you very much for the White Paper and for coming back. I won't say that I wished I had a long time ago, but it is really good. It makes it very clear.

At the last Subcommittee meeting, there were a lot of issues about the analysis that just weren't clear to us. So it has really helped a lot, and having all three of you here to talk about it is very helpful.

We are still stuck with we've got correlations that aren't causal or predictive, maybe. I have to say again what the other guys said. You need to test it against public safety significant events somehow. That is what we are really after.

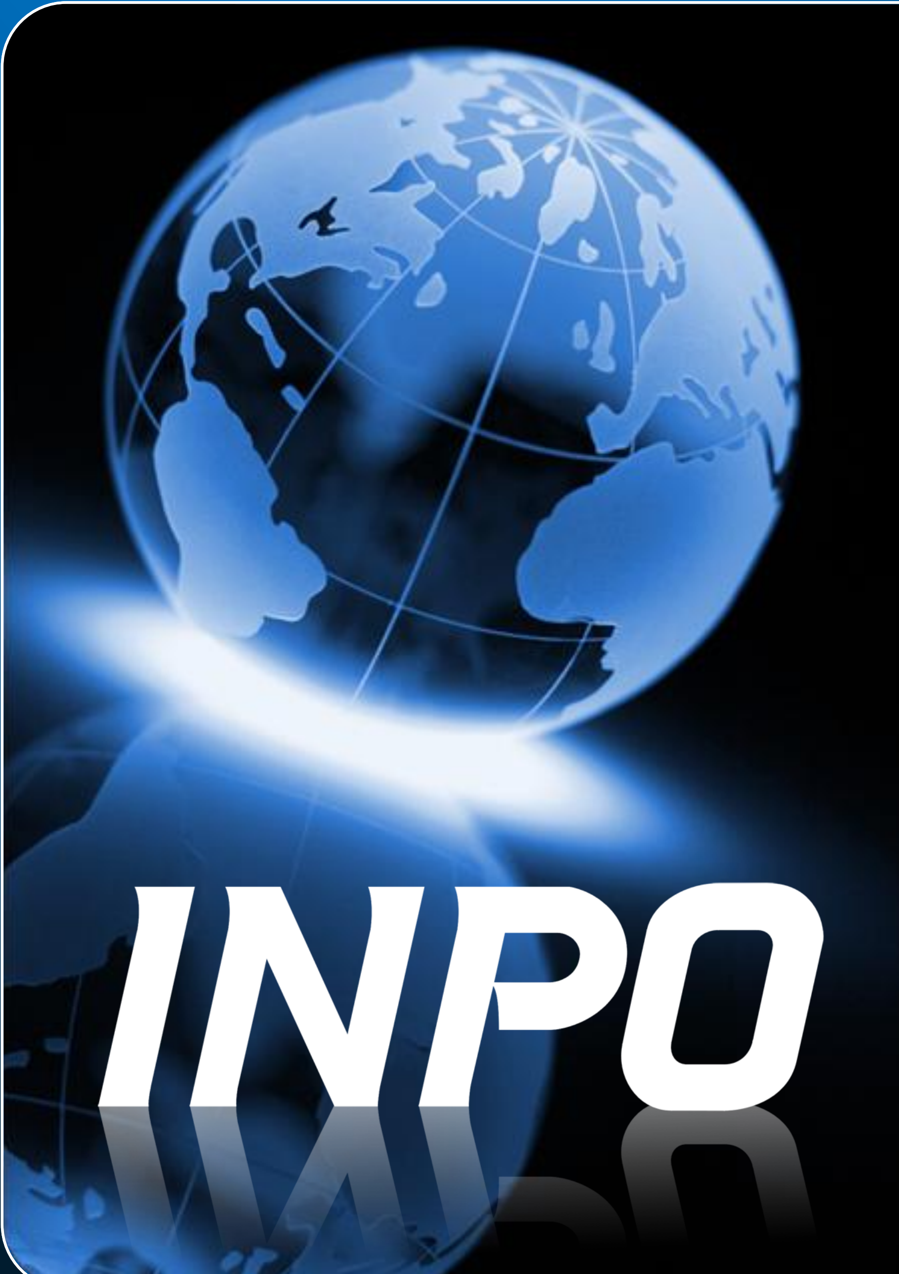
This last bit on the 2011, I got to think about that. I don't know what that is telling me, and I hate to see all those, albeit in my opinion weak correlations, sort of go weaker, and I don't know exactly why. I have heard speculations, and I can speculate as well. I don't know, maybe this is something that always has to go hand in hand in comparing the current culture with what is going on.

Anyway, it was a very useful day for us, and thanks very much for coming.

I forgot to ask earlier, which I should have done. I don't think we have any members of the public, but if we do, is there anyone here who would like to make a comment?

Hearing none, thanks again, and this meeting is adjourned.

(Whereupon, the foregoing matter went off the record at 3:49 p.m.)



SC Survey Validation using INPO Key Performance Indicators

G. Kenneth Koves, Ph.D.

ACRS Subcommittee on Reliability and Probabilistic Risk Assessment Meeting

2012, May 8,

Washington DC, USA

Central Question

- Does the measure actually measure what it purports to measure?
- Are the survey results related to other measures of safety?

INPO

Reliability and Validity

- Reliability
  - The consistency of a measure – it produces consistent results under consistent conditions
- Validity
  - The accuracy of a measure – it actually measures what it says it measures

INPO

Criterion Validity

- Criterion validity is the extent to which the measures are demonstrably related to concrete criteria in the "real" world
  – Concurrent
  – Predictive

Criterion Validity

- Concurrent
  - the extent to which the survey is demonstrably related to current results of the criteria
- Predictive
  - the extent to which the survey is demonstrably related to future results of the criteria

INPO

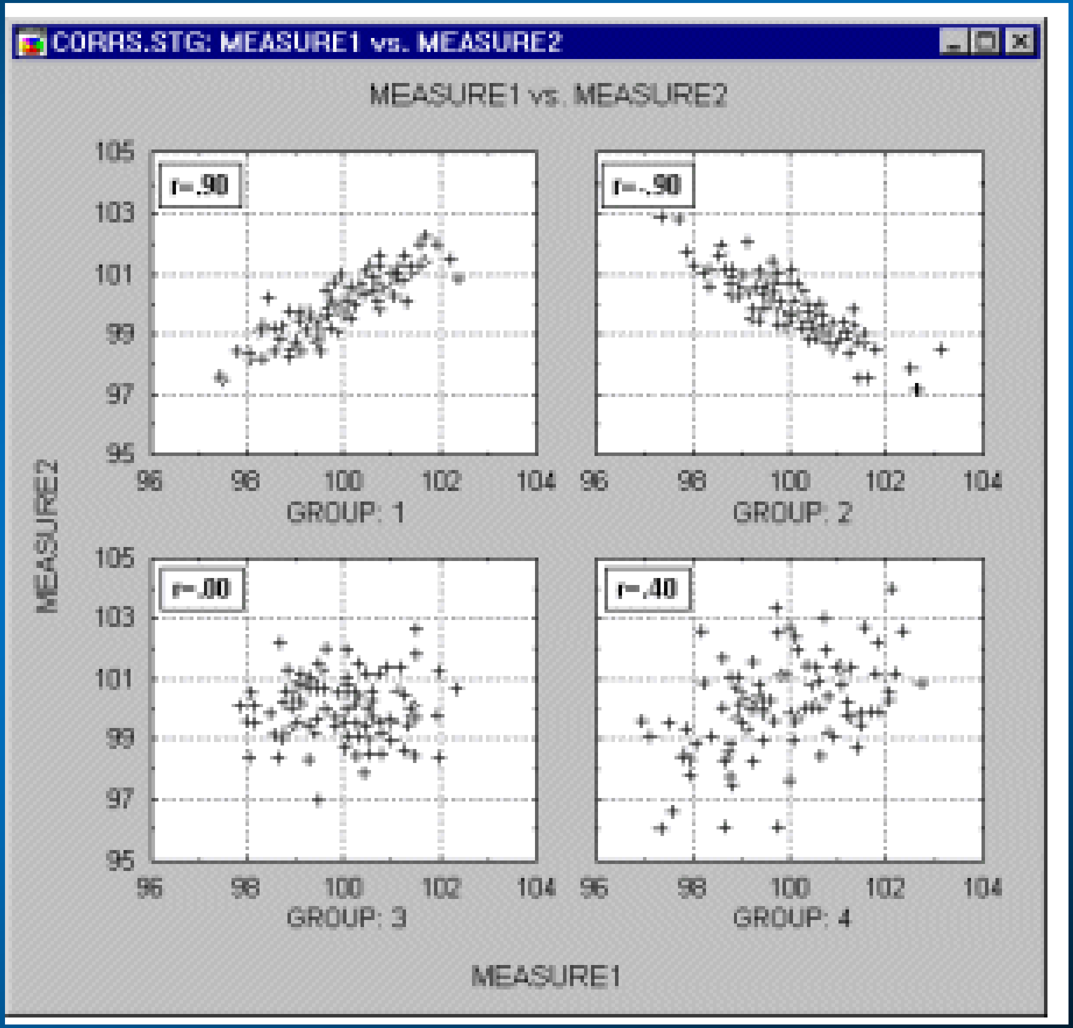
Correlation

- The degree to which two or more variables show a tendency to vary together
- The most familiar measure of correlation is the Pearson product-moment correlation coefficient. It is obtained by dividing the covariance of the two variables by the product of their standard deviations.

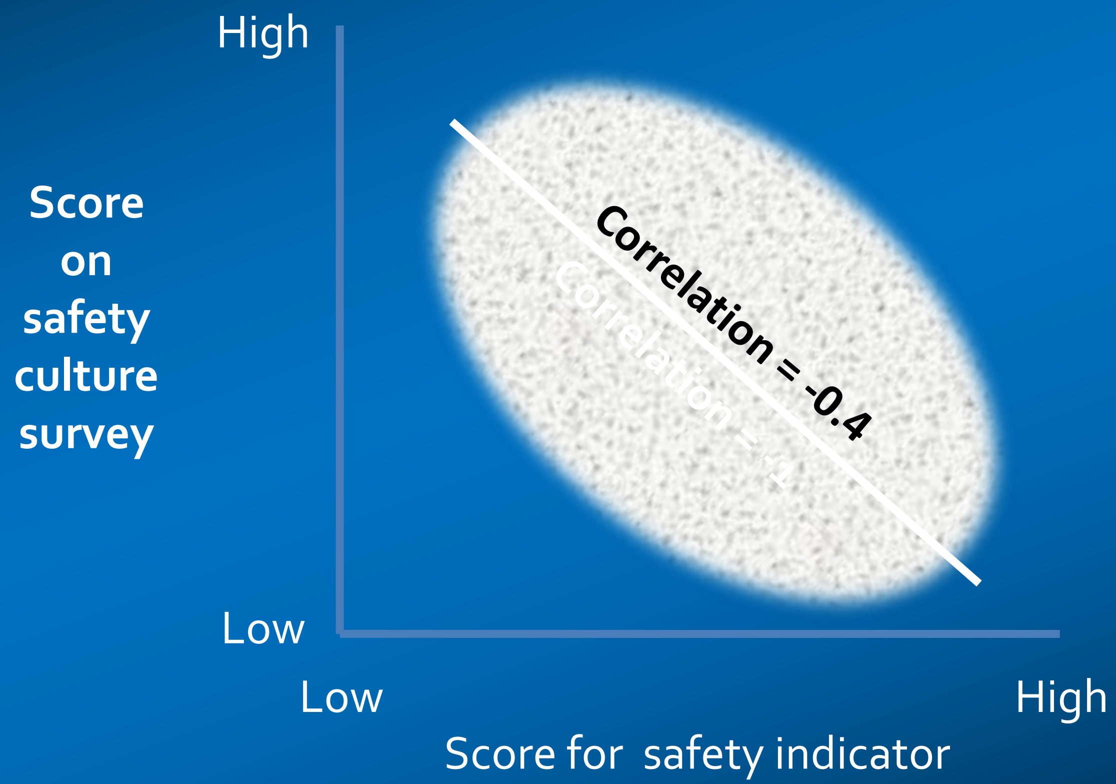
INPO

Correlation Examples

Correlations range from +1 to -1



Correlation Applied to SC Data



Social Science Correlations

- Engineers are typically accustomed to higher correlations than are seen in the social sciences
- Average correlations in previous meta-analyses were .22 and .31 (Clarke, 2006; Christian, et al, 2009)

Correlation Example

NPP # = 63	Survey Mean Score	Unplanned Critical Scrams	Unpln Auto Scrams	Emergency Power Unavailable	Per. Safety Index	Chem. Index	Hu Error Rate
NPP #1	6.12	0.00	0.00	1.74	6.25	1.00	0.00
NPP #2	6.06	0.22	0.28	1.50	3.13	1.05	0.01
NPP #3	5.99	0.44	0.16	1.42	0.00	1.00	0.02
NPP #4	5.93	1.00	0.21	0.91	18.75	1.00	0.02
NPP #5	5.73	0.33	0.31	3.29	40.63	1.12	0.02
NPP #6	5.58	2.67	1.30	2.39	40.63	1.00	0.03
NPP #7	5.52	2.00	0.64	1.13	25.00	1.01	0.04
NPP #n…							
Corr. with mean	--	-0.28	-0.34	-0.27	-0.26	-0.27	-0.36

Select SC Research Results

SC Survey Scores	NRC ROP	Unplanned Critical Scrams	Unpln Auto Scram	Emergency Power Unvailabil.	Per Safe Index	CY Index	Hu Err Rate
Mean Score for NPP	-.26	-0.28	-0.34	-0.27	-0.26	-0.27	-0.36
Manager Responsibility	-.30	-.29	-.34	-.26 (-.31)	-.23 (-.31)	-.27 (-.39)	-.38
Raising Concerns	-.25	-.17	-.24	-.27	-.22	-.22	-.37
Decision Making	-.32	-.28	-.38	-.24	-.25	-.28	-.36
Supervisor Responsibility	-.28 (-.35)	-.15	-.22 (-.40)	-.30	-.19	-.14 (-.32)	-.40
Questioning Attitude	-.18	-.27	-.26 (-.44)	-.37	-.32	-.26 (-.32)	-.28
Communication	-.20	-.32	-.34	-.27	-.27	-.28	-.39
Training	-.12	-.33	-.40	-.15	-.13	-.30	-.19

Subfactor correlations in parentheses

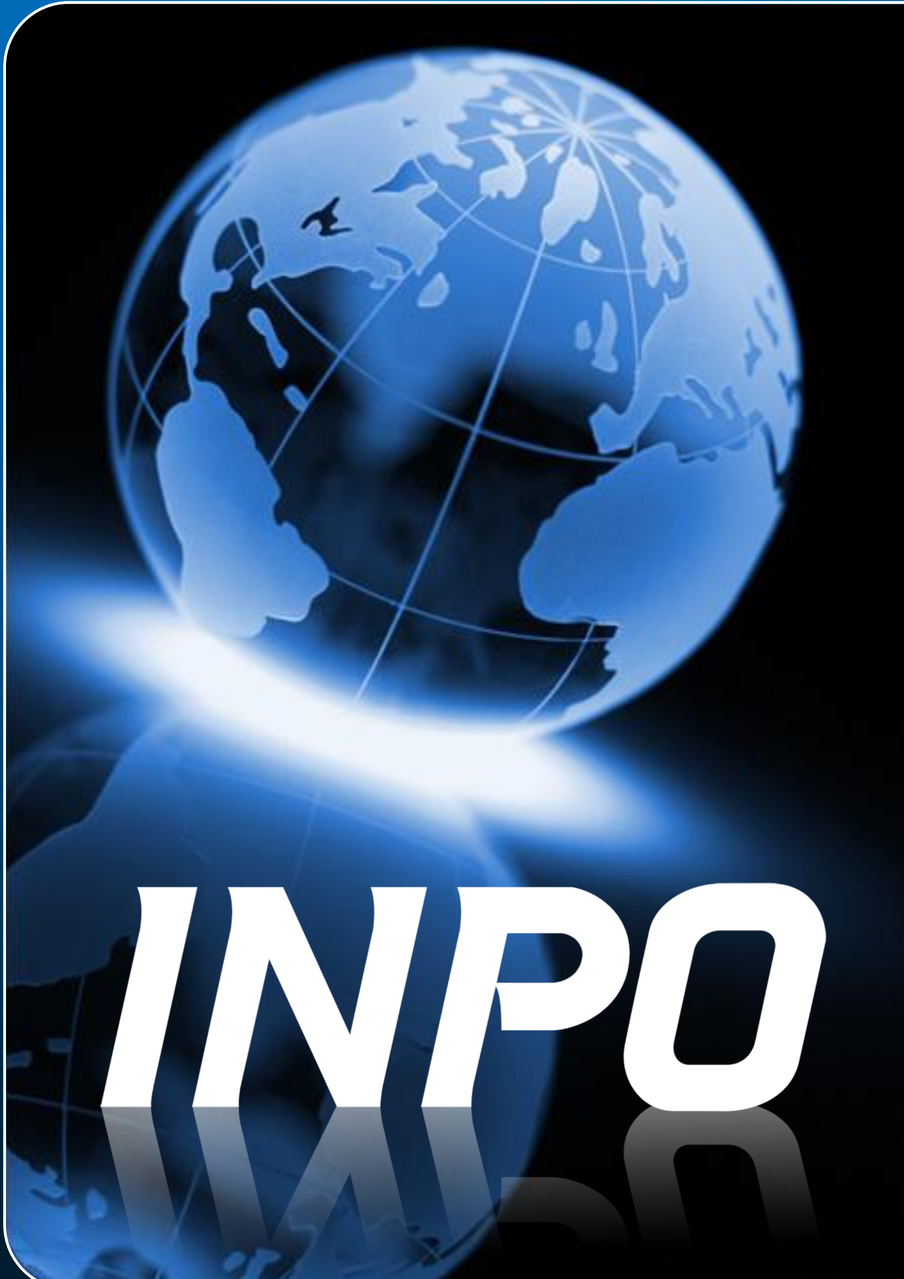
Predictive Validation Results

Survey Mean, Factors and Subfactors	Forced Loss Rate 2010	Forced Loss Rate 2011	ISA Rate 2010	ISA Rate 2011	Total ISA 2010	Total ISA 2011	Chemistry Perf. Ind. (CPI) 2011	Chemistry Perf. Ind. (CPI) 2011
Survey Mean	-0.11	-0.31**			0.02	-0.25*	-0.27*	-0.45**
Management Responsibility	-0.11	-0.26*			0.03	-0.25*	-0.27*	-0.45**
Respectful work environment							-0.26*	-0.45**
Continuous improvement	-0.08	-0.28*	-0.03	-0.27*	-0.02	-0.30*	-0.26*	-0.43**
Rewards	-0.14	-0.30*			-0.14	-0.29*	-0.39**	-0.45**
Willingness to Raise Concerns							-0.22*	-0.36**
Decision-Making	-0.08	-0.29*					-0.28*	-0.43**
Questioning Attitude	-0.19	-0.29*	-0.11	-0.32**	-0.01	-0.29*	-0.26*	-0.43**
*p < .05; **p < .01								

What does it all mean?

- The SC survey appears to be solidly related to other measures of safety
- The SC survey may be a predictor of some indicators of safety and SC, but more research is required

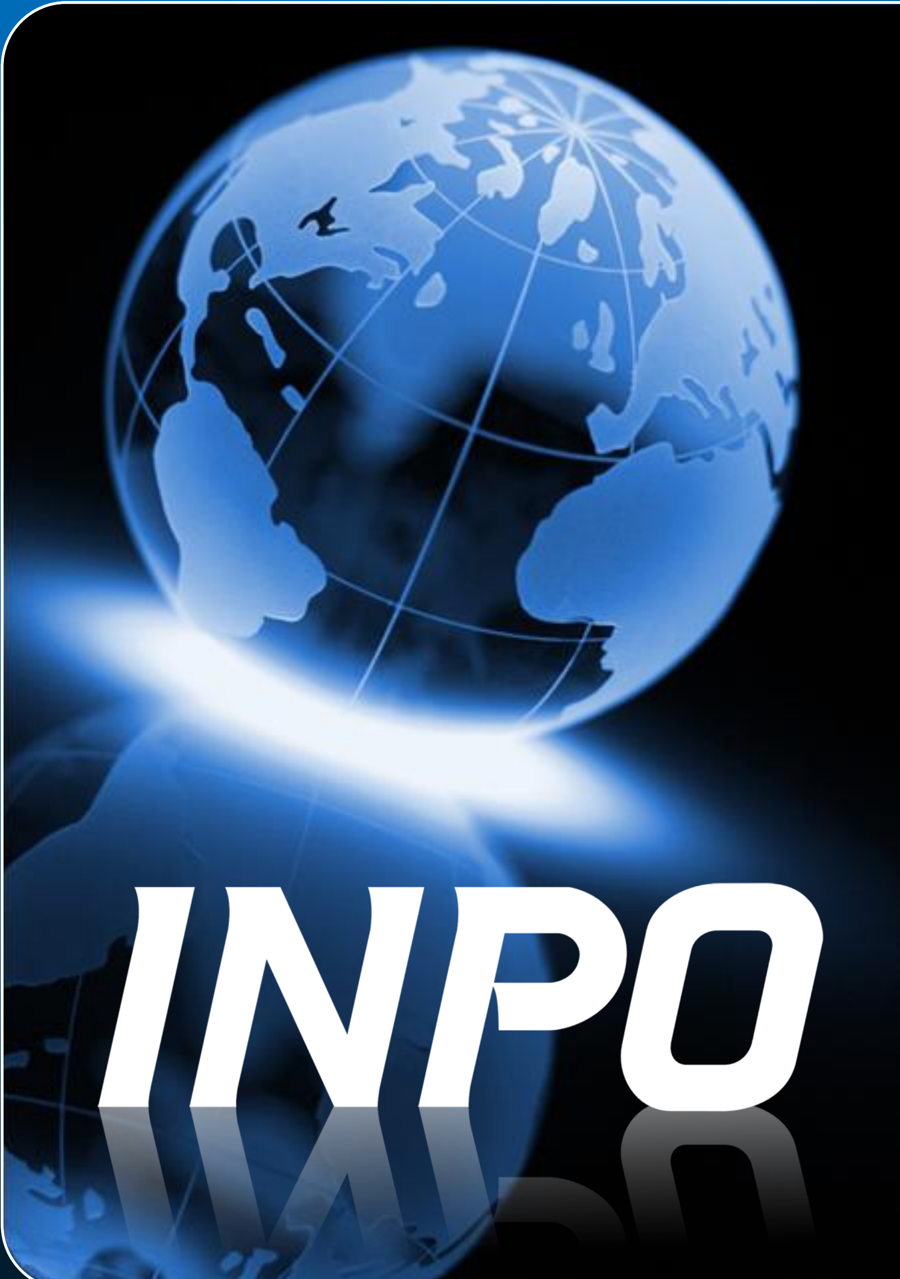
INPO



Questions and Discussion

G. Kenneth Koves, Ph.D.

koveskg@inpo.org
+01 770 644 8303



Survey Development and Factor Structure

G. Kenneth Koves, Ph.D.

ACRS Subcommittee on Reliability and Probabilistic Risk Assessment Meeting

2012, May 8,

Washington DC, USA

Introduction

- About Ken Koves
  - Senior Program Manager – INPO
  - Senior Organization Development Consultant – Sprint
  - Ph.D. I/O Psychology – Georgia Tech

Purpose of Study

- Validate the traits included in the Safety Culture policy statement for the nuclear power industry
- Determine relationship between INPO's safety culture survey results and concurrent measures of safety performance
- Determine relationship between safety culture survey results and the same measures of safety performance one year after administration

Roles

- INPO – survey strategy, planning, and analysis of industry indicators
- Vendor – Administered the survey
- NEI – paid for the vendor
- NRC – reviewed the work

Some Nuclear Models of SC

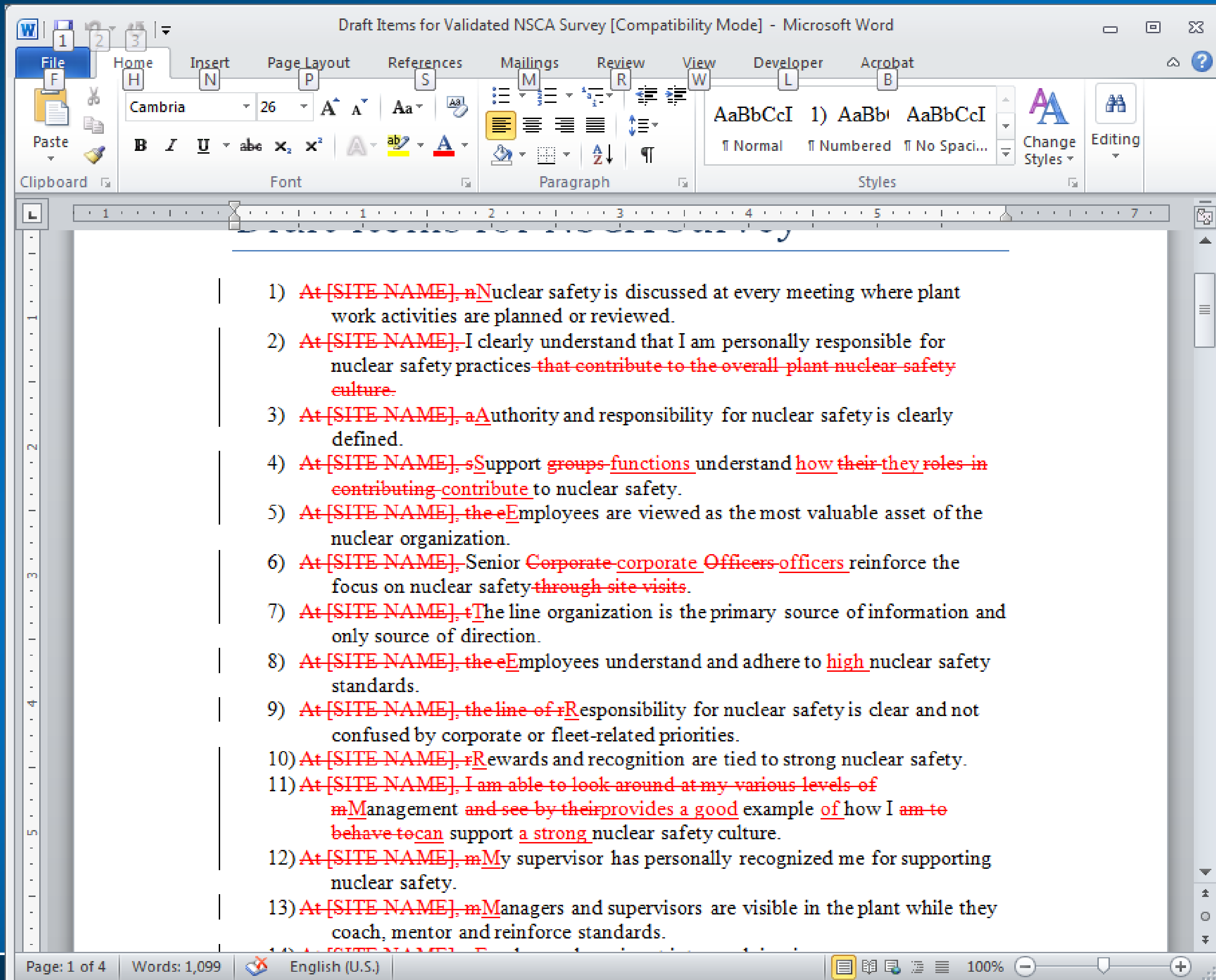
IAEA	INPO Principles	ROP Components *
Safety is a clearly recognized value • **Documentation** • **Resources** • **Planning** • **Decision-making** • **Social acceptance**	2. Leader Demonstrates 4. Decision Making Reflects Safety First 5. Nuclear is Special and Unique	1. Decision Making 2. Resources 13. Safety Policies
Leadership for Safety is Clear • **Commitment** • **Change Management** • **Conflict resolution** • **Trust**	3. Trust Permeates the Organization	8. Environment for Raising Concerns 12. Org Change Management
Accountability for Safety is Clear • **Roles and Responsibilities** • **Compliance** • **Ownership**	1. Everyone Personally Responsible	10. Accountability
Safety is integrated into all Activities • **Processes** • **Work conditions** • **Teamwork**		3. Work Control
Safety is Learning Driven • **Questioning attitude** • **Open Reporting** • **PI & R** • **EE development**	7. Org Learning 6. Questioning Attitude is Cultivated 8. Nuclear Safety Undergoes Constant Examination	5. CAP 6. OE 7. Self & Independent Assessment 9. Preventing Retaliation 11. Continuous Learning Environ

* 4. Work Practices too broad to categorize

Question Development

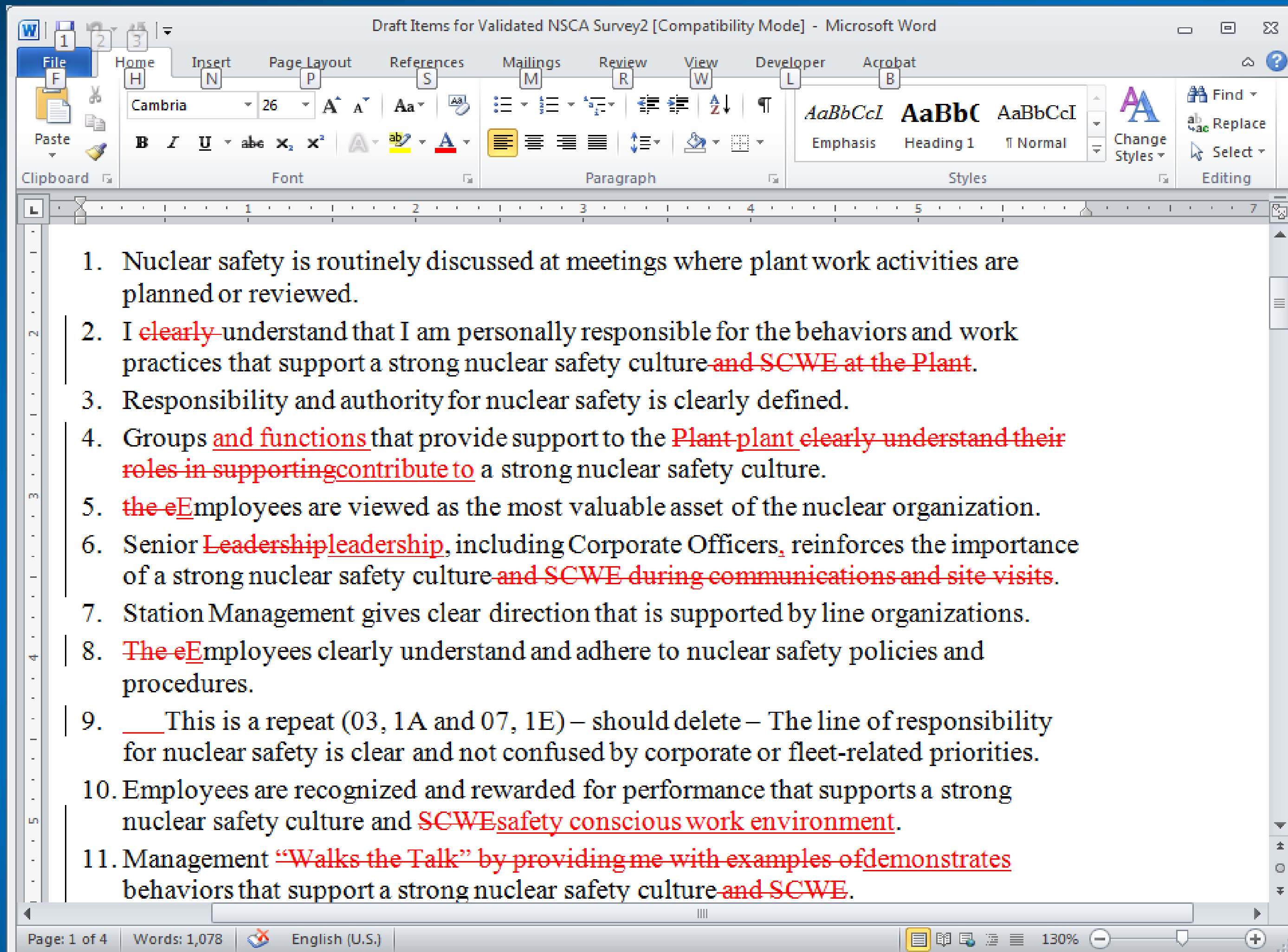
- Started with the Utility Service Alliance SC survey built from the INPO Principles for a Strong Nuclear Safety Culture
  - Similarity of models
  - Limited time
  - Only two nuclear models with existing questions

Development – Substantial Editing of Original 73 Items



1) ~~At [SITE NAME], n~~Nuclear safety is discussed at every meeting where plant work activities are planned or reviewed.
2) ~~At [SITE NAME],~~ I clearly understand that I am personally responsible for nuclear safety practices ~~that contribute to the overall plant nuclear safety culture.~~
3) ~~At [SITE NAME], a~~Authority and responsibility for nuclear safety is clearly defined.
4) ~~At [SITE NAME], s~~Support ~~groups~~ functions understand how ~~their~~ they ~~roles in contributing~~ contribute to nuclear safety.
5) ~~At [SITE NAME], the e~~Employees are viewed as the most valuable asset of the nuclear organization.
6) ~~At [SITE NAME],~~ Senior ~~Corporate~~ corporate ~~Officers~~ officers reinforce the focus on nuclear safety ~~through site visits~~.
7) ~~At [SITE NAME], t~~The line organization is the primary source of information and only source of direction.
8) ~~At [SITE NAME], the e~~Employees understand and adhere to high nuclear safety standards.
9) ~~At [SITE NAME], the line of r~~Responsibility for nuclear safety is clear and not confused by corporate or fleet-related priorities.
10) ~~At [SITE NAME], r~~Rewards and recognition are tied to strong nuclear safety.
11) ~~At [SITE NAME], I am able to look around at my various levels of m~~Management ~~and see by their~~ provides a good example of how I ~~am to behave to~~ can support a strong nuclear safety culture.
12) ~~At [SITE NAME], m~~My supervisor has personally recognized me for supporting nuclear safety.
13) ~~At [SITE NAME], m~~Managers and supervisors are visible in the plant while they coach, mentor and reinforce standards.

Development – More Editing



Draft Items for Validated NSCA Survey2 [Compatibility Mode] - Microsoft Word

1. Nuclear safety is routinely discussed at meetings where plant work activities are planned or reviewed.
2. I ~~clearly~~ understand that I am personally responsible for the behaviors and work practices that support a strong nuclear safety culture ~~and SCWE at the Plant~~.
3. Responsibility and authority for nuclear safety is clearly defined.
4. Groups and functions that provide support to the ~~Plant~~ plant ~~clearly understand their roles in supporting~~contribute to a strong nuclear safety culture.
5. ~~the e~~Employees are viewed as the most valuable asset of the nuclear organization.
6. Senior ~~Leadership~~leadership, including Corporate Officers, reinforces the importance of a strong nuclear safety culture ~~and SCWE during communications and site visits~~.
7. Station Management gives clear direction that is supported by line organizations.
8. ~~The e~~Employees clearly understand and adhere to nuclear safety policies and procedures.
9. ___This is a repeat (03, 1A and 07, 1E) – should delete – The line of responsibility for nuclear safety is clear and not confused by corporate or fleet-related priorities.
10. Employees are recognized and rewarded for performance that supports a strong nuclear safety culture and ~~SCWE~~safety conscious work environment.
11. Management ~~“Walks the Talk” by providing me with examples of~~demonstrates behaviors that support a strong nuclear safety culture ~~and SCWE~~.

Page: 1 of 4 | Words: 1,078 | English (U.S.) | 130%

Development – 10 Additional Items for Workshop Traits

69. ~~a~~ A broad set of performance indicators is utilized with a focus on early detection of problems.

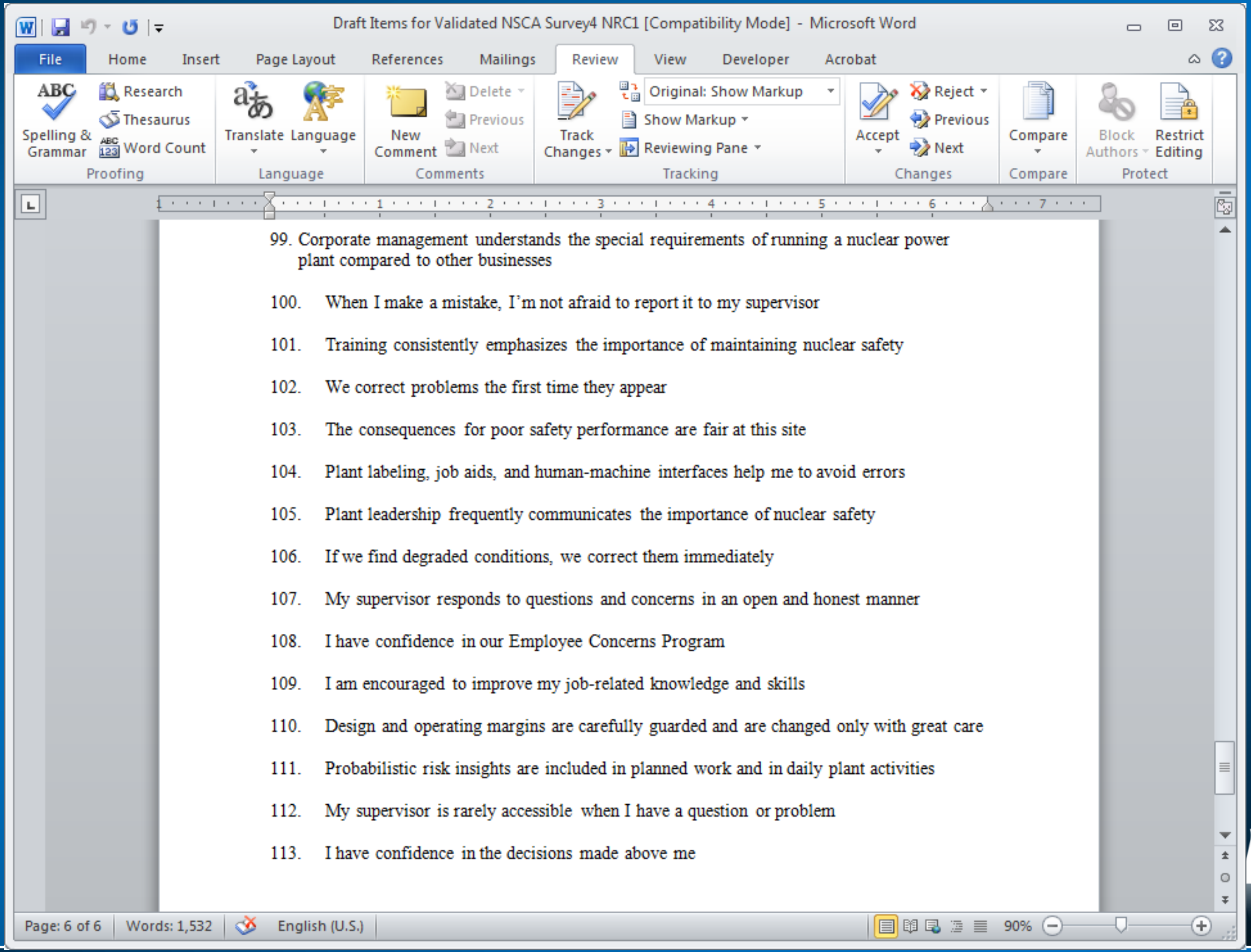
70. ___(Similar to 67) Delete – Senior management is periodically briefed by oversight on Station Safety Performance

71. ~~Senior management is periodically briefed by the Quality organization on Station Safety Performance.~~ We have a strong quality assurance process and organization.

Additional

72. We always go beyond the minimum requirements when performing work.
73. We never justify existing conditions.
74. Our nuclear oversight organization is effective in making positive changes.
75. We permanently correct problems the first time they appear.
76. The leadership considers safety more important than production.
77. If we find degraded conditions, we correct them immediately.
78. Our corrective action program is very effective.
79. I understand the plant's improvement focus areas for this year.
80. Plant leadership frequently communicates that safety is most important.
81. I can raise safety concerns without fear of retaliation.

Development – NRC Edits and 32 Additional Items



99. Corporate management understands the special requirements of running a nuclear power plant compared to other businesses

100. When I make a mistake, I'm not afraid to report it to my supervisor

101. Training consistently emphasizes the importance of maintaining nuclear safety

102. We correct problems the first time they appear

103. The consequences for poor safety performance are fair at this site

104. Plant labeling, job aids, and human-machine interfaces help me to avoid errors

105. Plant leadership frequently communicates the importance of nuclear safety

106. If we find degraded conditions, we correct them immediately

107. My supervisor responds to questions and concerns in an open and honest manner

108. I have confidence in our Employee Concerns Program

109. I am encouraged to improve my job-related knowledge and skills

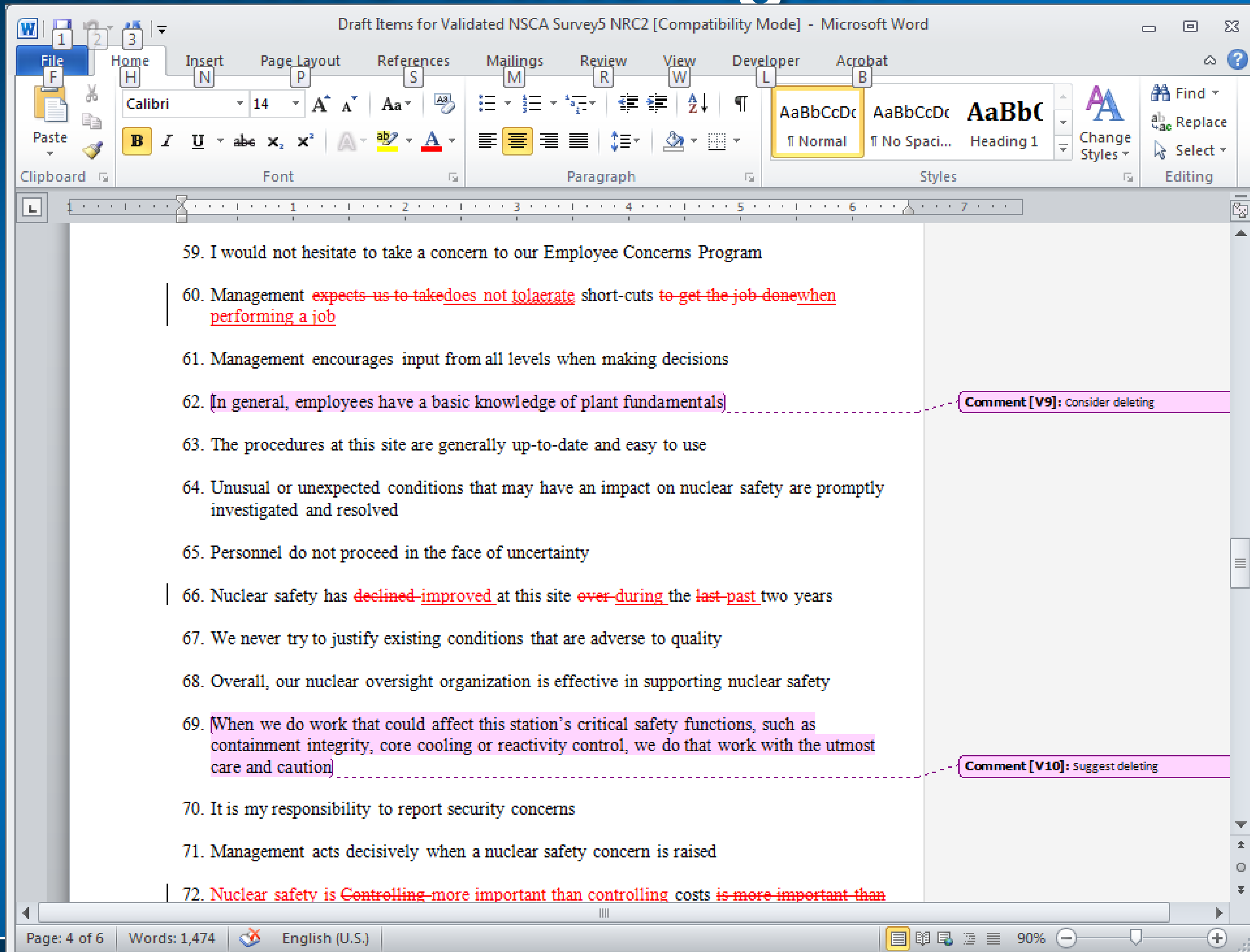
110. Design and operating margins are carefully guarded and are changed only with great care

111. Probabilistic risk insights are included in planned work and in daily plant activities

112. My supervisor is rarely accessible when I have a question or problem

113. I have confidence in the decisions made above me

Development – INPO Edits then NRC Edits Again



Draft Items for Validated NSCA Survey5 NRC2 [Compatibility Mode] - Microsoft Word

59. I would not hesitate to take a concern to our Employee Concerns Program

60. Management ~~expects us to take~~does not tolaerate short-cuts ~~to get the job done~~when performing a job

61. Management encourages input from all levels when making decisions

62. In general, employees have a basic knowledge of plant fundamentals

Comment [V9]: Consider deleting

63. The procedures at this site are generally up-to-date and easy to use

64. Unusual or unexpected conditions that may have an impact on nuclear safety are promptly investigated and resolved

65. Personnel do not proceed in the face of uncertainty

66. Nuclear safety has ~~declined~~ improved at this site ~~over~~ during the ~~last~~ past two years

67. We never try to justify existing conditions that are adverse to quality

68. Overall, our nuclear oversight organization is effective in supporting nuclear safety

69. When we do work that could affect this station's critical safety functions, such as containment integrity, core cooling or reactivity control, we do that work with the utmost care and caution

Comment [V10]: Suggest deleting

70. It is my responsibility to report security concerns

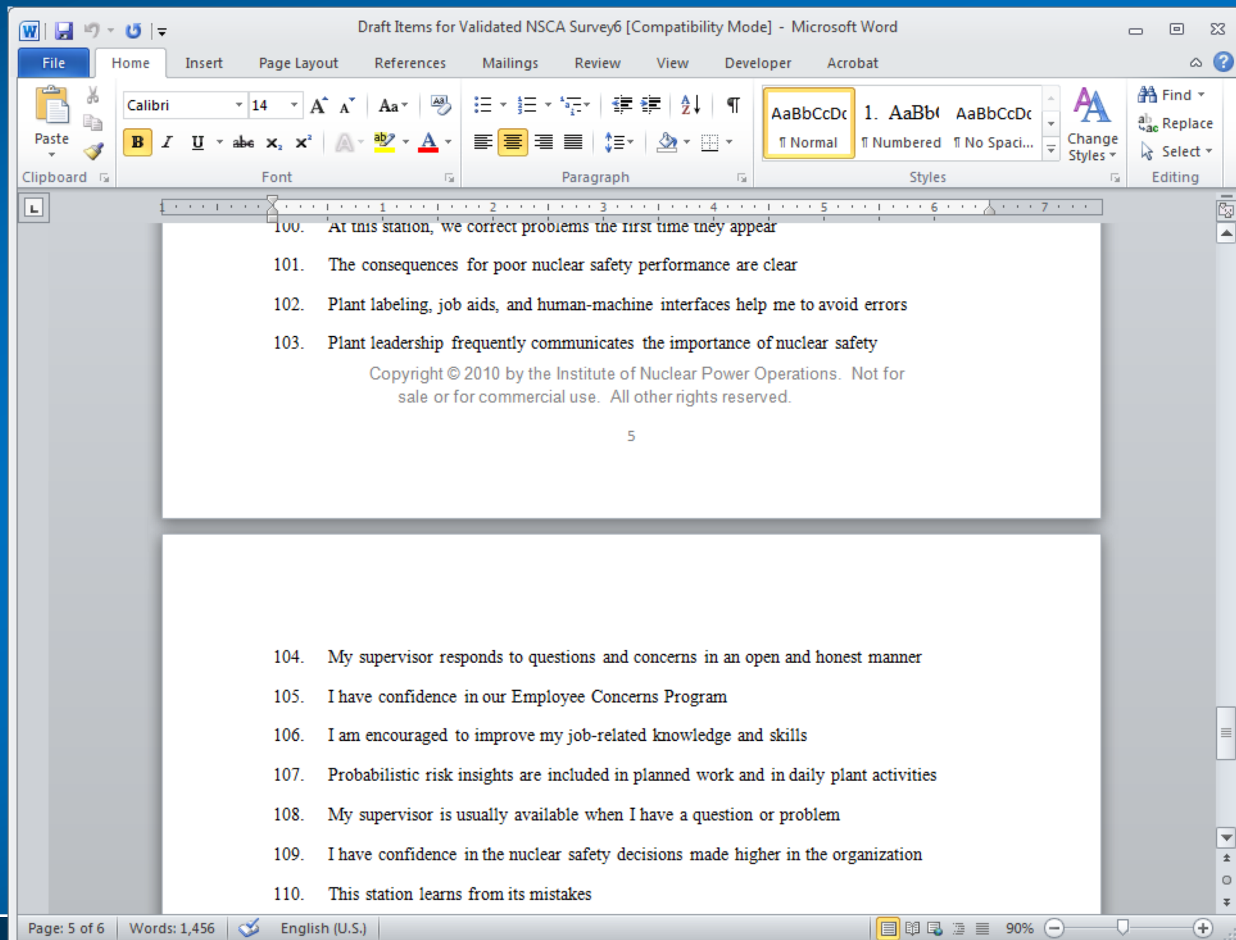
71. Management acts decisively when a nuclear safety concern is raised

72. Nuclear safety is ~~Controlling~~ more important than controlling costs ~~is more important than~~

Page: 4 of 6 Words: 1,474 English (U.S.)

Development – Finally, a Final Survey of 110 Questions

- 51% more items than original



100. At this station, we correct problems the first time they appear
101. The consequences for poor nuclear safety performance are clear
102. Plant labeling, job aids, and human-machine interfaces help me to avoid errors
103. Plant leadership frequently communicates the importance of nuclear safety

Copyright © 2010 by the Institute of Nuclear Power Operations. Not for sale or for commercial use. All other rights reserved.

5

104. My supervisor responds to questions and concerns in an open and honest manner
105. I have confidence in our Employee Concerns Program
106. I am encouraged to improve my job-related knowledge and skills
107. Probabilistic risk insights are included in planned work and in daily plant activities
108. My supervisor is usually available when I have a question or problem
109. I have confidence in the nuclear safety decisions made higher in the organization
110. This station learns from its mistakes

Survey Administration

- Online survey
- Administered by a vendor
- 7-point scale (strongly disagree to strongly agree)
- ‘Don’t Know’ option
- Randomly selected sample of 100 personnel from each site

INPO

Survey Response

- 63 nuclear reactor sites participated (97% of industry)
- An average of 46 individuals participated from each site
- 2,876 individuals provided valid responses to the majority of items

INPO

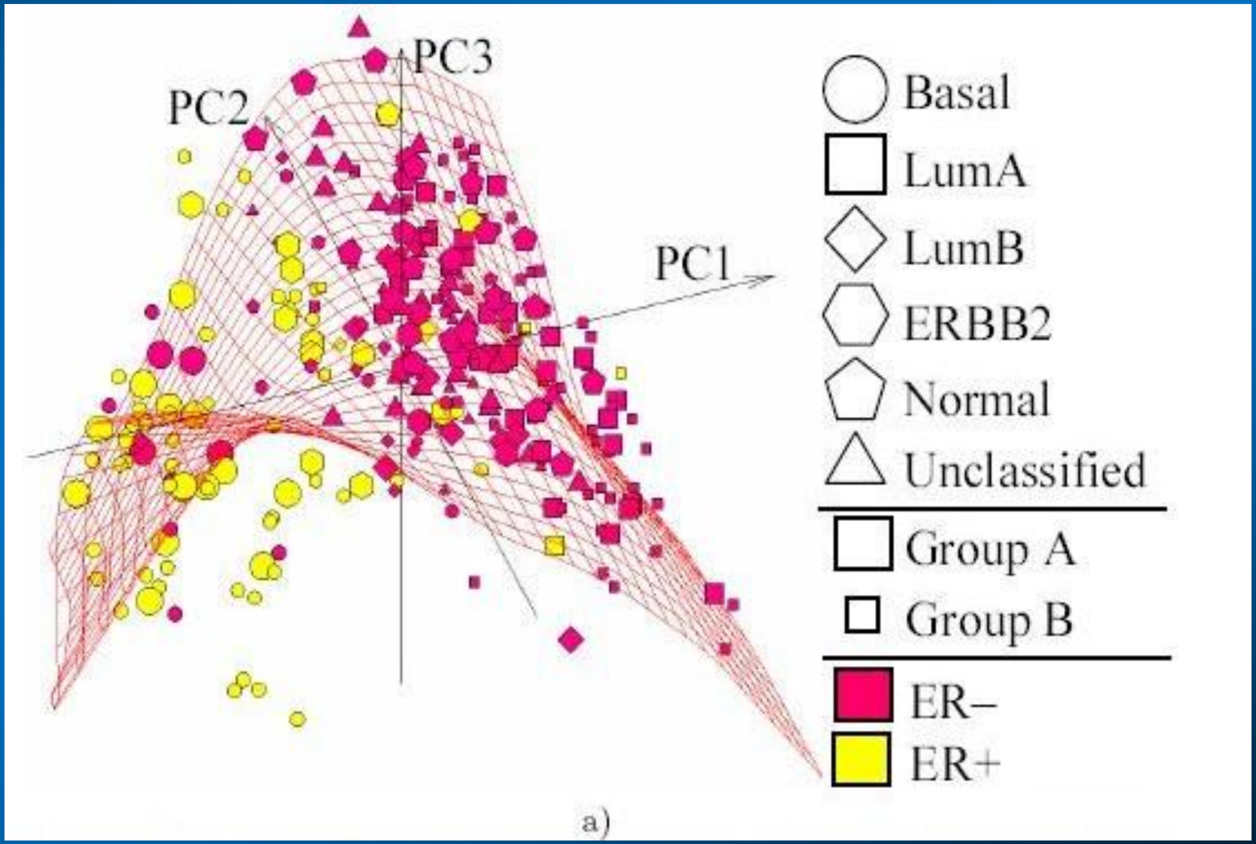
Principal Components Analysis – Why?

- Find patterns (underlying structure) in multidimensional data
- Data reduction
- Build better scales
- Often used synonymously with the term ‘factor analysis’

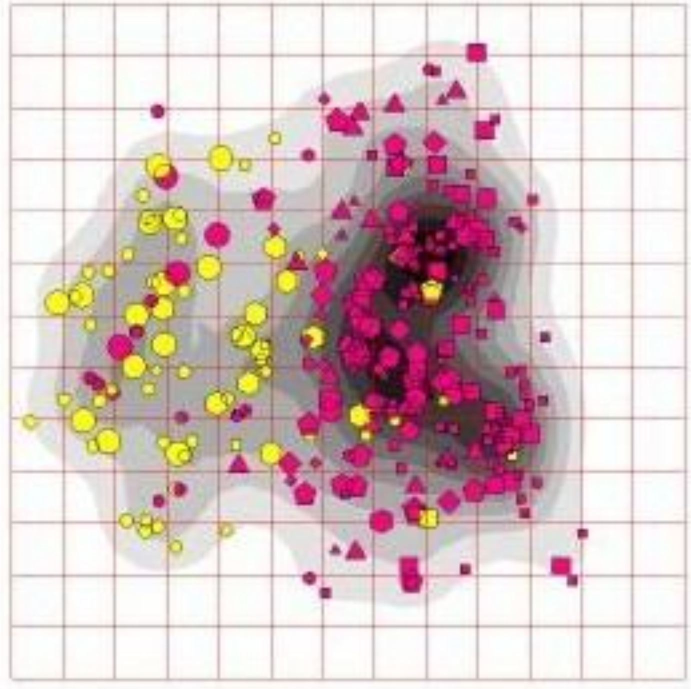
INPO

Principal Components Analysis – How?

- 'Clouds' of data in multidimensional space

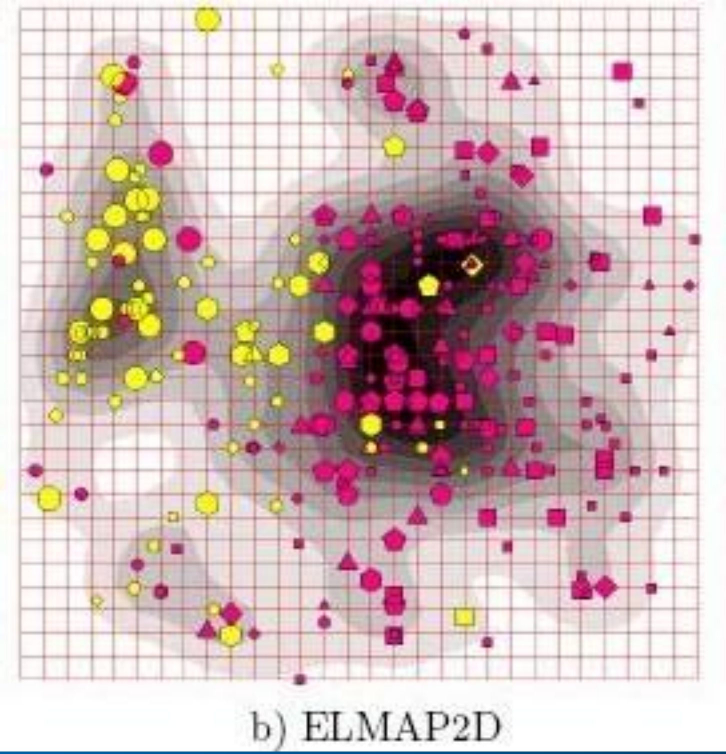


Principal Components Analysis – A Visualization



c) PCA2D

Principal Components Analysis – A Visualization



Principal Components Analysis – How?

- Enter the data into the statistics package (SPSS,SAS, R, etc.)
- Determine the number of components
  - Eigenvalue > 1
  - Scree plot
  - Coherence
- Label the groups of items
- Iterative process

Principal Components Analysis – How?

- Labeling groups of items

95 Dialogue and debate are encouraged	.390	.434	.436	.184	.157	.143	.134		.195		
26 My supervisor gives me useful feedback	.273		.172	.727	.111	.119		.147			
8 My supervisor observes me working	.188			.692	.198		.206	.125			
108 My supe is usually available when I have a question or problem	.254	.230	.204	.673						.288	
24 My supervisor has personally recognized me	.258		.184	.628		.149			.202	-.119	
104 My sup responds to questions in an open and honest manner	.204	.244	.354	.597	.104	.146				.222	.140
57 My sup discusses safety before I start work on a job	.122	.244	.114	.596	.230	.180	.168	.183			
82 When I need a decision, my management is usually available	.388	.285	.214	.546	.109			.106		.271	
85 Supervisors are responsive to employee questions	.316	.316	.359	.484	.179					.217	.117
41 Supervisors are visible in the plant	.404	.286		.471	.296	.134					.197
35 My sup supports senior mgt policies	.259	.290	.119	.368	.119	.311		.196			.207
50 Co-workers hold one	.278	.124	.183	.204	.636	.118			.191		.155

Principal Components Analysis – How?

- Labeling groups of items

mgt policies											
50 Co-workers hold one another to high standards	.278	.124	.183	.204	.636	.118			.191		.155
21 Workers usually follow procedures	.211			.139	.630	.229	.144				
65 Personnel do not proceed in the face of uncertainty	.246	.294	.151	.154	.610	.180	.122				
92 Workers maintain a questioning attitude	.251	.359	.233	.164	.561	.164		.165	.146	.149	
56 Workers follow procedures and make conservative decisions	.234	.275	.205	.243	.521	.188	.141	.146		.146	
55 Personal conflicts are not allowed to interfere	.317	.280	.335	.203	.372	.153		.114	.159		
75 It is my responsibility to raise nuclear safety concerns		.128	.214			.771	.105			.105	
70 It is my responsibility to report security concerns		.145	.148			.735	.161				
40 I am personally responsible for nuclear safety		.239	.137	.101		.687	.163				.169
43 I always use human error prevention techniques		.164		.141	.225	.641		.100			
49 If a procedure is incorrect, I report the problem			.136	.130	.297	.619			.117		-.203
29 Security is just as important as safety	.265	.170		.123	.108	.313		.138	.108		.187
1 Nuclear safety is routinely	.169	.192	.115	.117		.269	.624			.133	

Principal Components Analysis – How?

- Coherence in identifying factors

Rotated Component Matrix[a]

9 Train

10

	Component										
	1	2	3	4	5	6	7	8	9	10	11
45 My managers ensure high-quality training	.417			.344					.545		
17 Training helps understand how I contribute to safety	.364							.315	.495		
78 Continuous learning is expected of everyone						.368			.460		
93 Training at this station reinforces safe worker behaviors	.253		.271		.306	.342			.425		
34 New personnel know the differences working at a nuclear site	.271		.265						.285	.275	
18 I know how to enter an issue into CAP							.327			.588	
31 Mgt oversight is provided for safety-significant tests	.281		.322			.311				.354	
19 Safety culture assessment in the past two years	.265									.299	
58 This station has a knowledgeable and experienced workforce	.276				.343						.35

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 13 iterations.

Principal Components Analysis – How?

- Redo process on factors with large groups of items to look for subfactors

Survey Results

- Management Responsibility
  - Respectful Work Environment
  - Continuous Improvement
  - Performance Indicators
  - Resources
  - Rewards
- Willingness to Raise Concerns
  - Informally
  - Formally

INPO

'—' = Subdimension

Results

- Decision Making
  - Decisions are conservative, timely, safety-focused, and engender confidence
- Supervisor Responsibility
  - Communication
  - Presence/Availability
  - Coaching
  - Management Alignment

INPO

'—' = Subdimension

Results

- Questioning Attitude
  – Situation/Problem Awareness
  – Process Use
  – Plant Knowledge
- Safety Communication
  - ➢Safety communication is broad and includes plant-level communication, job-related communication, worker-level communication, equipment labeling, operating experience, and documentation

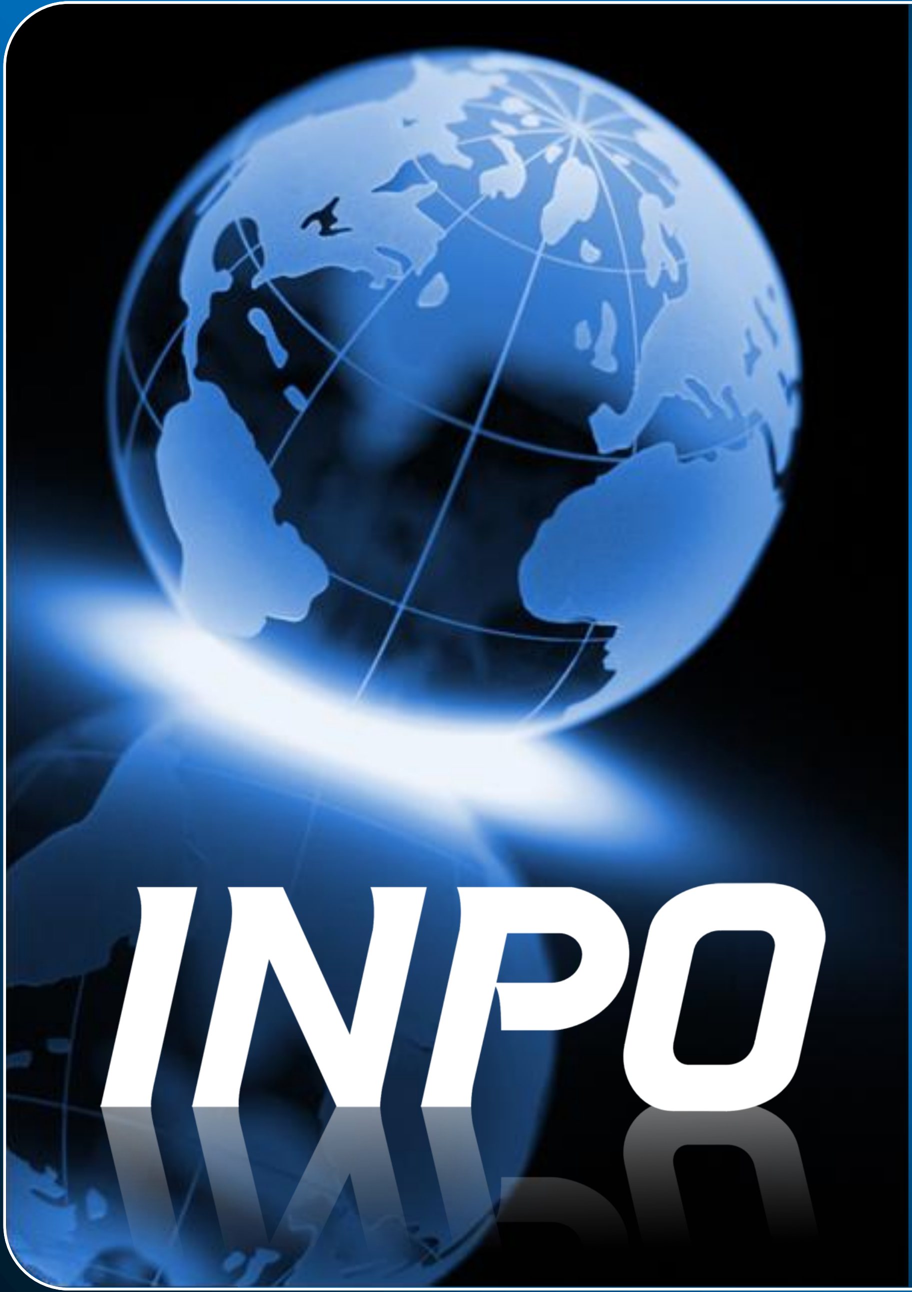
INPO

Results

- Personal Responsibility
  - It is my responsibility to report concerns and practice nuclear safety
- Prioritizing Safety
  - Nuclear safety is a priority that is seen in meetings, expectations, coaching, and decisions
- Training Quality
  - Training is high quality, supported by management and encourages nuclear safety

Item Reduction

- Use multiple sources of information to make a judgment about each item
  - Number of missing data points
  - Reliability of the scale with the item removed
  - Inter-item correlations
  - Item content
  - Correlations with KPIs



Questions and Discussion

G. Kenneth Koves, Ph.D.

koveskg@inpo.org
+01 770 644 8303

INPO Safety Culture Survey and Construct Validation Study

Presentation to the ACRS Subcommittee on Reliability and Probabilistic Risk Assessment

May 8, 2012

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

Presenter Biographies

- Dr. Valerie Barnes
  - Senior Technical Advisor for Human Factors, RES/DRA
  - Ph.D. in Social Psychology, University of Washington
- Dr. Stephanie Morrow
  - Human Factors Analyst, RES/DRA/HFRB
  - Ph.D. in Industrial/Organizational Psychology, University of Connecticut
- Dr. G. Kenneth Koves
  - Senior Program Manager, INPO
  - Ph.D. in Industrial/Organizational Psychology, Georgia Tech

Agenda

Presentation	Presenter
Recent Developments in Safety Culture Research	Dr. Valerie Barnes, RES
INPO Survey Development and Factor Structure	Dr. Ken Koves, INPO
Independent Evaluation of INPO Survey and Alignment with Safety Culture Policy Statement	Dr. Stephanie Morrow, RES
Validation of INPO Survey with INPO Performance Metrics	Dr. Ken Koves, INPO
Validation of INPO Survey with NRC Performance Metrics	Dr. Stephanie Morrow, RES

Recent Developments in Safety Culture Research

Dr. Valerie Barnes, Ph.D.

RES/DRA

A Brief History

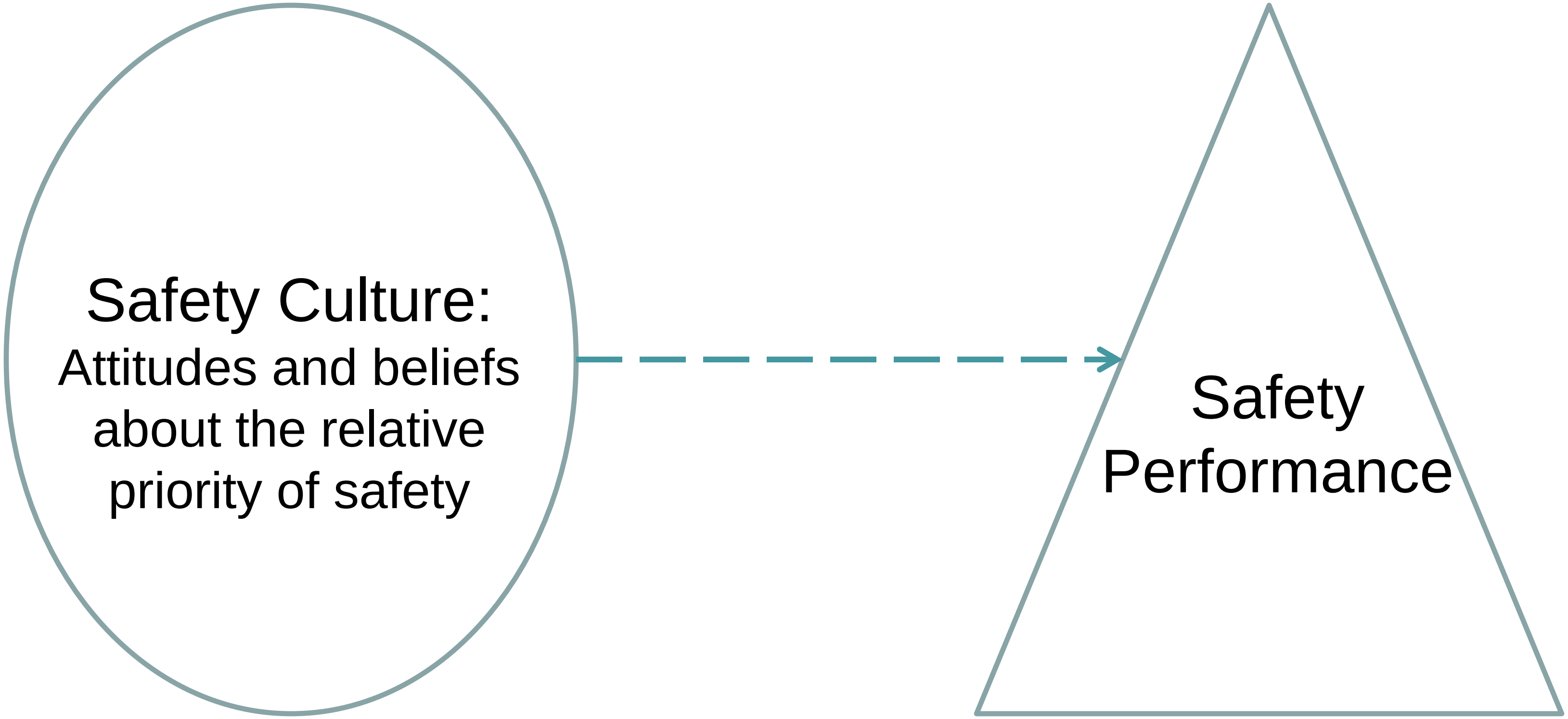
United States Nuclear Regulatory Commission
Protecting People and the Environment

- Organizational factors are important to safety.
- Concept of “safety culture” has improved research focus.
- First empirical study in 1980 – correlations were slight.
- Theory, research and utility continue to advance.

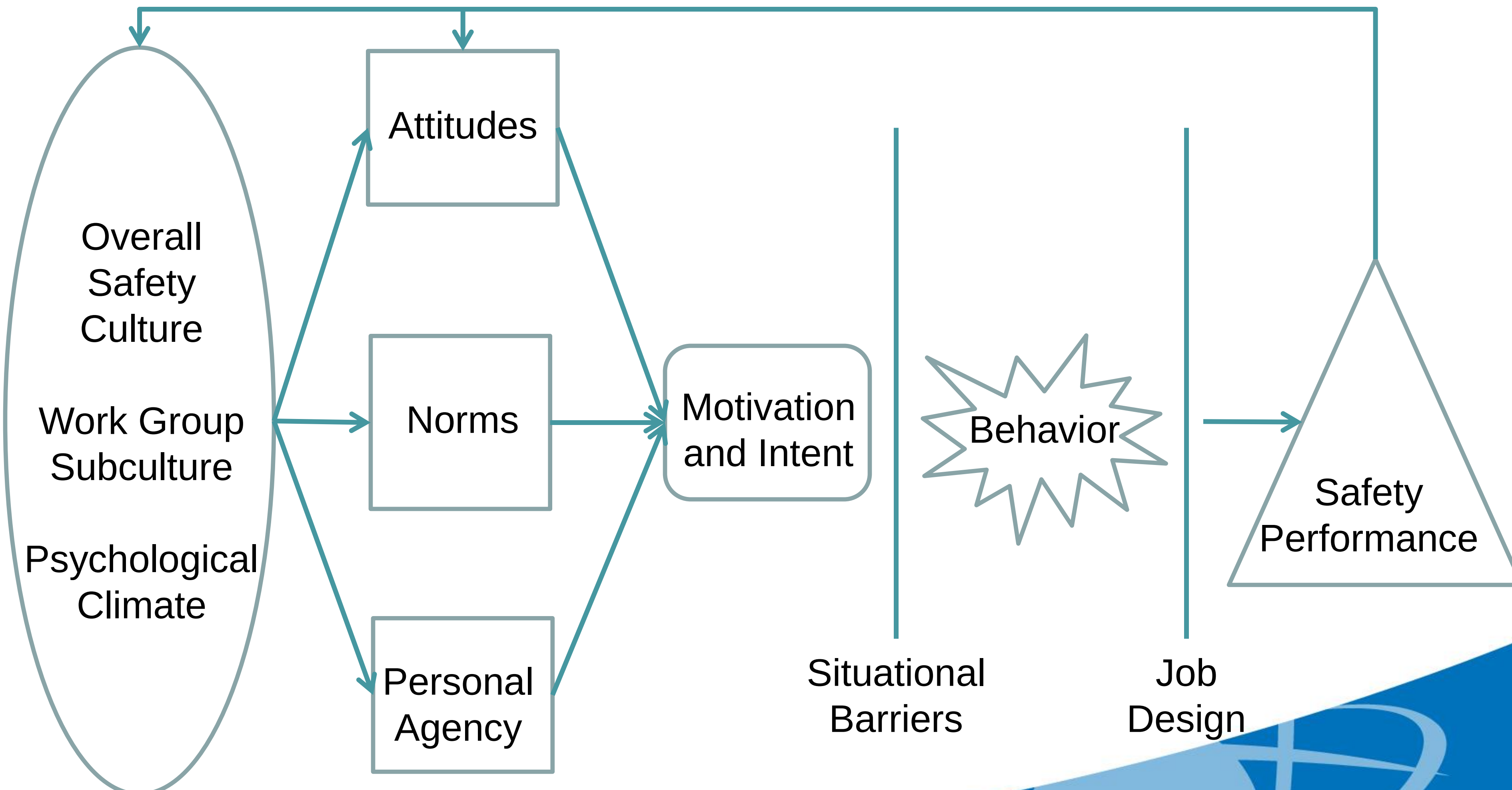
What is the theory underlying the empirical research?

- The attitudes and beliefs of members of an organization provide insights into the organization's safety culture
  - Culture is a social phenomenon that naturally arises among members of a definable group (e.g., an organization)
  - Members learn/experience it in their interactions with others within the group, and
  - Individuals develop attitudes and beliefs about the organization's values and norms from their social interactions, which can be measured
- These views about "what's important around here" (values) and "how we do things around here" (norms) affect individuals' safety-related decisions and actions in the workplace
- Workplace decisions and actions (human behavior) affect the organization's safety performance

In The Beginning…



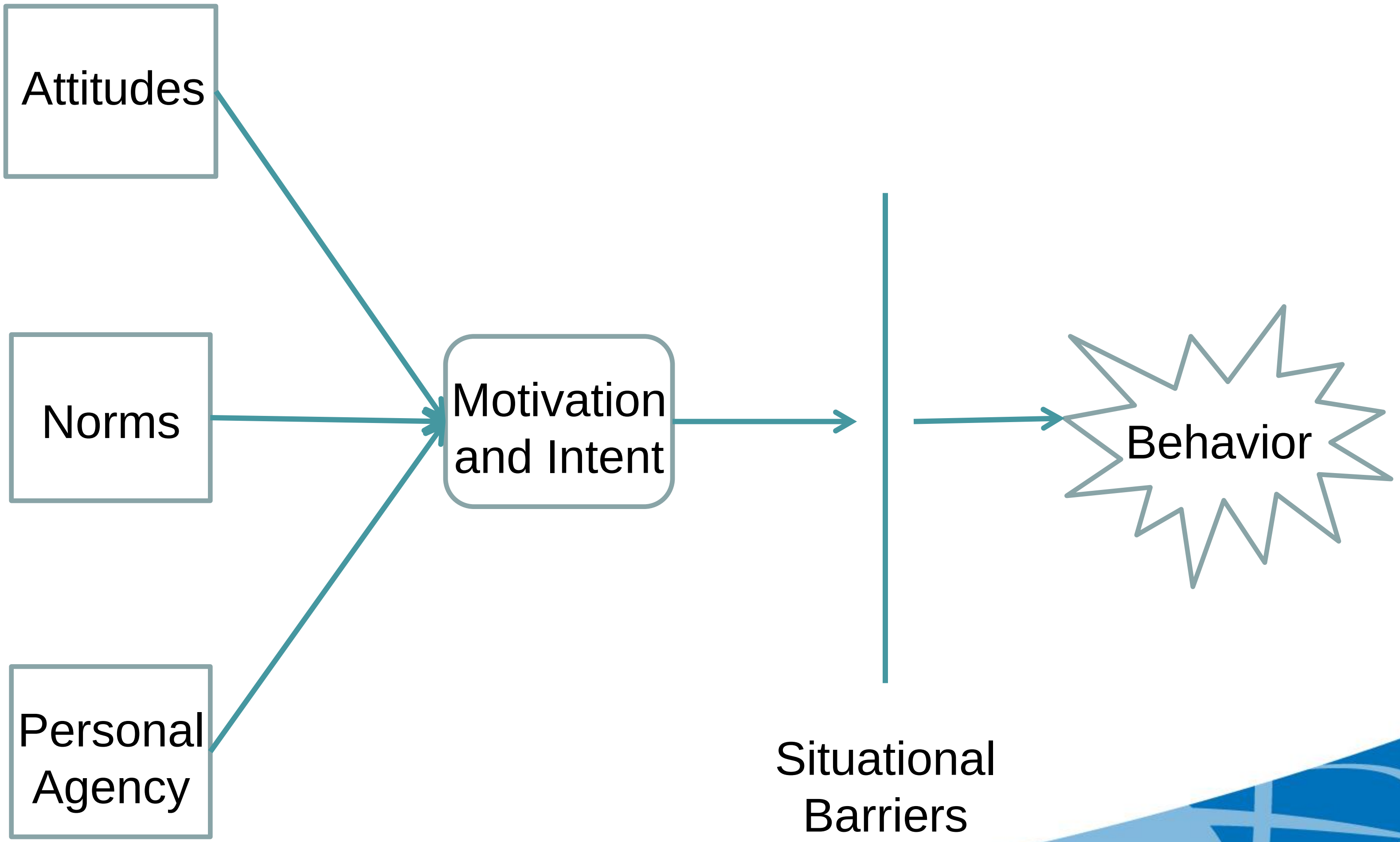
A Bit More Complicated



Attitude/Behavior Relationship

- Medium effect size - ~.30
- Strength of relationship moderated by
  - Specificity of the attitude and behavior
  - Social influences (norms)
  - Personal characteristics of individuals
  - Situational characteristics

Components of Attitude/Behavior Relationship

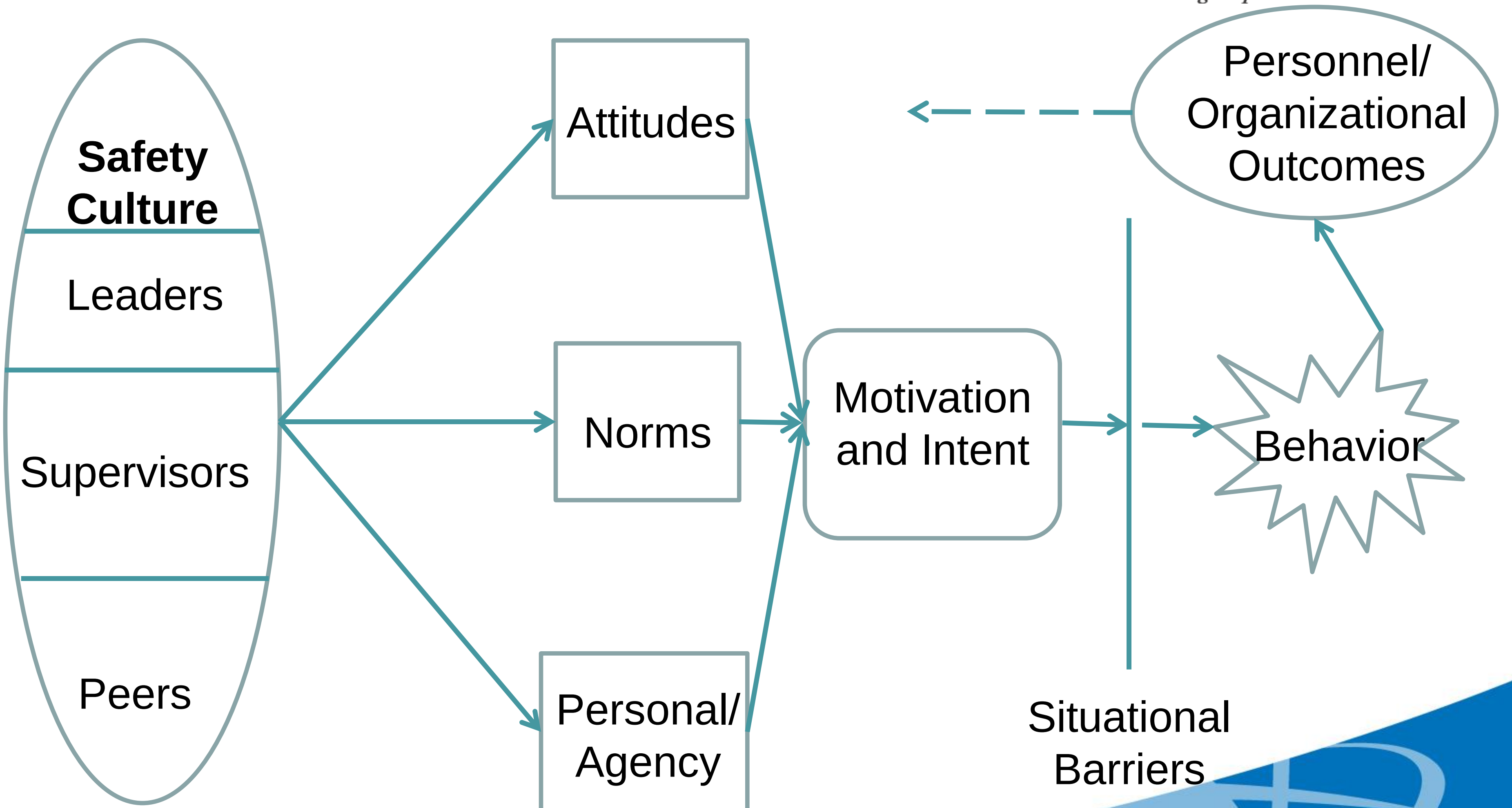


What Does This Have to do With Safety Culture?

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

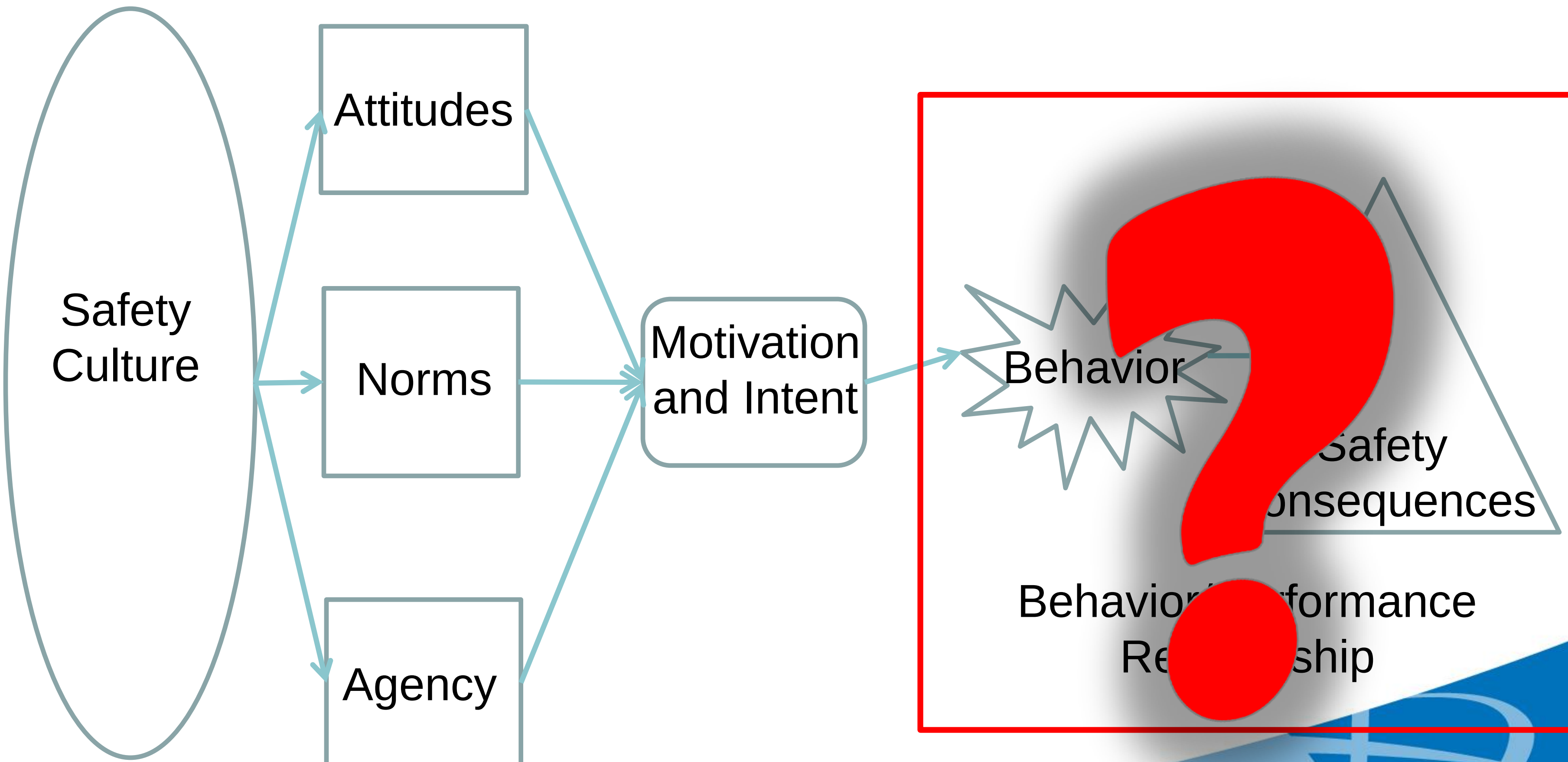
- Useful for understanding the influences on individuals' workplace safety-related behaviors.
- But, what shapes these attitudes, norms and personal agency beliefs at work?
- Concept of safety culture provides some answers.

Impact of Safety Culture on Attitude/Behavior Relationship

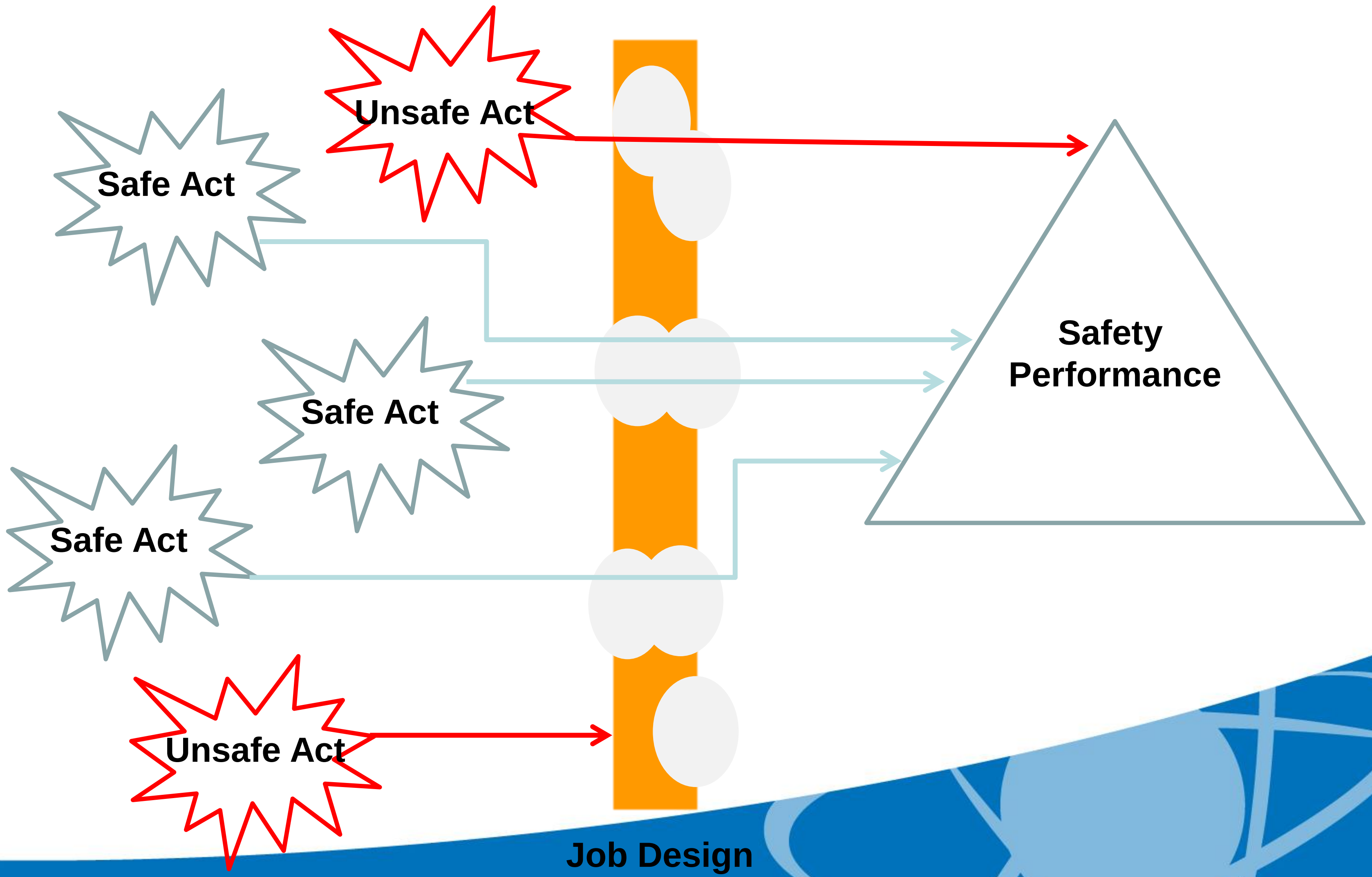


Another Complication: The Relationship of Behavior to Safety Performance

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment



Job Design Affects Behavior/ Safety Performance Relationship



Measuring Safety Performance

- Different categories of measures
  - Occupational safety
  - Patient safety
  - Near-misses/microaccidents
  - Human error/unsafe acts
  - Reports of injuries, errors, near-misses
- Different levels of analysis
  - Individual (self-reports, observations)
  - Work-groups (hospital units, departments)
  - Whole organizations (companies, hospitals)

Effect Sizes

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

Level of Analysis	Safety Behavior	Safety Performance
Individuals	Medium-Large	Depends on type of tasks
Work groups	Medium-Large	Medium-Large
Organizational	Small	Small-Medium

In Summary…

The safety culture/safety performance relationship is complex. The strength of the relationship depends on a variety of intervening factors as well as how both safety culture and safety performance are measured.

How are Insights About Safety Culture Being Used?

- Internal and external benchmarking
- Identifying areas for improvement
- Informing the design of safety interventions
- Checking the effectiveness of safety interventions

Example Intervention Results

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

- 30% decrease in unsafe acts during electrical work in an oil refinery (Zohar, 2007)
- 50% increase in safe housekeeping acts at a steel company (Zohar, 2007)
- Decrease in ICU patient bloodstream infection rates from 2.8/1000 to 0 (Pronovost, et al, 2010)
- 70% decrease in ventilator-associated pneumonia rates in Michigan hospitals (Pronovost, et al, 2012)

Expectations for INPO Study

- Design appropriate for purpose – to validate the safety culture workshop traits, but
- Expected small effect sizes with performance measures because
  - Survey participant samples from each site too small to permit analyses at the work-group level
  - Safety performance measures only available at the overall organizational level

Independent Evaluation of INPO Survey and Alignment with Safety Culture Policy Statement

Dr. Stephanie Morrow, Ph.D.
RES/DRA/HFRB

Purpose of White Paper

- Present results of RES independent evaluation of the INPO safety culture survey and construct validation study
  - Is the survey valid and reliable?
  - Does the survey show support for the Safety Culture Policy Statement traits?
  - Is there a relationship between the safety culture survey results and safety performance metrics?

Key Findings

- The INPO Safety Culture Survey demonstrates evidence of construct validity and reliability in the context of this study.
- Although there was not one-to-one alignment, the INPO Safety Culture Survey factors and NRC's Safety Culture Policy Statement traits shared many commonalities. Each trait was represented by one or more survey factors.

How Do We Evaluate Validity and Reliability?

- **Construct Validity**: Is the survey measuring what it purports to measure?
  - **Content validity**: Does the survey cover the breadth of the construct?
  - **Criterion-related validity**: Does the survey demonstrate a relationship with outcomes to which it should (theoretically) be related?
- **Reliability**: Does the survey measure the construct consistently?
  - **Internal consistency**: Do the items in the survey produce similar results?
  - **Within-group reliability**: Do respondents from the same groups have similar results?

Additional Evaluation Questions

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

- Were the data collection procedures and resulting sample appropriate for the research questions?
- Is the data analysis approach consistent with good practice in the social sciences?
- Given the data, would an independent researcher produce the same results and arrive at similar conclusions?

Content Validity of INPO Survey Items

- Survey items drawn from multiple sources (INPO, NRC, IAEA, non-nuclear research literature)
- 7 point Likert-type rating scale used (strongly disagree to strongly agree)
- Items are simply written, address single topics, and avoid double-barreled statements
  - "Our leadership frequently communicates the importance of nuclear safety"
  - "Our corrective action program is effective"
  - "Management acts decisively when a nuclear safety concern is raised"
  - "Continuous learning is expected of everyone"
  - "My supervisor discusses safety with me before I start work on a job"
  - "Staffing levels are adequate to meet work demands"
  - "Station management gives us clear direction"

Data Collection and Sample Characteristics

- Web-based survey
- Cross-section of nuclear power industry
  - 97% of operating nuclear power plants participated
  - 100 individuals per site randomly selected to take survey
    - 48% response rate, and at least 30 individuals per site
    - Final Sample: 2,876 respondents across 63 sites
- Representative of different occupational groups, for example…
  - 17% Maintenance
  - 16% Operations
  - 10% Security
  - 6% Radiation Protection
  - 5% Systems Engineering
- Inclusion of contractors
  - 7% Contractors

Factor Analysis Approach

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

- Principal Components Analysis (PCA)
  - 9 Interpretable Factors
    - Most Stable Factor: Management Responsibility/Commitment
    - Least Stable Factor: Safety Communication

Table 1 Results of PCA of 110 Items with a 9 Factor Solution

Factor Label	% Variance Accounted For	# Items
1. Management Responsibility/Commitment to Safety	15.7%	36
2. Willingness to Raise Concerns	6.9%	9
3. Decision-Making	6.3%	10
4. Supervisor Responsibility for Safety	6.2%	11
5. Questioning Attitude	5.9%	9
6. Safety Communication	5.3%	13
7. Personal Responsibility for Safety	4.8%	6
8. Prioritizing Safety	4.0%	6
9. Training Quality	2.7%	6

Factor Labeling

Factor/Sub-Factor Label	Example Item
1. **Management Responsibility/Commitment to Safety**	At this station, people are routinely rewarded for identifying and reporting nuclear safety issues
2. **Willingness to Raise Concerns**	When I make a mistake, I'm not afraid to report it to my supervisor
3. **Decision-Making**	Decision-making at this site reflects a conservative approach to nuclear safety
4. **Supervisor Responsibility for Safety**	My supervisor is usually available when I have a question or problem
5. **Questioning Attitude**	Personnel promptly identify and report conditions that can affect nuclear safety
6. **Safety Communication**	There is good communication about nuclear safety issues that affect my job
7. **Personal Responsibility for Safety**	It is my responsibility to raise nuclear safety concerns
8. **Prioritizing Safety**	At this station, nuclear safety takes priority over production goals
9. **Training Quality**	Training at this site provides me with the knowledge I need to perform my job

Reliability of INPO Survey Factors

- Cronbach's Alpha (α) used to measure internal consistency of items
- Values ≥ .70 indicate good reliability

Meta-factor/Factor/Sub-factor Label	Cronbach's α	# Items
SAFETY CULTURE	0.98	60
1. Management Responsibility/ Commitment to Safety	0.96	20
2. Willingness to Raise Concerns	0.90	6
3. Decision-Making	0.88	5
4. Supervisor Responsibility for Safety	0.88	6
5. Questioning Attitude	0.85	6
6. Safety Communication	0.87	7
7. Personal Responsibility for Safety	0.77	3
8. Prioritizing Safety	0.83	4
9. Training Quality	0.78	3

Within-Group Reliability

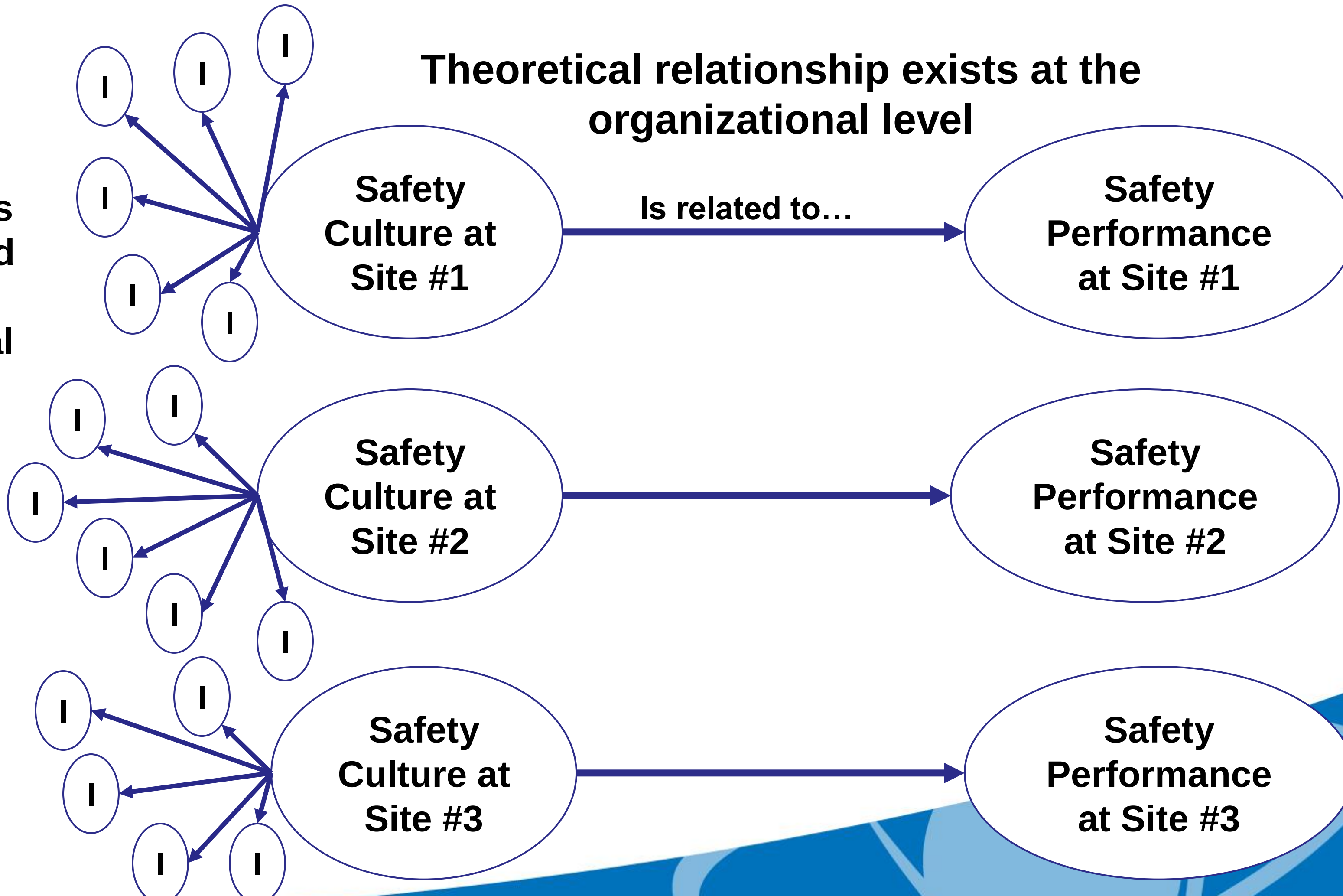
- Evaluated with Intra-Class Correlations (ICCs)
  - ICC(1) = extent to which individuals at a site had the same responses
  - ICC(2) = internal consistency of the mean score for the site

Table 4 Range of ICC(1) and ICC(2) Values for Safety Culture and Safety Culture Factors at Each Site

	ICC(1)		ICC(2)	
	Range	**Sig.**	**Range**	**Sig.**
SAFETY CULTURE	0.26 - 0.58	$p < .01$	0.95 - 0.99	$p < .01$
1. Management Responsibility/ Commitment to Safety	0.33 - 0.63	$p < .01$	0.91 - 0.97	$p < .01$
2. Willingness to Raise Concerns	0.38 - 0.81	$p < .01$	0.79 - 0.96	$p < .01$
3. Decision-Making	0.32 - 0.79	$p < .01$	0.70 - 0.95	$p < .01$
4. Supervisor Responsibility for Safety	0.31 - 0.72	$p < .01$	0.73 - 0.94	$p < .01$
5. Questioning Attitude	0.25 - 0.66	$p < .01$	0.66 - 0.92	$p < .01$
6. Safety Communication	0.24 - 0.65	$p < .01$	0.69 - 0.93	$p < .01$
7. Personal Responsibility for Safety	0.28 - 0.88	$p < .01$	0.54 - 0.96	$p < .01$
8. Prioritizing Safety	0.30 - 0.76	$p < .01$	0.63 - 0.93	$p < .01$
9. Training Quality	0.24 - 0.73	$p < .01$	0.49 - 0.89	$p < .01$

Why is Within-Group Reliability Important?

Safety Culture is measured at the individual level



Alignment of INPO Survey Factors and Policy Statement Traits

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

- Are the factors identified in the INPO Safety Culture Survey similar to the traits identified in the NRC Safety Culture Policy Statement?
  - Review factor labels and specific items within each factor
  - Compare factors/items to definitions of policy statement traits

Safety Culture Policy Statement Traits

- Leadership Safety Values and Actions
- Problem Identification and Resolution
- Personal Accountability
- Work Processes
- Continuous Learning
- Environment for Raising Concerns
- Effective Safety Communication
- Respectful Work Environment
- Questioning Attitude

Policy Statement/INPO Survey Crosswalk

Policy Statement Trait	INPO Survey Factor(s)
Leadership Safety Values and Actions	**1. Management Responsibility** **3. Decision-Making** **4. Supervisor Responsibility**
Problem Identification & Resolution	**1. Management Responsibility** (b. Continuous Improvement) **3. Decision-Making**
Personal Accountability	**7. Personal Responsibility**
Work Processes	**1. Management Responsibility** (d. Procedure Communication and e. Resources)
Continuous Learning	**1. Management Responsibility** **9. Training Quality**
Environment for Raising Concerns	**2. Willingness to Raise Concerns**
Effective Safety Communication	**6. Safety Communication**
Respectful Work Environment	**1. Management Responsibility** (a. Respectful Work Environment)
Questioning Attitude	**5. Questioning Attitude**

Key Observations

- Management Responsibility/Commitment to Safety
  - Accounts for most items and most variance in survey results
  - Similar to multiple traits in the Policy Statement
- Supervisor Responsibility for Safety
  - Separate factor accounting for supervisor's commitment and responsibility for safety
- Decision-Making
  - Elements of Leadership and Problem Identification & Resolution

Summary

- The design of the INPO construct validation study was appropriate for the research questions.
- The INPO survey demonstrated evidence of reliability and validity in this study.
- There were common themes between the factors that emerged from the INPO survey and the traits identified in the Safety Culture Policy Statement.

Questions & Discussion

Validation of INPO Survey with NRC Performance Metrics

Dr. Stephanie Morrow, Ph.D.
RES/DRA/HFRB

Key Findings

- **Based on accepted standards in the social sciences**, there are moderate, statistically significant correlations between the safety culture survey results and some NRC performance metrics measured during the same time period.
- Moderate, statistically significant correlations exist between the safety culture survey results and broad-based NRC performance metrics measured 1 year after the survey (e.g., elevated oversight in the Action Matrix, count of Allegations)

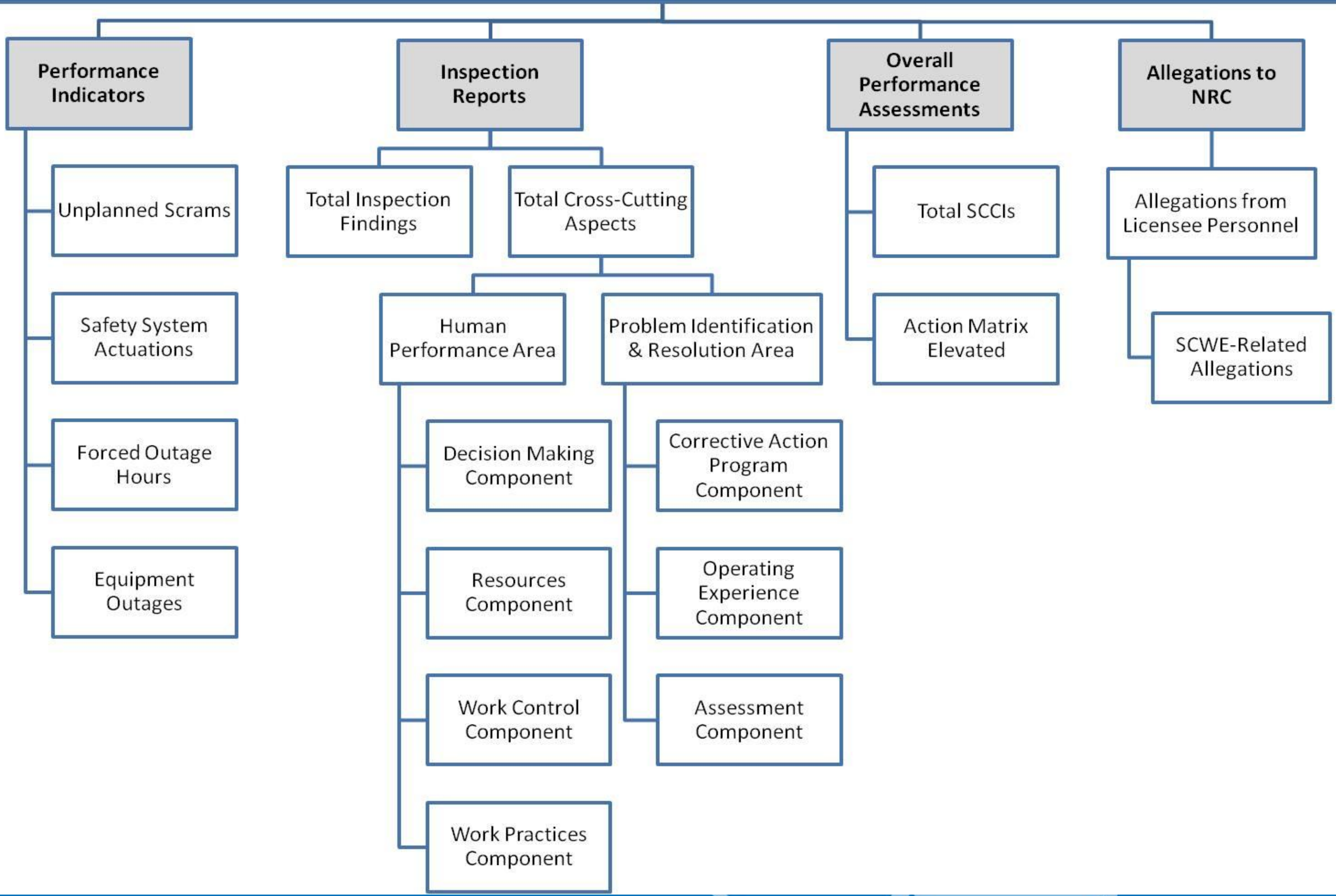
Limitations of the Research

- **Single study:** Need to replicate to show reproducibility and generalizability of results.
- **Correlational:** Cannot establish causal effects.
- **Cross-sectional:** Comparisons made between NPP sites, rather than within a single site over time.
- **Only nuclear power plants:** Results may not be generalizable to other regulated communities.

NRC Performance Metrics

- Performance Indicators
  - Unplanned Scrams
  - Safety System Actuations
  - Forced Outage Hours
  - Equipment Outages
- Inspection Findings
  - Total Inspection Findings
  - Cross-Cutting Areas, Components, and Aspects
- Overall Performance Assessments
  - Elevated Oversight in Action Matrix
  - Total Substantive Cross-Cutting Issues (SCCIs)
- Allegations made to the NRC
  - Allegations from Licensee Personnel
    - Safety Conscious Work Environment (SCWE)-related Allegations

NRC Performance Metrics



Background: Pearson's Correlations

- Describes the degree of association between two variables
- Most common statistic used to describe a relationship, chosen 95% of the time in research (Glass & Hopkins, 1996)
- Range of values from +1.0 to -1.0
  - Values close to +1.0 indicate a strong positive relationship
  - Values close to -1.0 indicate a strong negative relationship
  - Values close to 0 indicate no relationship
- Assumes variables are measured on an interval or ratio scale and are normally distributed
- Sensitive to outliers

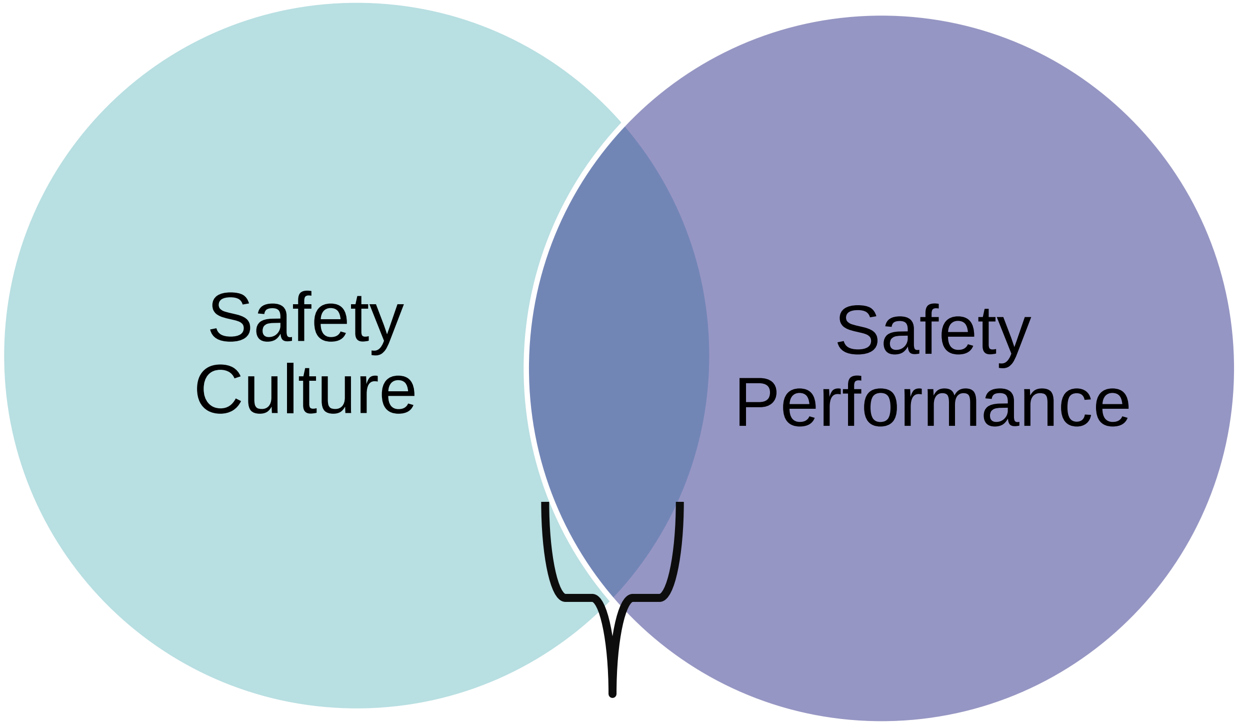
What is Standard in the Social Sciences?

- Effect size rule of thumb (Cohen, 1988)

Statistic	Effect Size	% Variance Shared
.10	Small	1%
.30	Medium	9%
.50	Large	25%

- Meta-analyses of safety culture studies (Christian et al., 2009; Clarke, 2006; Beus et al., 2010)
  – Correlation of **-.22 to -.39** between safety culture surveys and accident/injury rates (medium effect)
  – Correlation of **.43 to .61** between safety culture surveys and self-reported safety behaviors (medium to large effect)

Shared Variance



4% to 37% shared variance
Depends on how safety culture and safety performance are **<u>defined</u>** and **<u>measured</u>**

Concurrent Correlations (2010)

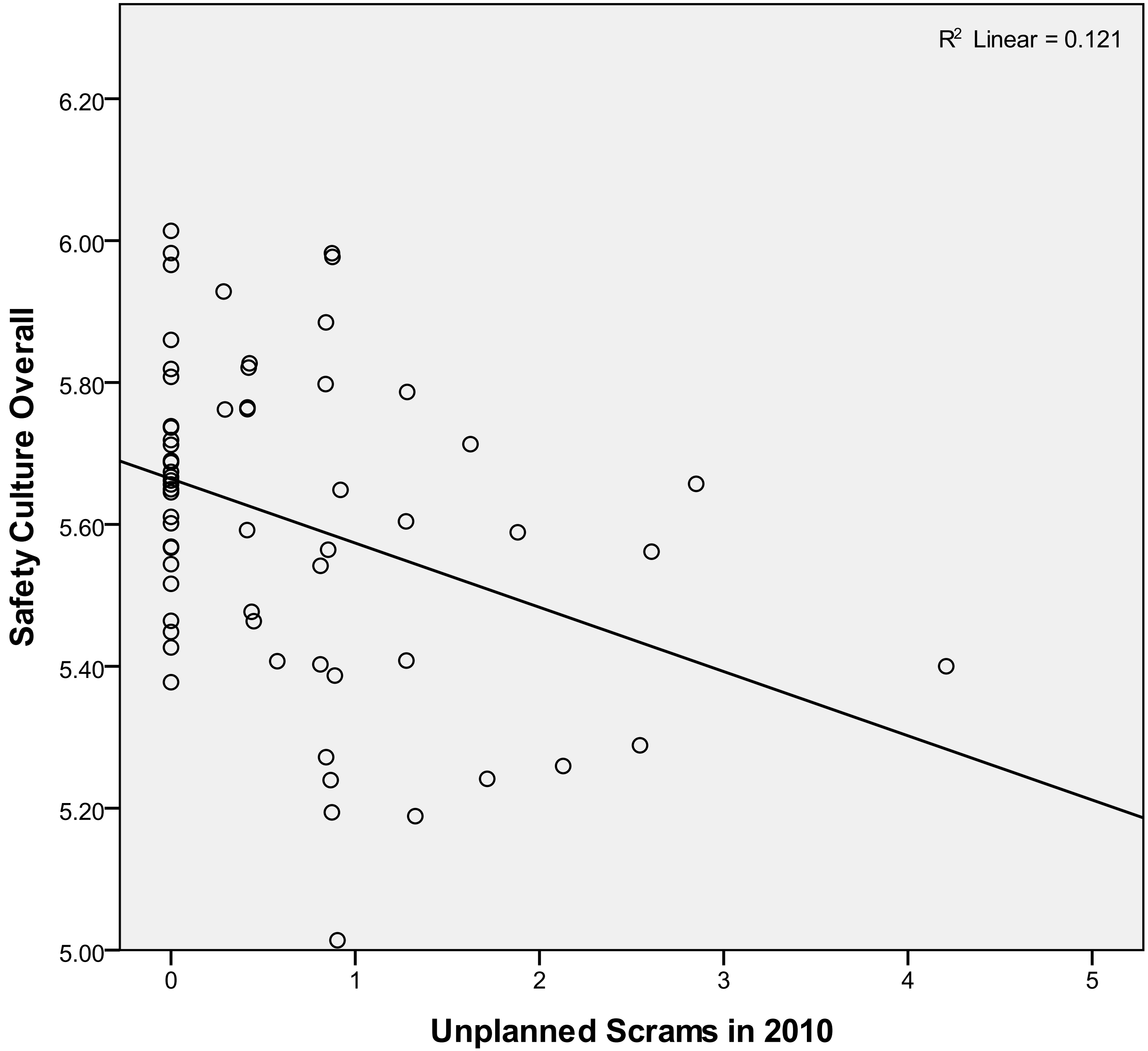
Table 7 Correlations between Safety Culture Survey and 2010 Performance Indicators, ROP Action Matrix, Inspection Findings, and SCCIs

	Unplanned Scrams	Safety System Actuations	Forced Outage Hours	Equipment Outages	Action Matrix Elevated	Total Inspection Findings	Total SCCIs
SAFETY CULTURE	-.35**	.01	-.27*	-.25	-.23	-.37**	-.23
1. Management Responsibility/ Commitment to Safety	-.34**	.03	-.26*	-.22	-.20	-.40**	-.28*
2. Willingness to Raise Concerns	-.29*	.01	-.24	-.30*	-.31*	-.21	-.10
3. Decision-Making	-.33**	.10	-.26*	-.25*	-.27*	-.34**	-.21
4. Supervisor Responsibility for Safety	-.26*	-.07	-.21	-.10	-.27*	-.34**	-.14
5. Questioning Attitude	-.29*	.00	-.22	-.24	-.17	-.41**	-.38**
6. Safety Communication	-.35**	-.04	-.29*	-.22	-.19	-.33**	-.14
7. Personal Responsibility for Safety	-.24	-.05	-.19	-.18	-.28*	.02	.13
8. Prioritizing Safety	-.23	.06	-.16	-0.23	-.11	-.27*	-.11
9. Training Quality	-.46**	-.03	-.39**	-.37**	-.23	-.28*	-.10

*p < .05; **p < .01

Figure 4 Scatterplot of Relationship between Safety Culture Survey and Unplanned Scrams in 2010

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment



Concurrent Correlations (2010)

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

Table 8 Correlations between Safety Culture Survey and 2010 NRC Allegations, ROP Cross-Cutting Areas and Components

	Allegation From Personnel	SCWE-Related Allegation	Total ROP Aspects	Human Perform. Area	HP 1: Decision Making	HP 2: Resources	HP 3: Work Control	HP 4: Work Practices	Problem ID & Res. Area	PI&R 1: CAP	PI&R 2: OE	PI&R 3: Assess-ments
SAFETY CULTURE	-.21	-.28*	-.39**	-.28*	-.27*	-.23	-.01	-.21	-.37**	-.35**	-.10	-.24
1. Management Responsibility/ Commitment to Safety	-.24	-.30*	-.44**	-.30*	-.29*	-.22	-.05	-.23	-.40**	-.40**	-.08	-.23
2. Willingness to Raise Concerns	-.16	-.21	-.21	-.16	-.13	-.15	.09	-.16	-.24	-.22	-.14	-.14
3. Decision-Making	-.17	-.25	-.37**	-.25*	-.23	-.25*	-.02	-.17	-.35**	-.33**	-.13	-.23
4. Supervisor Responsibility for Safety	-.15	-.24	-.28*	-.27*	-.25*	-.25*	.00	-.19	-.24	-.21	-.13	-.30*
5. Questioning Attitude	-.41**	-.48**	-.47**	-.40**	-.36**	-.30*	.04	-.37**	-.37**	-.36**	-.15	-.18
6. Safety Communication	-.12	-.19	-.31*	-.20	-.22	-.20	.05	-.14	-.32*	-.31*	-.06	-.28*
7. Personal Responsibility for Safety	.17	.14	.04	.05	.09	-.12	.02	.09	.00	.03	-.17	.00
8. Prioritizing Safety	-.09	-.15	-.28*	-.18	-.17	-.23	.01	-.09	-.30*	-.29*	.00	-.26*
9. Training Quality	.04	-.05	-.22	-.14	-.20	-.17	.07	-.05	-.25*	-.23	-.08	-.20

*p < .05; **p < .01

Figure 5 Scatterplot of Safety Culture Survey and Total ROP Aspects in 2010

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

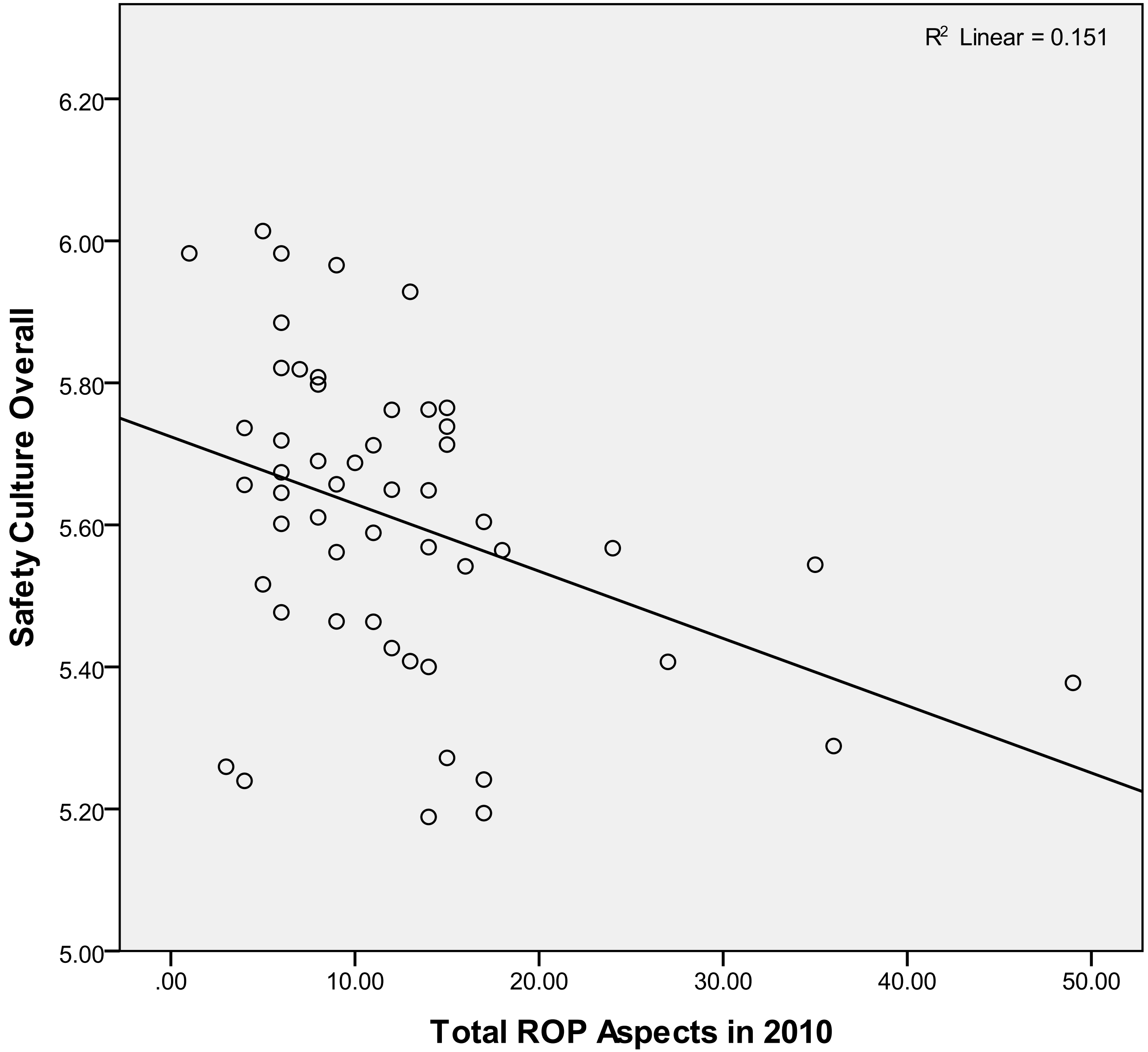
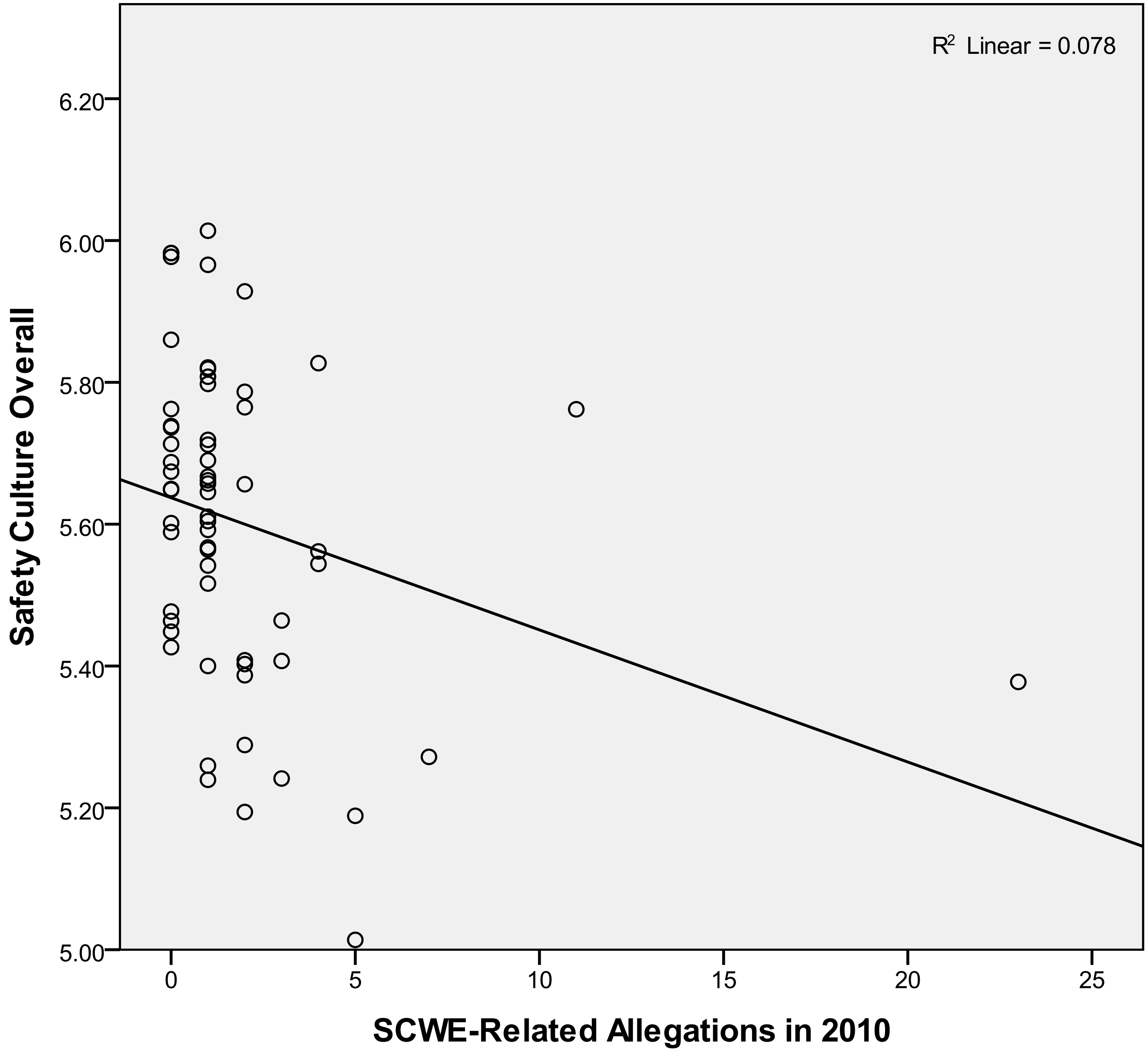


Figure 6 Scatterplot of Safety Culture and SCWE-Related Allegations in 2010



Key Observations: Concurrent Correlations

- Overall Safety Culture survey is moderately correlated with:
  - Unplanned Scrams (-.35)
  - Total Inspection Findings (-.37)
  - Total ROP Aspects (-.39)
    - Problem Identification & Resolution Cross-cutting Area (-.37)
- Questioning Attitude factor is moderately correlated with:
  - Total Inspection Findings (-.41)
  - Total SCCIs (-.38)
  - Allegations from Licensee Personnel (-.41)
    - SCWE-related Allegations (-.48)
  - Total ROP Aspects (-.47)
    - Human Performance Area (-.40)
    - Problem Identification & Resolution Area (-.37)
- Training Quality factor is moderately correlated with:
  - Unplanned Scrams (-.46)
  - Forced Outage Hours (-.39)
  - Equipment Outages (-.37)

Predictive Correlations (2011)

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

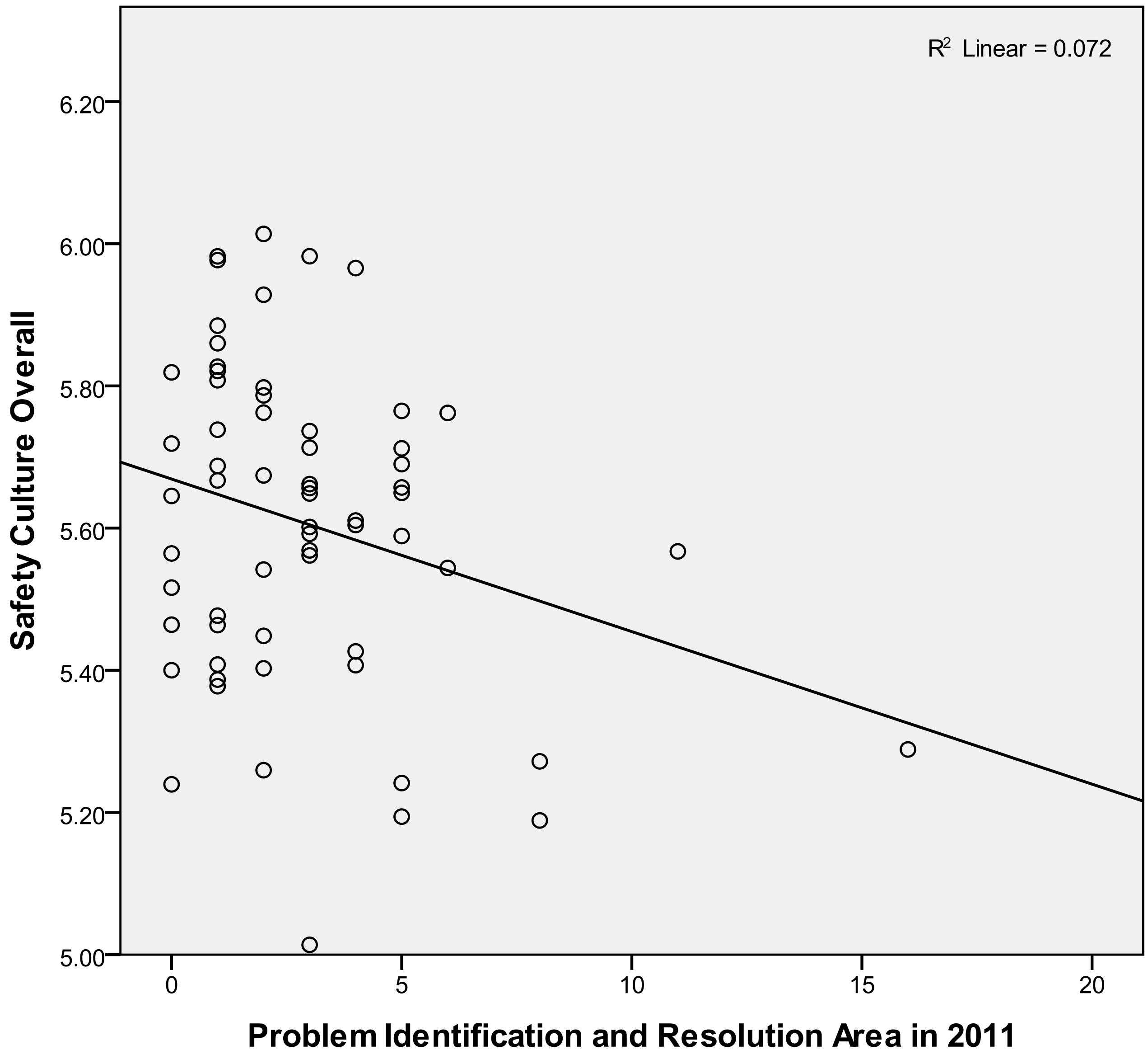
Table 9 Correlations between Safety Culture Survey and 2011 ROP Cross-Cutting Areas and Components

	Total ROP Aspects	Human Perform. Area	HP 1: Decision Making	HP 2: Resources	HP 3: Work Control	HP 4: Work Practices	Problem ID & Resolution	PI&R 1: CAP	PI&R 2: Operating Experience	PI&R 3: Assess-ments
SAFETY CULTURE	-.20	-.12	-.10	-.14	.01	-.07	-.27*	-.28*	-.06	.01
1. Management Responsibility/ Commitment to Safety	-.21	-.11	-.12	-.14	.05	-.06	-.30*	-.30*	-.10	-.02
2. Willingness to Raise Concerns	-.10	-.03	.01	-.08	.01	-.01	-.18	-.23	.03	.12
3. Decision-Making	-.21	-.13	-.10	-.12	.00	-.09	-.29*	-.29*	-.12	.03
4. Supervisor Responsibility for Safety	-.17	-.12	-.06	-.12	-.01	-.10	-.20	-.21	-.07	.06
5. Questioning Attitude	-.19	-.15	-.19	-.11	-.07	-.04	-.19	-.23	.04	.01
6. Safety Communication	-.18	-.14	-.10	-.13	-.05	-.08	-.20	-.23	-.01	-.01
7. Personal Responsibility for Safety	-.06	-.05	.03	-.07	.02	-.10	-.06	-.10	.06	.13
8. Prioritizing Safety	-.21	-.14	-.07	-.18	-.04	-.08	-.25*	-.25*	-.07	-.08
9. Training Quality	-.11	-.06	-.01	-.15	.06	-.04	-.15	-.16	-.02	.03

*p < .05; **p < .01

Figure 7 Scatterplot of Safety Culture and Problem Identification and Resolution Area in 2011

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment



Predictive Correlations (2011)

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

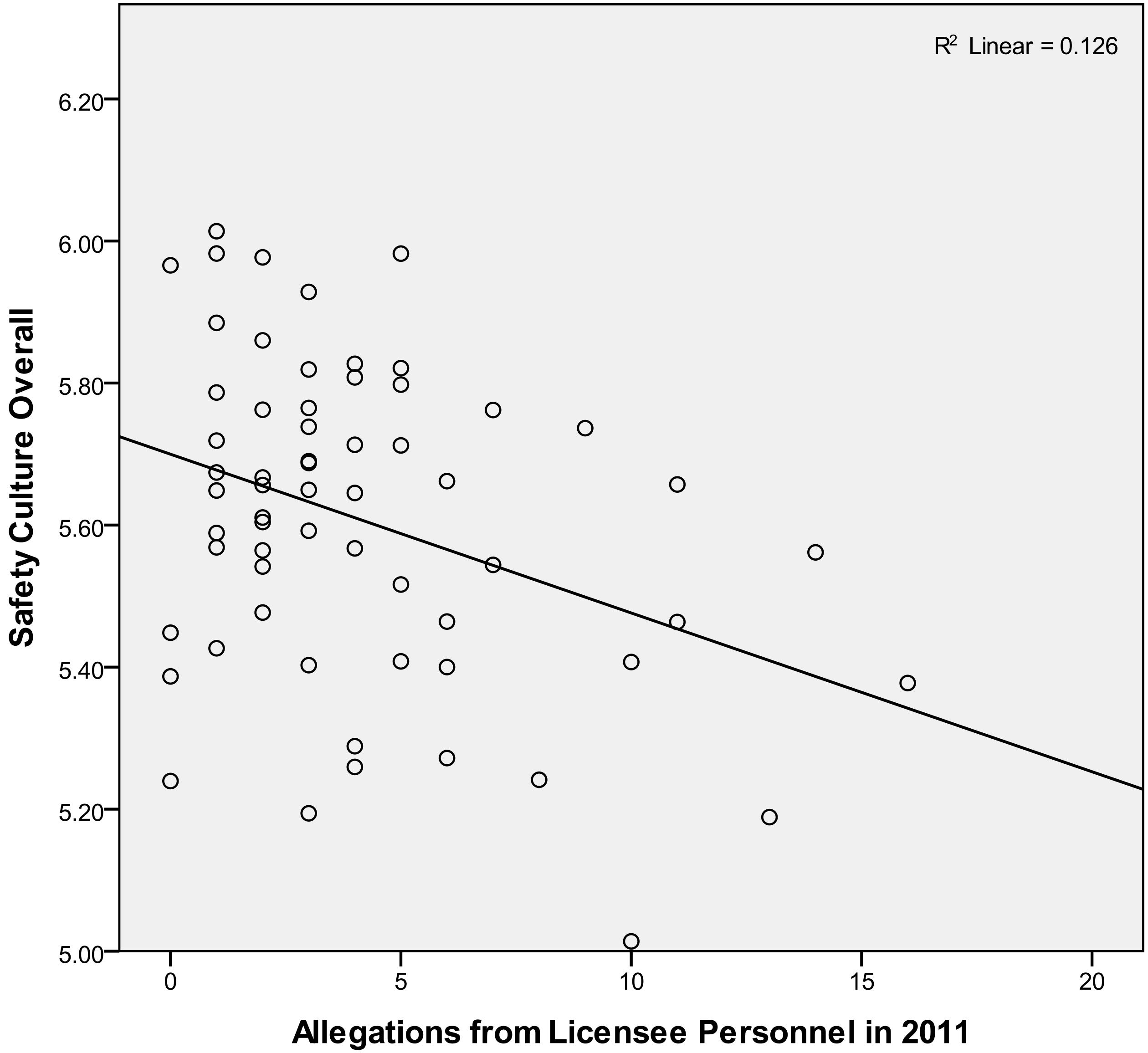
Table 10 Correlations between Safety Culture Survey and 2011 ROP Action Matrix, Inspection Findings, SCCIs, and Allegations

	Action Matrix Elevated	Total Inspection Findings	Total SCCIs	Allegations From Personnel
SAFETY CULTURE	-.30*	.07	-.26*	-.36**
1. Management Responsibility/ Commitment to Safety	-.29*	.07	-.27*	-.38**
2. Willingness to Raise Concerns	-.24	.16	-.18	-.38**
3. Decision-Making	-.32*	.01	-.32*	-.33**
4. Supervisor Responsibility for Safety	-.30*	.01	-.18	-.20
5. Questioning Attitude	-.26*	.02	-.23	-.48**
6. Safety Communication	-.29*	.07	-.21	-.28*
7. Personal Responsibility for Safety	-.05	.19	-.10	-.08
8. Prioritizing Safety	-.27*	.04	-.20	-.21
9. Training Quality	-.23	.02	-.20	-.07

*p < .05; **p < .01

Figure 8 Scatterplot of Safety Culture Survey and Allegations in 2011

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment



Comparison of Selected 2010 and 2011 Correlations

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

	Action Matrix Elevated		Total SCCIs		Allegations From Personnel	
	2010	2011	2010	2011	2010	2011
SAFETY CULTURE	-.23	-.30*	-.23	-.26*	-.21	-.36**
1. Management Responsibility/ Commitment to Safety	-.20	-.29*	-.28*	-.27*	-.24	-.38**
2. Willingness to Raise Concerns	-.31*	-.24	-.10	-.18	-.16	-.38**
3. Decision-Making	-.27*	-.32*	-.21	-.32*	-.17	-.33**
4. Supervisor Responsibility for Safety	-.27*	-.30*	-.14	-.18	-.15	-.20
5. Questioning Attitude	-.17	-.26*	-.38**	-.23	-.41**	-.48**
6. Safety Communication	-.19	-.29*	-.14	-.21	-.12	-.28*
7. Personal Responsibility for Safety	-.28*	-.05	.13	-.10	.17	-.08
8. Prioritizing Safety	-.11	-.27*	-.11	-.20	-.09	-.21
9. Training Quality	-.23	-.23	-.10	-.20	.04	-.07

*p < .05; **p < .01

Key Observations: Predictive Correlations

- Most correlations between safety culture survey and performance metrics from inspection reports (e.g., inspection findings, ROP aspects) were small and non-significant
  - Exception: Management Responsibility was moderately correlated with the Problem Identification & Resolution cross-cutting area (-.30)
- Overall Safety Culture survey was moderated correlated with:
  - Elevated Oversight in the Action Matrix (-.30)
  - Allegations from Licensee Personnel (-.36)
- Strongest predictive correlations between safety culture factors and Allegations metrics:
  - Questioning Attitude correlated with Allegations (-.48)
  - Management Responsibility and Willingness to Raise Concerns factors correlated with Allegations (-.38)
- 2011 Performance Indicator data unavailable at time of analysis

Summary

- **Concurrent validity**: Moderate, statistically significant correlations between safety culture survey results and some safety performance metrics during the same time period.
  - E.g., Unplanned Scrams, Total Inspection Findings, Total ROP Aspects.
- **Predictive validity**: Moderate, statistically significant correlations between safety culture survey results and some broad safety performance metrics measured one year later.
  - E.g., Elevated Oversight in the Action Matrix, Allegations from Licensee Personnel

Questions & Discussion