TECHNICAL GUIDELINES DEVELOPMENT COMMITTEE

MEETING DAY ONE

NATIONAL INSTITUTE OF STANDARDS & TECHNOLOGY

THURSDAY, MARCH 22, 2007

(**START OF AUDIOTAPE 1, SIDE A**)

MODERATOR: I just want to take a check here first to see if the TGDC members that are on line can hear us and we can hear them. Could you still identify if you're on the teleconference? This is Alex. Someone was just on. Good morning.

MR. PIERCE: Yes, good morning. It's Phillip Pierce.

MODERATOR: Hi, Phillip. Good morning. Anyone else on joining us?

MS. MILLER: Good morning. This is Alice Miller.

MODERATOR: Thank you, Alice. Anyone else?

(No audible response.)

MODERATOR: I'm assuming you both can hear us, and we'll identify people as they join us. Good morning, everybody. I'm Alan Eustis for the NIST Information Technology Laboratory here at NIST. Welcome to the Eighth Plannery Session.

Just some quick overviews that we like to do at the beginning of our meetings. We are now located in the Employee's Lounge which I will show you is here. We also have an overflow room in Lecture C, should we have a large attendance, which we are expecting. If we have a fire or a fire drill or emergency, you'll hear the sounds and you'll be warned, and here's the exit. You go out the door here and take a right and go down, keep going down the hallway, and you'll see the glass doors. And you can just exit right. That's the easiest from here. From Lecture C you'll want to go down the hall to the right, and then down and out the entrance, the main entrance is the fastest way there.

Please turn off your cell phones and pagers. There's a lot of RF in this room as it is, so it would be helpful if you'd turn those off, because it actually even affects the video and audio of the microphones. It's best not to have food, but I'm already violating that with my cup of coffee. So I probably can't tell you not to do that.

Particularly the public in attendance, please wear your name badge while you're here. If you're planning

on coming back on Day 2 and you're driving, if you bring your name badge and a license or a positive identification, you do not need to go back through the security shelter. You can just come right to the meeting tomorrow, and we're meeting here tomorrow starting at 8:30, to let everybody know. We'll be breaking for lunch at around 12:30 and the cafeteria is right across the way here. And we'll have some breaks in between.

So with that, welcome to Gaithersburg on a nice March day. And Dr. Jeffrey, the meeting is yours.

MR. CHAIRMAN: Thank you very much. And since I'm violating Rule Number 2, I guess I'll waive it for everyone else since this is not the auditorium. So first of all, good morning and everyone welcome. Dr. William Jeffrey, Director of NIST and Chair of the TGDC. And I hereby call the Eighth Plannery Session of the TGDC to order.

First I think we should stand for the Pledge of Allegiance.

(Allegiance recited by all.)

At this time I'd like to recognize a brand new

parliamentarian for the TGDC, Ms. Thelma Allen who will now do the official roll call.

MS. ALLEN: Okay. Williams? Williams? Williams not responding. Berger? Berger? Berger not responding. Wagner? Wagner is here. Paul Miller? Paul Miller? Paul Miller is not responding. Gayle? Gayle? Gayle is not responding. Mason? Mason is here. Gannon? Gannon is here. Pierce?

MR. PIERCE: Here by teleconference.

MS. ALLEN: Pierce is here. Alice Miller?

MS. MILLER: Here.

MS. ALLEN: Alice Miller is here. Purcell? Purcell is here. Quisenberry? Quisenberry is here. Rivest? Rivest is here. Schutzer? Schutzer’s here. Turner-Bowie? Turner - Bowie is here. Dr. Jeffrey is here. We have 11 in attendance. That is a quorum.

MR. CHAIRMAN: Thank you very much.

UNIDENTIFIED SPEAKER: Mr. Chair, if I could for a second, I forgot we have our signers over to my right. If anyone needs their services they will be here this morning and this afternoon. And please find a seat over here on the right side. Thank you.

MR. CHAIRMAN: Thank you. I'll mention that I know a couple of the TGDC members are stuck in various airports due to bad weather in the Midwest, and they hopefully will be here within a few hours. Well, I'd again like to welcome everyone back to the TGDC and to the Gaithersburg campus of NIST. I know everyone has been working very diligently and very hard over the past couple of months since the December meeting. At least, that's what my staff has claimed. So we've got a lot of work ahead of us before we get the next generation of the TGDC guidelines to the EAC on schedule in July of 2007. And we've really benefited from the advice and counsel that has been provided by this body, so I really do look forward to the next two days of continuing that, as we really try to wrap up and get sort of a little bit of a finishing product going.

I'm especially pleased to have representatives from the EAC here this morning, Commissioner Donetta Davidson, Commissioner -- is Gretchen Hillman here? Okay They should be here a little bit later. And two newly-confirmed members --

UNIDENTIFIED SPEAKER: They'll be here shortly.

MR. CHAIRMAN: They'll be here shortly. So we have two newly-confirmed commissioners, Carolyn Hunter and Rosemary Rodriguez who will join us shortly. And I'd also like to welcome back Executive Director Tom Wilke and his senior staff, who again have been an absolutely invaluable help to us.

So at this time, I'd like to entertain a motion to adopt the minutes from the December 4th and 5th TGDC meeting. Is there a second? There is a second. So, is there a motion for unanimous consent? I'll call for unanimous consent. Any objections? Okay. Hearing no objections to unanimous consent, they are accepted.

Also we have to entertain a motion -- we missed the minutes from the last meeting, I guess -- is that why -- of the March 29th meeting of the TGDC Committee.

UNIDENTIFIED SPEAKER: (Indiscernible.)

MR. CHAIRMAN: I'm sorry. So much for my notes. So the agenda for this meeting, we need to adopt. Sorry. My apologies. Any second to adopt the agenda that you all have in front of you? Okay. There are seconds. Any objections to unanimous consent? With this formalism we now actually have an agenda. Thank

you. Sorry about that confusion.

Since the last meeting in December of 2006, the three working subcommittees of the TGDC have drafted and edited sections of the next generation of the VVSG. And they will be reporting back at this meeting. Specifically as a committee we will review, approve and, where appropriate, provide supplemental direction to the subcommittees. This guidance is critical to the refinement of the final VVSG guidelines that we will hand to the EAC.

At the December 2006 session, TGDC members highlighted the need for the subcommittees to collaborate on issues of mutual concern to two or more of the subcommittees. And we're going to discuss the results of those collaborations tomorrow.

Now, the time required for us to actually go through this agenda means that the committee cannot take public comments at this meeting, however there are opportunities and will be continued opportunities for the public to comment. In fact, I'd like to emphasize that point, that the documents, draft documents are on the web and are available for users, vendors, the

public. If you have any initial feedback, please e-mail us. We'd be happy to accept that.

In addition, I'd like to mention two additional things that we're going to be doing. As we get closer to handing over our draft guidelines to the EAC in July, we want to make sure that those guidelines are as good as possible, that we've captured as many of the needs of the community and that they are attestable and are drafted as well as could be done. So what I'm asking is that for each of the three subcommittees, I've asked that there be co-chairs to each of the three subcommittees where the co-chairs will actually be representatives of the end users, the people out in the states and localities who actually have to make sure that this is implemented.

And I'm very happy to say that Paul Miller has volunteered to be co chair on the Core Requirements Team. Alice Miller has agreed to be co-chair of the Human Factors and Privacy Team. And to prove that you don't have to have the last name Miller to be a co-chair, Helen Purcell has volunteered to be co-chair on the Security and Transparency Subcommittee. So thank

you for your work and again, I want to make sure that we don't have any gaps in what we put forward.

Along those lines, an important component is that when we have the guidelines that they need to be testable and verifiable as they get implemented. And Whitney has had some discussions with us. I think they've raised some really valuable points. And so what I've asked, since NAVLAP works under NIST, I've asked the NAVLAP representative to start participating actively in the subcommittees to ensure that as the guidelines are drafted, they're drafted in such a way that they translate easily into something that can be tested and verified. So I want to make sure that we don't have some inconsistencies across that boundary. And so the NAVLAP folks have agreed to that, and I appreciate that.

Any additional comments and position statements about the work of this committee or the TGDC draft guidelines can be sent to voting@NIST.gov. That's voting at N-I-S-T, dot, G-O-V, and they will be posted on the website. And comments that we've received to date are already posted there and been reviewed by

members.

So at this time, it's my great pleasure -- and let me just mention that Commissioner Gretchen Hillman has joined us. Welcome. In fact, perfect timing because this is the opportunity now. I'd like to invite both commissioners, Donetta Davidson and Gretchen Hillman, to address this committee.

COMMISSIONER DAVIDSON: Good morning. Well, I want to start by thanking each and every one of you sitting here today and on the telephone for the next two days and all of the time that you have served, and definitely giving your guidance in this extremely important task. And your opinions are valuable in this process and we really do appreciate it.

Before I start, I think -- well, maybe I'd better go ahead and then I'll introduce them if they get here before I finish. Commissioner Hillman has been introduced, and as you can tell she will speak right after me. She's going to inform you how we're going to start getting our Standards Board and our Advisory Board involved with this process so that we have been definitely educated as far as we can, and because

they're a real important part of this process.

As we begin our meeting here today and begin to discuss the next iteration of the VVSG, I think it is important to note where we are in this process and where we have to go. You've worked hard since the 2005 VVSG in getting to where we're at today. And after the deliberation of the next two days, NIST and the leaders of the TGDC are looking to hold another meeting in mid-May to finalize the details that will be coming to us, which is planned July of '07 to be delivered to the EAC.

The delivery of the TGDC draft version is an extremely important step, but it only marks the beginning of the next part of the process. After reviewing the TGDC draft, the EAC has the responsibility and mandated under HAVA to conduct a deliberate and thorough review of the document.

First, EAC will review and vet the TGDC document itself. Second, HAVA mandates that the EAC publish their draft version in the federal register and receive public comments for a minimum of 90 days. Also HAVA requires that the EAC Board of Advisors and the Standard Board get a minimum of 90 days also to review and give

comments. After the close of the comments the EAC staff must review, catalog, and incorporate the comments submitted by the Boards, by the public, and by all members that have interested in giving those comments.

For the 2005 version, I think you've heard me say it before -- we got over 6,500 comments that had to be vetted. And we worked very hard along with NIST to make sure that we had the very best product we could have at the time we adopted the 2005 in December. This was only a partial rewrite. This time the VVSG is a complete rewrite. So amongst the steps of HAVA requiring EAC, it also has us holding public hearings to meet with our stakeholders, our major stakeholders.

For instance, we need to know what its election officials need from the machines. We need to know thoughts and concerns from advocacy groups. We must engage the voting system manufacturers to understand the technology available and a timeline for development. This includes an open and honest discussion about how much it's going to cost to develop, manufacture, and test. In order for these guidelines to be functional, they must be affordable.

The point of all of this is that the next iteration of the VVSG is going to take time. And to do it properly, it should take time. Unlike the 2005 -- as I said -- this iteration is a complete rewrite. Anything short of a methodical, systematic, and thorough review by the EAC is irresponsible. With the support of NIST, it is our goal to create a set of standards that won't need to be looked at again for four years. EAC's goal is to end the cycle of constant change between elections. The creation of a comprehensive set of guidelines is the only way to accomplish that goal.

VVSG is only one element in the process though that we have to consider at the EAC. Elections is more than just a machine. We have been working hard in conjunctions with the VVSG to make sure that our election management guidelines aid the election officials in administrating the most transparent, accessible, and trustworthy elections possible. Where the VVSG ends off, the technical guidelines ends, the management guidelines begin in taking the best practices and advice on the administration of elections.

Currently the EAC is working on five new quick-

start guides for the officials. These five guides will cover election certification, developing an audit trail, public relation, disaster planning, and change management. The goal is to release all of these prior to the 2008 election. So the fall, we'd like to have everything out by September of this year.

In conjunction with the management guidelines, the quick starts, the EAC is working to develop several new chapters in the management guidelines. These new chapters will cover everything from military overseas voting to polling place management, and also mail and absentee ballots. The election center -- and we've offered the same to other election training areas -- is going to hold a meeting in Kansas City in April And at that meeting they are introducing a lot of our management guidelines.

In conjunction with all of this is obviously, what is our top priority for 2007. It is to increase public confidence in election. To achieve that goal we must increase voter confidence in voting equipment and the process. That means a vigorous system of testing and certification of the equipment, educating the public and

the voters about the process, and continuing to examine the way we conduct elections, and making improvements as we go. So we have a huge job ahead of us, but we're confident that we can meet that goal with all the help that we have in our group that we are working with.

I want to tell you how much I look forward to the next two days and learning everything that's coming forth. And let's see if we have the new commissioners here yet. No, but when we do we'll make sure that they're introduced.

Thank you very much. I appreciate it. Ms. Hillman?

MR. CHAIRMAN:: Thank you very much.

MS. HILLMAN: Good morning everybody, and thank you for the opportunity to be with you this morning. As Commissioner Davidson said, I mean, we are appreciating the enormity of the task before us over the next several months to perhaps a year to get through the next iteration of the voluntary voting system guidelines. And perhaps the two biggest challenges we have are helping our Standards Board, our 110-member Standards Board and our Board of Advisors to prepare for the role

that HAVA mandates, they perform, in reviewing and commenting on the guidelines.

But beyond that we've got to undertake the task of helping the public digest what we are doing, and doing that in a way that the public can understand. And I think a predecessor to that is to make sure that even among the groups that are intimately involved and familiar with the VVSG, that the scientists and the technical experts and the election officials can communicate and speak the same language. And quite frankly, we're not so sure that's happening right now.

In response to that, the Standards Board is starting now to prepare for its work, and we will be joined today by one member of the Standards Board who is serving on what is being called the VVSG Ad-Hoc Committee of the Standards Board. And basically that committee right now consists of three members of the Standards Board, but in a very short period of time it will grow to a larger-sized committee. This subcommittee, this ad-hoc committee will work with the Executive Board of the Standards Board to really review what its task is, how to help the Standards Board

members receive the information in small enough bites that they can adequately chew and digest it before getting to the full main course after the recommended guidelines are ready for public comment. And I expect that the Board of Advisors will be doing the same thing. And the Board of Advisors Subcommittee and the Standards Board Ad-Hoc Committee will be working together over the next several months to accomplish this.

And in many ways they are important spokespeople about this subject in the states to the more than 7,000 election officials, state and local election officials there are in this country, as well as to the grass-roots community. I mean, as we know the public has never been more interested in the very specifics of how voting systems work than they are today. Irrespective of whether the issue is accessibility for persons with disabilities, security, you know, functioning, human interaction, whatever the situation might be. And so we certainly want to make sure that those two important resources are adequately prepared to have discussions in their communities with their county officials, state officials, governors, whoever it may be, to help

everybody appreciate the implications of the Voluntary Voting System Guidelines on the future of voting and democracy in America.

And thank you so much. I look forward to the conversations as well.

UNIDENTIFIED SPEAKER: Do we have Bill Campbell? No? All right. Oh, Bill Campbell is here. Please let me introduce Bill Campbell who is the City Clerk from the city of Woburn, Massachusetts. And he is a member of the Standards Board, has been since the Standards Board was organized in 2004, and has just completed a tenure on the Executive Board. He's here for the two days to observe, and will be an important reporting mechanism back to the Executive Board. Thank you.

MR. CAMPBELL: Thank you very much. I definitely appreciate the absolutely strong support and good working relationship that this committee has had with the Election Assistance Commission. And we really appreciate the comments from the commissioners.

At this point I'll ask Mr. Mark Skall to review the summary of activities since December 2006. And I believe that all the information in his briefing is

contained in the three-ring binder marked workbook. And I will also reiterate one of the comments that Commissioner Hillman made, is during these presentations, the closer to English we can get some of the technical briefings the better we will all be served. So with that challenge, Mark --

MR. SKALL: You know, for someone from Brooklyn, New York, it's very difficult to meet that challenge. Good morning. I'd like to tell you about the voting activities that NIST has been engaged with over the last few months.

So this is an overview of what I'm going to speak about. First of all, since December 4th and 5th, which was the last TGDC meeting, we've been very, very busy. As you know, the TGDC itself makes recommendations to the EAC with respect to Voluntary Voting System Guidelines. NIST of course provides the research and actually drafts the words that go into the VVSG. We of course cannot do this without very close coordination with the TGDC.

Outreach is a very important area. In doing our research, we want to make absolutely sure that we meet

with everybody we can meet so we understand the environment we're working in, so that our research is as thorough as it can be. And those of you who have been involved with the TGDC from the beginning know that during the first iteration of the VVSG, that was very difficult because we were time constrained. During this iteration we really are trying as well as we can to reach out to as many different people so we can learn everything we can in order to do our research.

Lastly, I'll talk about the resolution matrix that the TGDC asked us to keep up to date. The agenda and aims of the meeting are to -- I'm going to talk a little bit about the focus of this meeting, the strategy of this meeting, and then go over the agenda.

The first bullet is to remind me of an issue that I did want to mention. In speaking with the EAC, they were concerned that we were referring to this upcoming iteration as VVSG 2007. For many good reasons, this iteration, by the time it goes through the public reviews will almost definitely not be adopted by the EAC until at least 2008, I guess 2008 at least. I hope it doesn't go another year. So we've been asked to refer

to this as something generic, like the next iteration of the VVSG or the new VVSG. Perhaps we can have a contest to name this, but not 2007.

I'd like to quickly go over the research that the sub-groups have been doing over the last few months. HFP has been working very arduously to update the usability performance, benchmarks, which are of course very, very important to get benchmarks that are performance based rather than constraining design, updates to usability requirements, and updates to accessibility requirements. They've also been looking at software independence and accessibility, the relationship between the two.

The CRT has continued to do research in reliability in benchmarks, quality requirements, electromagnetic compatibility requirements. STS of course has been doing a lot of the work on software independence and auditing research, innovation class research, coordination with HFP on software independence and accessibility, and paper record usability issues, and then more traditional security requirements such as updates to Crypto, set up validation access control, and

system event logging.

Now again, we work very, very closely with the TGDC. Obviously we've had 21 telecons since the last December meeting, joint telecons between the committees which we think is a very good idea. We obviously don't want the committees working in a vacuum so the joint telecons increase that synergy. We prepared much discussion papers, draft material, and of course numerous individual discussions among ourselves and with TGDC members.

We have what we call the draft build. The is essentially the draft of the VVSG. It's on the web. Every time we do our research we -- and we do drafting, we fill in the sections of the VVSG. We have over 500 pages now. We believe that drafting -- and this is a very rough number, so don't hold me to it -- the drafting can be about 80% complete as far as actually putting pen to paper, but that last 20% is going to be very, very challenging. And we're continuing to work with a newer and more usable format for the VVSG.

Again, outreach and support of NIST research, we of course have very close coordination with the EAC

including monthly meetings and countless telephone calls. We clearly are conscious that the vendors play a very important part in this, and they have to implement the VVSG. So we have reached out to them and we do have regular meetings with them via the Information Technology Association of America. There's actually a sub-group there that's devoted to voting system vendors.

We've made numerous presentations and discussions at conferences and meetings such as the Election Center Advisory Board. The Standards Board is an incredibly invaluable resource for us I believe. I was just there a few weeks ago speaking, and meeting with the very important election officials, the Secretaries of States, and others has been just invaluable for us to get information about how to process work. And we've had some more formal coordinations with the Standards Board that ran afoul of (indiscernible) rules, but we are clearly informally trying to liaise with them as much as possible. And we have a very good relationship, we believe, with them.

Other meetings as well, outreach to election officials on paper auditing issues starting with the

election officials on the TGDC to understand that issue. And we sent some correspondence to NAS and NASET to get more information on benchmark for reliability.

There has been a resolution for NIST to create a matrix and update it to map back our research and our drafting to the actual resolutions. And we're very conscious of that and update that regularly, and the website is listed under the third bullet for that.

So the aims of this meeting: after this meeting we're going to propose one more meeting, probably in the middle to end of May. So there is one meeting left between now and the delivery to the EAC in July. It's very important to us at NIST and to the TGDC to reach as much closure as we possibly can at this meeting. It will be very difficult to change the direction in May at the next meeting, so we would like to get all things as much as possible resolved now so that NIST can have clear direction to develop the VVSG drafts, which of course will be the TGDC product.

So our goal at the next meeting in May is essentially to have a complete draft. So we would like, as I said, closure. And I'm going to ask the NIST staff

as well when they're up there to make sure they have all questions, all issues that they don't have clear direction answered, and to please speak up if in fact you need further guidance. So that's the goal of the meeting that I hope you all agree with. The aims of the meeting, again to make substantial progress of finalizing existing material and to discuss remaining open issues, and of course to get a consensus.

So the presentations are broken up into two days. Day 1 will be subcommittee presentations, subcommittee consensus issues and material. And we hope to limit the discussion to that material and save some of the cross-cutting issues and perhaps more volatile issues to the second day so we can achieve consensus with the material in hand. The second day will be cross-cutting issues. The discussion will probably be a lot more open ended, because not as many decisions certainly have been made in those issues. So we need further guidance.

Today's presentations will begin with an overview of the draft VVSG by John Whack. The Security and Transparency Committee will talk about the many things it's doing: audit architecture, electronic paper record

requirements, crypto requirements, access control, software distribution and set-up, core requirements and testing, QA and configuration management, EMC requirements, review of CRT changes from the previous draft, and benchmarks for reliability. And then human factors and privacy, we'll be discussing usability requirements, accessibility requirements, privacy requirements, and the usability research update.

Tomorrow we'll begin with Mary Saunders giving the presentation on NAVLAP activities as Dr. Jeffrey mentioned. Although NAVLAP is clearly not within the scope of the TGDC, NAVLAP does activities that relate to the work we're doing since they have to assess testing labs for competence in testing the VVSG. Certainly we want a dialogue between them and all of us to make sure that the requirements we put in there are in fact testable and acceptable to be assessed.

The cross-cutting discussions will begin with the innovation class, accessibility and software independence, paper rolls, VVSG scope and ballot activation, and then there's time for resolutions and future TGDC meeting planning.

That's about it. Any questions? Thank you.

MR. CHAIRMAN:: Now let's call John Whack up for the next presentation which is the draft VVSG recommendations, the EAC overview. Hopefully I've got the right order.

MR. WHACK: Thank you. First though, I'd like to ask Commissioner Davidson to come back up and do some more introductions.

MS. DAVIDSON: Our commissioners have arrived, so I'd like to take a moment to introduce them. Rosemary Rodriguez is filling the vacancy of Ray Martinez, and she's with us. And then Caroline Hunter -- I'll get on this side where I really can see. Caroline Hunter has been selected to replace Paul DeDivorio (phonetic spelling). So welcome and thank you both for being here. They plan on being here most of the day, so please everybody introduce yourselves to them so that they can get to know you. And they'll be very interested in meeting everybody. Thank you.

MR. WHACK: Thank you, and welcome to the new commissioners. We look forward to a lot of work together over the next several months.

Okay. Good morning. It's always a pleasure to be here and talk with you. And what I'm going to do is -- actually we have little bit more time. We're a little bit ahead of schedule. We don't actually take a break until 10:30 so I can speak very slowly, which is pretty easy for me to do. What I'll do is I'm going to go over essentially kind of where we are in the schedule, what we have remaining for the next couple of months, what we expect to be doing after the TGDC delivers this to the EAC. Then I'm going to go through the document itself and just simply try to point things out a little bit. The document is getting very big. I don't expect that all of you have read the thing from cover to cover at this point. Mark Twain was a book critic and he was reviewing somebody else's book, and he said once you put it down you can't pick it back up. And it's sort of like the VVSG right now. So with that I will launch into this. At the end I'm going to talk about response to TGDC Resolution 2305. So I'll get into that.

Okay. I'll just start with what's going to happen right after this meeting. We will make changes undoubtedly, and we still have some general areas we

still have to complete, some remaining core material on security and CRT. We need help on the innovation class requirements and open-ended vulnerability, and some other areas that we'll talk about in more detail tomorrow, final updates to the usability that Dr. Leskowski is going to get into later today.

Then we have to go through a process of essentially harmonizing a lot of material. We have some overlap right now and we just have a very large document, and we're highly interested in the document being as usable and readable to the community as possible. We want to give to the EAC a document that doesn't saddle them with a lot of reformatting or a lot of restructuring. We've got a lot of guide material to write, things like that. So we are hoping to more or less be done with the document by the end of May. And then we could spend a leisurely -- you know, that's a joke -- leisurely June and July going through and, you know, maybe boiling it down to fewer requirements. Right now we have roughly about 920, 930 requirements right now, and while we have more to add, the final number could actually go down a fair amount. There may be better ways to present the

requirements than we've done.

We will -- well, I guess NIST and the TGDC upon delivery to the EAC in July will post the draft recommendations on our website, and as before the EAC will review this. They may make adjustments. They will put a version out for public review. We expect that we will be involved in vetting this with the Standards Board and other groups as requested. I think some TGDC members -- I remember Ron and Whitney and some others were helping last, two summers ago, right? Two summers ago out in Colorado doing the same thing. Final version likely in 2008. Maybe this final bullet might get discussed a little bit, I don't know. There's pending congressional legislation that may affect things.

Okay. With that let me -- let's see if I can do this with a mouse. Okay. This is the document right here. And what I'm going to do -- and I will try to speak fast actually, and I'll go through it rather quickly. Please feel free to raise your hand and stop me if you have a question. I promise as much as possible to the NIST people that I will try not to make up like new projects or things we need to do as I'm

talking. I do want to say that this document right now, the structure of it kind of reflects the communication that we have within our project right now. And that has to be ironed out somewhat, the structure of the document, that is. We don't always agree on the material, but I do have to say that everybody here on a project at NIST really cares about this material. It's the first time I've ever been involved in a project where I've seen such dedication from people. People actually really care about the material. And we want to do as good a job as we can. We need clear direction from a number of -- well, on a number of different items today and tomorrow. And we're looking to you to get as much help as we can.

Okay. So I'll start with just a quick overview of the document, and hopefully you can see it there and you don't get motion sickness if I go through it fairly quickly. Essentially it says six volumes. Really volume 6 is just a bibliography and a summary of requirements. But essentially the guidelines overview, you'll notice Frank Lloyd Bright's great-grandson designed the cover for us. I don't know if we'll stick

with that, but essentially this is going to be just an introduction to the other four volumes, volumes 2 through 5. We haven't written much material for that yet, but that will be essentially a guide to the standards themselves.

Terminology section standard is essentially the glossary. And I'll go through it briefly. Maybe I can blow that up a little bit. Scope, applicability, essentially the big changes here maybe from the 2005 version is we have stuck to only the terms that we're using on the VVSG, in the draft VVSG. The final version of this, these will all be linked. There will be a lot of cross-referencing. These definitions build upon themselves. So we've got that. Volumes 3 and 4 are actually the volumes that have most of the requirements in them. Volume 5 has some requirements as well, but Volume 3 is really the requirements that apply to the equipment basically, requirements for vendors. So I'll go through the introduction a little bit.

Again, let me blow that up a little bit. Standard things up front, starting off with a description and rationale of significant changes from Reference 6, which

is the 2005 VVSG. And I want to just point out a few things. Maybe some people in STS may not realize some other stuff is in this particular area in the core requirements. But we'll go through the conformance clause a little bit. Discussion on marginal marks, those in CRT remember a fair amount of discussion there. Actually, let me expand that a little bit. Coding conventions, a lot of updates to coding conventions, structured programming, a number of things in there. Discussion of COTS, how COTS is being handled. And I think if I'm not wrong, Volume 5 has more discussion on COTS.

Reference models, right now we have a couple different reference models at the very end of the chapter. The section on deletions, what's not going to be standardized, a number of different things. So this is I think a good introduction written mainly by David Flater. It will be augmented a good bit with more material from SES and HFP before we're done.

Let me move down here. The conformance clause -- and it's not a single clause, it's actually a chapter. There's a lot of clauses in it. Basically going over

the structure of requirements, what's normative, what's informative, implementation statement. I'll talk a little bit about the structure of requirements in a minute. The terminology we're using for classes. Now, classes as you know really are in essence things used in the requirements to distinguish what the requirements apply to, what sorts of equipment a requirements applies to. Classes are arranged hierarchically. There's some pictures right up here. This is the voting device class here. And for example, a VVPAT is related to DRE; DREs can be related to tabulator, and basically starting with voting device at the very top.

And then when we get to the actual requirements, I will show you a little bit more about that. Semantics of classes, how they are joined together, various extensions that can be added on. And them I'll get into the various chapters of the volume. Right now we've got -- is it eight different security chapters? Let me go down to a smaller number. We haven't figured out the arrangement in general, but let's say there's security represented in front of you. There are core requirements, HPF usability/accessibility. There are

more requirements written by CRT, requirements by voting activity, and some reference models.

And I won't go through these in much detail at all. It's my intention just to show you where they are right now. Access control, for example, I believe we'll be discussing some of that. But why don't I just stop here and show you this requirement to give you an idea of the structure. Basically every requirement has this arrow here. Every requirement now has a title. If you're skimming through, the titles hopefully -- well, I think by and large they are fairly descriptive of the requirements, so it's a short hand for being able to skim through quickly and find what you're looking for. Green text, you know, just sort of make the requirement's body stand out more to people. Here's a class. This applies to the voting system class, which means it applies to all voting systems at this point. Test reference, this points to Volume 5, Section 5-2. Why don't we go there quickly. Volume 5, Section 5-2. Okay. Chapter 5, functional testing. So it's basically saying that requirement will be tested via techniques and functional testing. How do I get back to where I

was?

Okay. So test reference applies to a discussion field. Many of the requirements have it. Where did the requirement originate from, many of these are new requirements. Many of these have their origins in VVSG 2005 or the VSS 2002, or other areas of the standards. The impact field probably isn't a field that will appear in the final version of the standard. This is more a note to us, but it's just in general describing what sort of impact this requirement may have on equipment or new technology.

Now, let me find a sub-requirement here. Not a whole lot of sub-requirements in this section here. I think there's a couple down here about passwords if I can find them. This is the part where I was saying motion sickness. Oh, here we are. Okay. User name and password management requirement. Okay, so we've got a general, a more high-level requirement here. We have two choices. We could then have made this requirement very long and maybe put a table, or we could have put a number of sub-requirements. So we chose the sub-requirement route here that get into more detail. So a

sub-requirement has this symbol here. It's one level of sub-requirements basically There aren't sub-sub-requirements, just sub-requirements.

Okay. So that's the requirements structure. I think I'll skip ahead to the CRT general requirements. Don't look at those yellow things there. Well, you can look at them.

(**END OF AUDIOTAPE 1, SIDE A**)

* * * * *

(**START OF AUDIOTAPE 1, SIDE B**)

MR. WHACK: -- good idea, and it's basically to identify those glossary terms that are used and provide a link to them, you know, back to the glossary. I think tabulator device that counts votes -- and it's a way of making sure that these are understood correctly. So a couple of other things. So there will be a lot of cross referencing and linking here. And the idea gain is to make this as usable as possible. It could be that developers, testers end up using a paper version of this, so the hyperlinks at least will identify that it is a glossary term.

CRT general requirements, what are they? You can

think of them as the basic core functional requirements for voting systems. Another way of looking at them is that they're everything else after security and HFP. And some requirements in here may actually end up leaving and getting covered by some of the other subcommittees. But, you know, just going through some of the requirements here for voting variations, you can see what we're covering here, cross-party endorsements, so on and so forth.

What else are we going through? Hardware, software performance general requirements, reliability, accuracy. We'll be talking about these, and Goldfine will be talking about electrical workmanship. Software engineering practices for those people in STS, there is a lot of material in here on structured programming, various coding practices, techniques, things that need to be used, a lot of material in here that I would recommend taking a look at. Quality assurance, quality -- Alan Goldfine again will be discussing some of that. Durability, security, and audit architecture -- John Kelsey will be discussing some of that. When I talk about overlaps, this is one example. It could be that

the audit material may go here. It could all end up in the security section, just as long as it's in one place so you can find it.

Archival requirements, so on and so forth -- interoperability, right now we have interoperability requirements in this section. They may migrate over to STS. These essentially deal with a standard format for data.

Okay. Usability, I won't go into too much because Sharon Leskowski will do a good job of that. Essentially, just in case you don't know already, people think of this as usability and accessibility, but it also has the privacy requirements and also has new material that I think we discussed a little bit last time on usability for poll workers. So again, important material to pick out.

Requirements by voting activity, another arrangement of requirements basically necessary to support different activities. So basically election preparation, equipment setup, opening polls, casting, closing, counting, reporting, you know, various requirements on what reports ought to look like. Audit

status and readiness reports may be covered here. It might be covered in much more detail perhaps in STS. How the reports are formatted may be covered more in HFP. So we've got the basic requirement that there shall be reports here. Things about the security, whether they be digitally signed would be in STS. Ways of representing the data so that it is readily usable and accurately read would probably be more in HFP.

So with that we have reference models at the end of Volume 3 which talks about the process model being used in here with various diagrams. And I think we have the UML down below. And this part you will not be able to put down. I'm sorry. I shouldn't say that. It's tough to read this without a computer interpreting it for you. Various vote capture device state models, things of that -- and any work we do and, you know, discussion of threats would probably go in this general area as well.

Okay. Volume 4 is the other big volume with requirements in it, but these are requirements for vendors and test labs, documentation requirements. And scrolling through this again -- let me blow that up -- again the introductions are well written and it's good

just to read through quickly and just take a look in general. And it does a good job of telling you what's in the general volume as well.

Requirements for what the vendor has to deliver to the test lab, the technical data package, this again is an important area for security. It's an important area for all subcommittees. It's cross cutting, but I think security and CRT have a lot of involvement here. Voting equipment, user documentation, this has been a big topic of discussion as well in all three subcommittees; how well it's written, how readable it is, the things it covers. So these are things to look at as well per discussions and all the telecons regarding some of this. Certification test plan, test report for the EAC, the public information package, Dave Flater has looked at the EAC's certification plan and material, and done his best to harmonize this as well as he can.

I'm doing pretty well on time here. I'll conclude here with Volume 5, and Volume 5 here is the testing standard. It does not contain the tests themselves to test specific requirements. It basically contains everything but that. It is in essence kind of an

introduction to the different types of tests, and it has information in the test protocol section more about how the tests will be conducted in general. There is an informative section, Chapter 2, on the conformity assessment process, which is an overview of that. So, just to good to read in general.

And then Chapter 3 is introduction to test methods, the different types of test methods that will be used here. Vulnerability testing; another name for that is open-ended vulnerability testing. And that will probably have more material in it. Discussion of interoperability testing -- so on and so forth.

Some requirements for documentation and design reviews in Chapter 4, and discussion a little bit of COTS, physical configuration audits. And actually, one thing I may have passed over if you don't mind me jumping back to the introduction, I'm not sure -- we had some discussion about COTs and the different types of categories we're going to use. So that is in the introduction. I just wanted to point that out since I think I saw some requirements pertaining to that in Chapter 4.

And then Chapter 5, test protocols -- test protocols sometimes these terms are confusing, but a test protocol here is essentially how the test in general is being done, but not the specific test, so how functional testing will be done, various general guidelines, pass/fail criteria, assertions, missing functionality, things in here about what vendors have to report on such as the number of should requirements they meet or they don't meet, you know, things of that sort. And in general that's what we have there.

The bibliography -- well, I've got almost half an hour left. I could just go right through and read them all, but --

MR. CHAIRMAN:: You're not obligated to.

MR. WHACK: -- but I won't do that. But we do plan to have an extensive summary. Now, right now we have just a summary of the requirements table. And we'd encourage feedback for the sorts of tables we could put in there that would make this easier to read, and/or ways we could format this better so people can find things. I guess Acrobat links this already, so of course we can get to the requirements by clicking on the

page numbers. But if there are better ways of presenting summaries to the material and you have advice for us, we'd certainly like to hear it. We're working with the EAC as well on this format. We want to give them something that they can rather immediately use.

That is kind of it. Are there any questions I can answer quickly on this, you know, pertaining to the structure or the document?

MS. QUISENBERRY: Hi, Whitney Quisenberry. I have a comment and a question. The comment is that I think I was one of the people that set you down the path of trying to make the document usable by, I think nagging is an appropriate word. And I'd like to commend you on the results. I know that this was not an easy task and we've gone a long way from, this is technical so it should be hard to read. And I think that the layout and structure of the document is really quite usable and very attractive to read. I mean, you open it up and I feel like I can scan through it quite quickly. And so I'd like to -- I'm sure a lot of people worked very hard on that, so a round of applause for all of them and for you for sticking with it.

And the question is in Volume 3, whether the order and organization of the chapters within that section as presented here is determined, or is that simply mushing it all together and getting it into the document?

MR. WHACK: It was mushing it in together and getting it in the document really. If the TGDC has preferences as to the order, that's fine.

MS. QUISENBERRY: I'd love to make a pitch for, starting with usability and accessibility. Of course it's mine so that's an obvious place to start, but I'd like to give a reason why. And it's that this is a technical standard. It's an equipment standard. And because we're so focused on the details of the equipment, it's easy to lose track of the fact that the purpose of this standard is to support humans and human activity. And so starting with that and then talking about the technical requirements that support that activity I think would help us all remember why we're all here.

MR. WHACK: In the VVSG 2005 there was HFP Chapter 3, so there may be some good reason just to continue that as well. Any other comments that I can get to?

None? Okay. Well, I want to reinforce that this is out there on the public site of the website. Anybody can get to it, anybody can read it, any vendors, anybody else out there in the community is welcome to read it. And we welcome comments to the TGDC. We post those comments, they are available. The slides of course are available as well.

Since I'm ahead of schedule, I anticipated that I would go over the response to Resolution 2305 after the break, but how about if I do that before the break? And I think I'll still be ahead of the schedule at that point. So if that's okay --

This was a resolution that, if my memory serves me right, we started discussing in December of 2004. And coming up, those of you who worked on VVSG 2005 remember that we had a flurry of activity in December and January to develop resolutions. And the idea was basically to put electronic data into some interoperable format, hopefully something along the lines of ASCII that people could read easily. This is what we came up with and this is the resolution. I won't read it out loud to you since you know it pretty well.

So we did some research into this area. We think it's an extremely important area that, for a number of reasons, security is where I worked in mostly, but in core requirements, for all sorts of reasons it's important to have some sort of standard format to represent data. Right now, OASIS EML IEEE Project 1622 -- 1622 hasn't adopted anything yet. Current rev of EML is 4. Dave Flater submitted some issues and a number of those were incorporated into the version that's out right now, EML 5, which may be an OASIS standard by summer 2007.

What we need to do at NIST is make sure that we get all the information regarding our requirements for a format to both organizations, to OASIS and IEEE, to assist in moving this forward, to make sure that we can at some point reference a standard in the VVSG, and that everybody can start using. I think this is, as we've talked a little bit with some people, kind of a chicken-and-egg situation where we can't wait forever for one standard to emerge because it probably needs some pushing. At the same time we need to wait a little bit longer for both areas to mature fully, and for us also

to work more closely with them and make sure we've communicated our requirements there.

So that's what we're doing right now. And in the VVSG we're going to do what we did in VVSG 2005, which is basically have requirements for an interoperable format and what information goes into it. And in discussion fields we referenced EML in 2005. We will do the same again in the draft VVSG recommendations.

With that --

MR. CHAIRMAN:: Let me check. Are there any questions on what John just described in terms of the impact on 2305 then? Pat?

MR. GANNON: This is Patrick Gannon. I wish I could provide some additional update to the status information you provided there. The OASIS Election Voter Services Committee did have some interaction with David Flater, has been requesting closer participation from NIST staff on that committee to move that forward. The election (indiscernible) version 5 has been approved by the committee as a public review draft. It is out for a 60-day public review. Once that public review is completed, it will then be submitted to become an OASIS

standard around the June timeframe. And then shortly after that, the plan is to submit it to become an ISO standard for that. Also, the evaluation, there are representatives from the IEEE P1622 working group on the OASIS Election Voter Services Committee. They have reviewed the requirements under P1622 and find that they are a subset of the capabilities provided in the election market language standard, and that the new version 5 meets all of those requirements under 1622, even though the P1622 doesn't actually have any standard that they have adopted to meet the requirements. So right now the email seems to meet all of those requirements too, and are looking forward to setting up some testing or demonstration, interoperability demonstrations in the middle of this year.

MR. CHAIRMAN:: Thank you. Any other questions or comments for John? Okay. With that, I appreciate him getting us ahead of schedule. I'm sure we're going to lose it later in the day. So with that, let's take a 15-minute break right now and come back -- I'll be realistic -- 10:20 according to the official atomic clock up there. Thanks.

(Break.)

MR. CHAIRMAN:: Okay. Thank you very much. At this time I'd like to ask Nelson Hastings, Bill Burr, and John Kelsey to -- I assume one of the three of you, at least -- to come up and present security and transparency progress.

MR. HASTINGS: I'm going to give the security and transparency progress report. And John Kelsey will do a presentation on auditing, and Bill Burr will do a presentation on cryptography. So the overview is -- I'll review the development process that we're using to create draft requirements. Then we'll go through very briefly, very quickly the status of the different security requirements, different topics, grouped by topics. And then we'll open up for discussion.

So to give you a perspective on where things are as it's presented in the presentations, we first create draft requirements based on the TGDC resolutions and telecons. It's distributed within this for review. We revise those requirements and then distribute that to the Security and Transparency Subcommittee for review, and then we revised the requirements based on those

comments. And then we distribute those revised requirements to the TGDC at large for review.

So these are the ten different topic areas that we're working on currently. The ones at the top are a little less mature than the ones towards the end of the list there. And you'll see that in the presentation.

So since the last TGDC meeting, we've developed some draft requirements on physical security. Those requirements relate to physical keys, tamper-proof seals, external ports, door covers and panels, and encasements. Those requirements are being reviewed at this point by NIST staff to be revised, and will shortly be distributed to STS for review. Also since the last TGDC meeting, system integrity management requirements have been developed, and they cover areas such as communication security, malicious code protection, platform configuration management, and error conditions and how to alert people to those in handling of those.

Those requirements need to be mapped to the previous version of the VVSG to understand the impact, how far are we stretching the requirements in this iteration. In addition, they need to be harmonized with

the security and non-security related requirements. At this point it's in the process of being reviewed and updated internally to NIST and will shortly be distributed to STS for their review and feedback.

The innovation class has come up since the last TGDC meeting as part of a resolution. Some initial research and development has been conducted into creating some high-level requirements and entry criteria into the innovation class. We're working with the EAC to address how the innovation class type system could be certified, how to integrate that into their testing and certification program. The real question is how are innovative techniques going to be reviewed and tested. A discussion paper was recently distributed to STS for review, and I believe tomorrow we're going to have an extensive discussion on that topic.

Security documentation requirements - since the last meeting we've developed a few high level, very high-level requirements. These requirements need to be polished up to map to the previous version of VVSG. The low-level requirements, those have been developed as the different sections or the different areas of security

requirements have been developed. And once those areas become stable, we'll take and pull those out and consolidate them and put them into Volume 4 of the VVSG.

In general, there's three areas of documentation related to security. Some general security documentation related to security architecture and the threats that the systems are to mitigate, some technical documentation related to how the voting equipment is designed and implemented to provide the security features. Those documentations really feed into the testing labs to help them in order to perform testing. User documentation is related to how voting equipment security features are used. In addition to that, it requires kind of the assumed policies and procedures that were envisioned by the vendors when this equipment was created, so that if certain policies and procedures aren't implemented other mitigating policies and procedures would have to be put in place to mitigate those issues. The distribution of this to STS will probably be more in kind of chunks. As the general security requirements become -- as the high-level security requirements become stable, we'll let those out

for STS review. And then as the low-level requirements become available, we'll also let those out.

Software distribution requirements have been developed since the last meeting. They cover issues such as the creation of software distribution package master copies where software distribution packages have digital signatures on each file contained in that software distribution package. Requirements related to the witness build of the software, requirements based on types of repositories and the services that they provide, access control requirements -- this is a kind of cross-over area with access control in relationship to software installation and limiting software installation to the preloading mode. Requirements need to be mapped to the VVSG 2005 and harmonized once again with the security and non-security related requirements. This was very recently distributed to the STS subcommittee for review.

The next set of requirements relate to system event logging requirements. These requirements have also been developed since the TGDC meeting, the last TGDC meeting I should say. And they cover the types of events that

need to be captured by the log entry information, such as date, time, the type of event that occurred, protection of the logs through the use of cryptography, and log management. On this slide I should have put that these requirements have been mapped to VVSG for impact. And basically that's showing us that the types of events that need to be captured are at a much more detailed level than in the previous version. Also the introduction of the use of cryptography into protection of the logs. This was distributed to STS for feedback and was updated based on that. One of the comments that we had was to put the events into a tabular form so that it would be easier to read and understand. So we did that, as well as to cut out requirements that, to simplify the requirements and make them less complex. One of the big questions that came up is how configurable should system event logging capabilities be. In general-purpose operating systems, the configurability of the event logs is built in pretty much into those systems. However, limited use operating systems such as single-process, single-user operating systems or embedded operating systems probably don't

have those capabilities. And so we're working with STS to scope these requirements appropriately. And once that scooping is done, we'll redistribute it to the STS for review.

Access control requirements have been updated since the December meeting. They cover things such as authentication mechanisms, access control and (indiscernible) mechanisms, management of identities and rights and limitations of rights during (indiscernible) modes of the voting system. They have been mapped to the VVSG for impact analysis, and one of the things here is that in the previous version, authentication mechanisms really focused on the use of passwords and those types of things, very detailed requirements related to passwords. So we've tried to open it up and give better requirements in terms of use of hardware tokens or tokens for authentication. Reviewed the impact of software independence on this.

Originally these requirements were developed before the passage of the software independence resolution at the last meeting. And it turns out that software independence really didn't have an impact on those

requirements. We distributed the requirements to STS, we updated it based on their feedback, and once again there was a requirement on how or why should the access control policies be so configurable, how flexible should those be. Once again, general-purpose operating systems have these capabilities available, limited operating systems don't. And once again we're working with STS to scope these requirements properly. Once those are scoped we'll redistributed tests just for their review.

Set-up validation requirements have been updated since the last meeting. They deal with software identification and verification, inspection of register variables, and registers and variables, and other equipment property such as the levels of power that's left in back-up power supplies, being able to determine if the communications capabilities of the system are on or off, does the equipment have the correct level of consumables in it such as paper and ink. They've been mapped to the VVSG 2005. A lot of the 2005 (indiscernible) was very centric on software identification and verification as well as having the variable and register inspection capabilities. So a lot

of the new requirements in this section relate to voting -- the other properties in a sense. We wanted to expand the scope away from just the software and the registers.

Again, these requirements were developed before the passage of the software independence resolution, so we went back and looked at what areas that software independence actually impacted these requirements. And as you saw, there were software identification and software verification requirements. And the software identification requirements, it really didn't have too much of an impact on it. However, on the software verification requirements, it did have some impact. And one of those is that it seems acceptable to allow internal verification of installed software for non-network, vote-capture devices. Non-networks is kind of a misnomer here in the sense that it could be limited -- a limited network is more descriptive of what it is. And what we mean by limited network capability is that a vote-capture device could communicate with one election management-type system, or one other vote-capture device. So very limited communication with other devices.

What does this mean? It means that an external interface to check the installed software is not required on those limited network-type of vote-capture devices. So then the external verification is required for election management systems and network vote-capture devices, fully network vote-capture devices or more completely network vote-capture devices. And the reason here is that election management systems and network vote-capture devices do communicate with several different devices during the process of the election. And in that case, there is more chance of those systems getting infected with viruses and stuff. It seems somewhat appropriate to have Election Management Systems have this, because in most cases those systems are on general-purpose PCs that already have external interfaces on them. So that was the justification for that.

These requirements have been distributed to STS for review. They've been updated based on the feedback received. Some of the feedback that was received is to reduce the complexity and possibly try to raise the level of the requirements higher. So we discussed that

a little bit, and what it did is if you had different types of verification techniques, some that are cryptographic-based and some that are not, you need that level of granularity. So what was discussed is, should the VVSG support means other than cryptography for verification techniques. And it was decided that in this iteration, because the non-cryptographic based techniques are at a very infant stage in their development, that this iteration will explicitly call out cryptographic-based techniques. Those updates will need to be redistributed to STS for review.

Auditing requirements have been developed, focusing on how to achieve software independence through auditing. The requirements developed since the last meeting are the capabilities of the equipment to support auditing, requirements related to electronic records and paper records. And that was recently distributed to STS for feedback. John Kelsey will give you a more detailed presentation on that topic.

And cryptography requirements, the requirements have been significantly updated since the last meeting. It eliminated the tutorial style that it used to have,

the tutorial flavor of that section. It still focuses on using FIPS 140-2, validated cryptographic modules, and it really focuses on key management requirements and trying to make key management a workable solution and simplified. It was recently distributed to STS for feedback, and Bill Burr will be presenting more detail on that topic.

So that's what I have.

MR. CHAIRMAN:: First I'd like to acknowledge John Gayle's arrival. And John, any questions?

MR. GAYLE: Thank you, Dr. Jeffrey. As you would probably realize, this is always quite a test for those of us that don't do this type of work on a regular basis, and you speak a different language than we do in the election community. And therefore I would ask if you could put this in maybe a succinct description of exactly what you're trying to accomplish when you talk about your setup validation and some of the cryptographic changes and granularity. I mean, those are things I don't deal with on a daily basis. I'd like to know what it means and what the implication is for the equipment and for the election officials who will

use that equipment.

MR. HASTINGS: Okay. Do you want me to go ahead and take that one? So my understanding of what we're trying to do with setup validation is to provide the capabilities on the systems that allow election officials to inspect properties of the system so that they can be confident that it's ready for use at the polling place.

MR. GAYLE: Are these directions that you intend to be given by the vendors to the election officials, or are these standards that you're attempting to define that will be distributed universally, kind of a universal design for how set up will be validated?

MR. HASTINGS: The goal is to provide the capabilities in the systems, and it's up to the jurisdictions to decide whether they will use those capabilities to validate the systems, the different areas of the system.

MR. GAYLE: But what I guess I'm trying to understand is, I see what you're attempting to do. But are you saying that this is the singular method in which the local officials can validate the setup? It's the

only method that they can use, or is that something to the management guidelines of the EAC, or is that subject to state law, or are we setting up here a validation method that is a singular method that everyone must follow or else it won't meet the guidelines?

MR. HASTINGS: I hesitate to say that they're methods as much as they're capabilities of the system that are available for use if election officials wish to have them.

MR. CHAIRMAN:: David, do you want to add to that?

MR. FLATER: Maybe I can comment just to help understand a little about what the setup validation is trying to achieve. It's requiring the machines to have a capability so that officials can, if they choose to, to inspect the machines to check some of the things that you might care about if you wanted to check to make sure they're ready for use. So for instance, that includes things like what's the supply of consumables: is there enough ink, if ink is a consumable, for instance. It's checking things like, was it configured as you thought it should be, for instance. It would provide you the capability so you could do that if you choose to. And

one of the other pieces there, which is security relevant, is it includes a requirement to allow you to check what software is currently resident on the machine, what software is currently installed on the machine to allow you to check and to confirm, is that indeed the certified version of software, the version of software that ought to have been there, that hasn't been tampered or replaced, or hasn't been accidentally replaced with an uncertified version.

So the setup validation requirements in the standard would require vendors to provide those capabilities. Some of these have appeared in the 2005 VVSG. It's worth pointing out that this security part of it that allows you to check what software is resident, what Nelson described the current proposal on the table would be a partial relaxation of the requirements. So compared to the 2005 VVSG, the 2005 VVSG required, as I understand it, all machines to have that capability. And this would be a step back from that to say, a subset of machines are required to have that specific capability to check what software is resident, but not all of them. I don't know if that

helps.

MR. GAYLE: Well, that is helpful. And let me ask one other question. Are we talking about the initial setup of new equipment upon delivery, or are we talking about the setup each time that the equipment is going to be used for an election?

MR. FLATER: My understanding is that this is a capability that you would envision could be used before every election. For instance, checking before every election if there's a sufficient supply of consumables might be something you want to do. And so the vendors would have to provide that capability so you could use it if you decide to.

MR. CHAIRMAN: This -- oh.

COMMISSIONER DAVIDSON: One of the things I think that might help you understand that is, a lot of times new software is installed on existing hardware. And a lot of times, the election official is not sure it's been installed on every piece that is within that precinct or within that county. And it has been found before that software has been installed, but it wasn't installed in all of them, and then that caused problems

on election day. So this gives you the ability to verify what software you have in the equipment. Does that help?

MR. GAYLE: Well, that does help, Commissioner. But I guess I'm still struggling with the thought of whether these are for the purpose of new equipment that's being set up for use to ensure that all these different configurations are present as opposed to what you just suggested, and that's existing equipment that maybe wasn't certified under the new 2005 guidelines that's going to have new software installed. It sounds like you're talking about equipment that's under ongoing use and upgrade and update, and so these setup validations would apply to that as well?

COMMISSIONER DAVIDSON: No. What we're talking about, I mean, with the standards that we're talking about developing now is the future. How it's working right now obviously is going to continue working with the VVS 2005, or if you're certified to 2002. But what they're doing here is talking about the future. It may be four years out before your equipment will be able to tell you if you've got software that is updated. You

know, I don't know what the timeframe will be right now, because as we've said it probably won't be adopted until '08, and then we've got to have the meetings with the manufacturers and everybody to see how long it's going to be before you can develop this, and how long is it going to be before we can expect this to be purchased by jurisdictions. So we're talking about the future. It's not changing the past, what we're using right now. This is the future.

MR. GAYLE: Then I guess you would be the right one to answer my question. On e of my concerns is that we maintain a bright line or a fine line between election administration and election management guidelines. And so there are a lot of ongoing setup requirements that are going to be election management issues, not equipment issues.

COMMISSIONER DAVIDSON: Correct.

MR. GAYLE: So if I'm clear about the setup validation, we're talking about precisely the equipment to ensure that it has what it's promised to have --

COMMISSIONER DAVIDSON: Correct.

MR. GAYLE: -- post-certification and testing, or

post-testing and certification?

MR. RIVEST: Maybe I could add for clarification?

MR. CHAIRMAN: Yes. If I could ask people to also give their names just so the record will be easier.

MR. RIVEST: Ron Rivest speaking. To try to help out, I view the setup validation as being sort of the updating of the zero tape situation. With a zero tape, you're checking that certain portions of the machine are set properly at the beginning of election day, that the counts are correct. But with the modern machines, there's many more moving parts to them, software, many other things. And you want to make sure that those are in the proper state as well. And sometimes you could do it as analogous to the capability of printing out a zero tape, but it's just checking many more things.

CHAIRMAN: Are there any other questions for Nelson? Okay. Thank you.

MR. HASTINGS: Thank you.

MR. CHAIRMAN: Next is the discussion on cryptography requirements.

MR. BURR: Good morning. I'm Bill Burr, and I get to do the exciting part, I guess, to try to make

comprehensible the incomprehensible, something like that. When you do a talk like this, the sad fact is that you're either talking over or under somebody all the time. And I don't want to do a talk where I say, this is all magic, I'm a wizard, trust me. By the same token, for all of us at some level of cryptography we cease to be wizards and we have to rely on somebody whose expertise is deeper or better than ours, because various things become very specialized. And there are only a few people often in the world who really seem to understand the guts of certain things.

In any event, what I'm going to talk about here is the cryptography section as it stands now in this current draft. I'm going to walk through what's in it. It's going to be a fairly high-level walk through. I've actually got probably more detail on the slides than I'm going to try to address specifically. But I'm going to point out what I think are the major implications. And in this particular draft of the document, as Nelson noted we have taken the tutorial stuff that was in earlier versions out, I guess largely because in a standard of this sort organized requirement by

requirement by requirement, it's hard to see how to fit a tutorial in. And in any event, you know, the straightforward way to specify this is to write the specification for people who are knowledgeable in the art, and that's what I've tried to do. So I'm going to try to explain here the logic behind what I'm doing. I don't expect election officials to read the cryptograph section and get a tremendous amount out of it directly. I expect people who implement cryptographic stuff to read it and understand what I'm talking about.

So the first part of the cryptography section just sets some basic ground rules. And the first and most fundamental one is that all the cryptography will be done in a validated cryptographic module. FIPS 140 is a Federal Information Processing Standard that outlines a schema for testing cryptographic modules, and it includes a list of approved modules or approved algorithms. And we have a bunch of labs that are actually quite practiced at doing this, and so it seems an obvious thing to do to take your cryptography and plug into the existing federal –

(**END OF AUDIOTAPE 1, SIDE B**)

* * * * *

(**START OF AUDIOTAPE 2, SIDE A**)

MR. BURR: -- have. And then you know at least that the guts of it is good, sound, cryptography. Now, it's never the least bit difficult to take good, sound cryptography and use it in a way that's totally insecure, but at least that's a start. The other sort of general requirement is we specified a minimum of what we call 112-bit cryptography. And all that really means is the generation of cryptography that we're requiring for federal use, which we think will be good for at least another 20 years or so -- beyond that it's actually very hard to make long-term projections. And things like quantum computers cause almost a see change in what's secure and what isn't. And now we have stronger stuff we could give you that might be secure longer than that even, but there doesn't seem to be much point to it. And so that's a list of the algorithms up there. I'm not going to walk through them. The NIST (indiscernible) 157 tells you in some detail what this is. You could use the stronger stuff if you wanted to. It wouldn't bother us.

So what's a crypto module? I thought it would be worth talking about that specifically. And it's basically a separate, distinct program or a device, a piece of hardware, in which you do just basically cryptography. And we have mentioned a test program. And the big distinction I want to make here is that, roughly speaking, you can break modules into two kinds of things: software modules and hardware modules.

And so a hardware module is its own, dedicated little piece of hardware in which you do nothing but cryptography. And typically it's a little microcomputer. Basically, inherently not very different than any other little microcomputer, a $2 part basically, in many cases, once you reach sufficient volume. What we're doing here is fairly conventional, which is to say most of the cryptography that you do in a voting system you can do in software as a part of the general software system that you're running. However, we're identifying particular digital signature functions that are what actually protects audit information as having to be done in a dedicated hardware module. And the reason for that is basically because it gives you an

extra measure of protection against problems in the overall software system and the possibility that there's malicious code in the overall voting system code. It isolates that in a separate little, fairly well-protected sandbox.

So I wanted to give -- yes, Dan?

MR. FLATER: Just to clarify what I think I heard in one aspect of it. So we're specifying certain cryptographic algorithms and we expect over time -- your estimate currently is like 2010 or something, but it could be sooner, it could be later -- one might have to upgrade those algorithms?

MR. BURR: No. Well, let me -- the 2010 date is the date by which we're trying to kill the old generation of stuff that we've had in the federal government since the 1990's.

MR. FLATER: Okay. Got you.

MR. BURR: And that isn't even included here.

MR. FLATER: Okay.

MR. BURR: There's no point in introducing in a standard that probably won't even come to use until 2010 cryptography that we'd like to kill off.

MR. FLATER: But we are talking about something which can change over time, the algorithms. And you introduced the idea of modules. So are we introducing the idea that the guidelines will be that these systems should be designed such that they're modular enough that at given points in time you'll be able to swap the algorithms?

MR. BURR: Well, certainly that's relatively easy to do with software modules. Actually this is something we should get clear. The notion of the hardware module is that certain signature functions are built into the hardware and they don't get changed at all during the life of a machine. And so at some point you'd have to have a new voting machine to replace that, and I'm saying we think we've got at least 20 years of good security in the cryptography. And the thing that really puts the damper on all of this is the possibility of Quantum computers, which fundamentally affect the security of all the public key algorithms we use today. And that's why I'm not willing to go any more than 20 years in my estimate.

MR. FLATER: So you're not requiring the ability to

upgrade it? You're just saying that at a certain period of time (indiscernible).

MR. BURR: I'm saying that at some point, you know, in another 20 years, then -- this is the computer world and I realize that the election machines have traditionally been used for very long periods of time. But I can't project security well enough to want to specify anything beyond about a 20-year period or tell you it's good.

So I wanted to say a little bit about public key cryptography.

MR. CHAIRMAN: David, did you want to say something?

MR. WAGNER: David Wagner. I'm not sure if you got an answer to your question. I don't know if we've discussed this among STS. My feeling is no, it shouldn't be necessary to require the ability to do field upgrades on your crypto algorithms to voting machines, that crypto, as Bill explained as well enough understood, that the crypto algorithms put in place ought to last for the lifetime of the voting machine.

(UNIDENTIFIED SPEAKER): Right. That's the

clarification I got.

MR. BURR: But the only thing that really this really requires that puts a limitation on that is the signature part of the hardware module, because it will be easy enough to replace anything that's done in software, which in most cases is what people will choose to do, because the truth is, the processor in your overall voting machine is likely to be rather more powerful than the one in the signature module itself. And because basically the -- what we're requiring the hardware module to do is very specialized. So there's a lot of things that could be upgraded but, you know, I don't see any real that they should be any time soon.

So now, public key cryptography, this is something that's actually pretty recent. It didn't exist really 30 years ago, but I think by now just about everybody has heard of it one way or another. And of course it goes with this sort of awful initials, PKI, that kind of terrify people. And the concept is really pretty simple. You have two mathematically related keys. There's a public key that you can make public, and it's usually presented in something called a public key

certificate. And you can use this public key to either encrypt data or to verify a digital signature. And then associated with it you have a private key, and that really has to be kept a secret. With that private key you can either decrypt encrypted data or you can use it to sign a digital signature.

And it's the digital signature operation that's the key operation for what we want to do here. If you think about it, for most -- I wouldn't say for all, but for most election systems, I don't think there's an awful lot you're going to actually be encrypting. Possibly if you send results electronically back to an accounting center or whatever, you'll want to encrypt them while they're traveling over a network, or something like that. And in some kinds of schemes, we might get into a -- when we go beyond the innovation class systems, there'll probably be more uses for encryption. But the big thing here that we really want to use cryptography to do is to protect and authenticate records in terms of guaranteeing their authenticity and they haven't been altered.

So I guess I've already talked about digital

signatures at this point. So what we do with the digital signature basically is first we generate something called a hash of whatever it is we're going to sign, a relatively short, compressed representation of it typically. In the next generation of stuff that we're introducing here, 256 bit digest of the message. Then we apply the private key to it and we get out the signature.

And so if you want to think about what actually goes in on the voting machine typically, the general software of the voting machine probably does this hash. And then it passes the hash to the little hardware module that I've already mentioned to actually perform the signature operation. And that's basically what's going on. With signature verification, whoever's verifying the signature takes a look at the message, generates that same hash from the text of the message, then applies the public key to the signature field on the message, and at least in the simplest scheme, compares the hash that it then gets as a result to the hash that's on the message. And if the two are the same, the message verifies.

The verifier then knows that he has an authentic message if he had the right public key, and that not a single bit of that message has been altered in any way since it was signed. So this authenticates the message. And the practical implication of it is that it largely eliminates chain-of-custody issues. If you've got a good, signed message you really shouldn't care how you got it. Whether it was sent by passenger pigeon or just given to you or you found it on the street, if you can verify the signature, you've got a really strong check that it's an authentic message and it hasn't been altered.

So the point is, up until you apply a digital signature or some other cryptographic technique to data, there's nothing in the world more fungible, alterable, forgeable, changeable than data. But once you put a good electronic signature on it in a signature scheme, you've really locked it down in a way that's actually stronger than you typically get with paper, because with paper you're looking at a document and it's providence and how it was cared for. And then if you're worried about somebody altering it, you're looking at very

detailed forensic evidence to see if you can find evidence of alteration, or evidence that the paper is forged somehow, that the whole thing is a fabrication.

But with a digital signature you really lock things up in a good, solid, very easily-verified format. So we're interested in this because we want to produce electronic records, particularly audit records, that we can sign and be pretty darned sure that they haven't been messed with, fabricated, forged, altered, changed in any way since they were signed.

MR. CHAIRMAN: Hold on a second, please.

MR. GAYLE: John Gayle from Nebraska, Secretary of State. Just so I can catch up with you then, what you're talking about here is really an auditing, post-election function, the transmission of information by some method that needs to be encrypted to ensure its integrity while it's being transmitted? So we're not talking about a function during the election.

MR. BURR: I'm talking about signing the data as you create the audit records during the election. And then whenever someone, say, examines those audit records, being able to verify those signatures and

verify the authenticity of that data.

MR. GAYLE: And we're primarily talking about DRE equipment then, I presume?

MR. BURR: We're talking -- yes.

MR. GAYLE: That has a digital scheme.

MR. BURR: If you look at what's -- and what will follow in John's talk, what we're interested in being able to do more than anything is with DRE equipment and the human-verifiable paper audit trails, we want to be able to rigorously cross check them. And we want to be sure that the electronic audit records that we're cross checking have not been diddled with somehow. And so --

MR. GAYLE: I'm going to ask a lot of probably what sound like kindergarten questions, but I've got to ask them on behalf of election officials who don't understand this any better than I do. And I've been reading the minutes and the resolutions in the material prodigiously, and some of these things are still confusing. But what we're talking about is the outcome, I guess, the result that you want to audit. We have the voter-verifiable question as one form of auditing, but that's not what we're talking about, where the voter

attempts to verify what he cast. You're talking about a different form of authenticating the outcome, is that correct, with your cryptology?

MR. BURR: What I'm talking about is using the cryptography to create electronic records that can be carefully authenticated or fully or completely authenticated, so that in a layer-audit stage where you're actually comparing typically the paper to the electronics and making sure that they're consistent, then that you can be sure that the electronic part of it is authentic.

MR. GAYLE: So its' -- let me just finish my question. John Gayle. So we're talking about a triad here of voter verification on one hand. In the course of the election cycle if errors arise, hopefully a voter checking their ballot representation might see an error. Then you also have the paper trail so that you have another form of verifying a digital vote cast that can be used for recount, for example, or for partial audit. And what you're talking about is, I think, a third thing which is to make the source code or the security of the digital imprinting so safe and so secure that it's

beyond question, beyond debate as securing the signatures, I guess you call them, with integrity for recount or for some other purpose.

MR. BURR: What you want to know is what machine it came from and that since it was produced by that machine it hasn't been altered at all. And what we're actually worried about being able to reliably catch more than anything is the possibility of malicious code in the voting machine of printing one thing on the paper and putting something out electronically that's different. And overall we're just trying to come up with a system of -- I mean, this is really John's talk, not mine. And he'll go into this in some gory detail.

MR. GAYLE: As we say in court, though, you're our present witness.

(Laughter.)

MR. BURR: I'm the present witness but the present witness isn't as well prepared to --

MR. CHAIRMAN: Feel free to have John come up next to you to help answer the questions as well.

UNIDENTIFIED INDIVIDUAL: Could I comment?

MR. BURR: Certainly.

UNIDENTIFIED INDIVIDUAL: Maybe it would help to understand the purpose of the cryptography here. Maybe I can relate this to some things that we currently do procedurally in our current voting system. When you close the polls on many voting machines, typically there's some memory card or movable storage media where the votes are stored electronically on that memory card. And then it's very common in many places that memory card will be transported by poll workers back to county headquarters. And many places have a chain-of-custody requirement. There have to be two people accompanying that memory card or other requirements to ensure that it's not tampered with while it's in transit.

One of the things that cryptography can do for you is to protect that data using mathematics in a way that prevents tampering with the data on that memory card while it's in transit after you've closed the polls while it's being transported to the county headquarters or being stored there. So what that does is it reduces or maybe even eliminates the requirement for this two-person control on the memory card. It eliminates the opportunity for swapping of memory cards or sort of

modifying the data on the memory card while it's in transit. This is a different issue from the voter verification. This doesn't ensure that the vote was recorded initially as the voter intended. It just means that while it's in transit, it's not going to be tampered with.

MR. GAYLE: Well, that's why I asked the question about the transmission issue. This is a transmission issue --

MR. BURR: It's a transmission --

MR. GAYLE: -- moving the end result to some other tabulating or counting center. Is this what then we talk about in terms of hard wiring each machine so that it has a very specific encryption and can only be used for that particular precinct? It doesn't have mobility. Is that what you're talking --

MR. BURR: I'm not trying to do that in this. We're certainly, definitely not trying to do that. We are trying to be sure that you can always tell which machine actually generated these records. But we're not trying to actually, in the cryptography section for sure, specify anything about whether one machine is

producing records that are somehow tied to a particular polling place or not. That's kind of beyond the purview I think of certainly a cryptography module. And I suppose people might want to choose to set things up that way, but it's not the intent of this document to pin you down like that at all.

MR. GAYLE: Thank you. I don't have any other questions other than this is very helpful to me, and I appreciate your explanation then.

MS. QUISENBERRY: This is Whitney. If I could just ask, I know there are some questions about fingerprints, but are you essentially saying that each machine has a unique identity that can be known? And so that if I've decided that Machine 1 is in Precinct 25 and I get results back and it says it's from Precinct 26, I know that that -- let me back up -- I know that the results I got came from that machine and not from other source?

MR. BURR: You know what machine it came from. If you thought the machine was in Precinct 25 and the ballot on it says it's 26 somehow, the other information on it, you've got an inconsistency in your records somewhere and you ought to be looking into something.

But this is above the level of the cryptography spec. The cryptography spec is just saying it's authentic.

MS. QUISENBERRY: Right. So again, you're providing a capability that can be used as part of election process, rather than requiring election process?

MR. BURR: Right. That's the idea. And so I've already talked a good bit about the signature module. It's a separate chip, a separate microcomputer, a couple-dollar part once you get it in volume, but of course there'll be some serious development costs associated with making sure you've got it right in the first place. And one of the things that we've chosen to do in specking this is to require that it generate its own keys, because we want to try and make the operation of this thing as sort of seamless and transparent as possible, and also because having it do that actually eliminates a number of ways that people could attempt to diddle the system and manipulate the keys. So the private keys that are used in the signing operations, the idea is to design the module so that the private part of the key is generated on the module and it never

leaves the module. And this is one of the ways that we help to prevent the effects of malicious overall system code or code that's been compromised from tampering with the results successfully.

MS. QUISENBERRY: Sorry. When we say hardware, I know that I have some pieces of software that require me to plug something into my machine and the software won't run unless I can prove that I have the one and only little hardware thingy that I plug in. Is that the sort of -- I mean, I know they're probably not the exactly same strength, but is that the sort of thing we're talking about?

MR. BURR: What I mean in this particular case is that this is not something you plug onto the machine. It's something that's actually permanently soldered into the machine, so that you can think of it as -- or when we reach a high enough level of integration in these parts, it might be just a separate little piece of the actual chip that does everything. But it's a separate, physically distinct device that is probably in most cases a separate little microcomputer with its own memory.

MR. CHAIRMAN: Bill -- I think the answer is yes.

MR. BURR: Okay. Fair enough. The answer is yes. So this is just a list of the capabilities of the signature module. It has to generate the key pairs that imply some stuff about requiring a random number generator on the device. It has to be able to -- it is a public key certificate that identifies the public key and it has to be able to store and output that and to create those. And everything else really can be done that you want to do cryptographically. And there's a surprising amount of cryptography that goes on in any computer system, whether you know it or not. It can just be done in general software on the voting machine.

UNIDENTIFIED INDIVIDUAL: Just for amplification, generally speaking it is good practice to have this signature module be in hardware that cannot be accessible by anything else, because what's fundamental is that that private key that he's talking about can't be known by anybody or accessible by anybody because that's what's taking these elements of the record and binding it, signing it, so that in that transit it can't change it. So you want to make sure that nobody can

have any software where they can sort of access that key or make that new signing, because then that could tamper it. I mean, then I could find a key and somebody's doing the transit and I could modify and they could resign it. But if it's embedded in that machine and you can't pull it out and you can't access it and find out the private key, they you've achieved the objective. So for most applications in banking and so forth, we always insist that the private key be in hardware not accessible by anything.

MR. GAYLE: If I may, Dr. Jeffrey. John Gayle, Secretary of State, Nebraska. Well, this I guess is a bit of my concern, and maybe it's not justified. But in virtually every election, there's a lot of mobility of equipment between precincts. One precinct has doubled in size over the course of this summer and equipment needs to be assigned to that precinct. I just want to be sure that we're not encrypting machines in such a way they can only be used in a particular precinct and don't have the mobility that we're used to having.

UNIDENTIFIED INDIVIDUAL: If I can respond to that, yes, that's an interesting concern, but it's not one

that’s a problem here. The goal here is really to give every machine its own identity so you know what data comes from that machine. The roles and the positions, the precincts, what those machines do, whether they’re agnious (phonetic sp.) tabulators or vote-capture devices or something, all of that is a higher level of management. And there’s absolutely no intent here that any of this should eliminate any of that flexibility that you may want to have in managing election to the best use of the equipment.

MR. GAYLE: Thank you. That’s very helpful.

MR. BURR: So the next couple of slides talk about the way we’re managing keys that we’re creating on these modules. And basically there’s two kinds of keys: there’s a long-term device signature key that basically we’re requiring it comes from the device with the factory, it lasts the life of the device. And then there’s a short-term signature key that’s created for each separate election as a part of the start-up process for the election. And it used to sign the records for a single election, and then when you close the machine out at the end of the election the key is destroyed. So it

doesn’t exist anymore. It can’t be used by somebody, even if he gets possession of the machine to later fabricate another version of the records for that election.

So that’s the scheme. It’s intended that if somebody does a good job of implementing it, it’s almost transparent that this is going on in terms of, you have to set machines up for elections anyhow. And it should be an automatic process to generate these keys. You have to close machines out at the end of elections, and the destruction of the key should be an automatic process. And the necessary records should be automatically created as part of these things so that the actual people doing this should hardly even be aware that this is going on. That’s the intention.

The device signature key, which is the permanent key, the one other requirement is that it include in the certificate that goes with it the manufacturer model and serial number, at least whatever the actual identifier of the machine is, and that same nomenclature appear on the outside of the device so that then it’s relatively easy to match the two up. If you have the electronic

version of the certificate that tells what the public key is, you should be able by just looking at the outside of the machine to know which voting machine that applies to is the intention.

The election signature key as I said is generated per election. And one important point is that you keep counts of the number of keys of these that you generate and the number of times that each private key is used. And as a part of the auditing process, you should be able to account for every certificate you create. And every time you sign something you'll have to be able to produce the record that was signed when you did that.

When you close the election out, you produce a signed-out that tells how many times that the key was used and the key gets erased. So the idea is to be able to account for all of the audit records. And they give you automatically in this little hardware module everything that you need to do that.

So the basic summary of this is we are calling for a hardware module to do signatures. That adds some extra cost to the voting machine as opposed to doing it in software. It gives us extra protection against

software that's been tampered with or people who actually get physical control of the voting machines. There's a permanent key that is associated with the device. There's a new key for each election. And we've tried to do everything so that it adds, that the management of these keys adds very little extra to the overhead that's already involved in setting up the machines, running them, and closing them up.

And that's basically the whole talk.

MR. CHAIRMAN: Thank you for Crypto 101. Are there any additional comments or questions? Obviously it's very detailed. It's a very important section, certainly not something in the normal vernacular and issues that people deal with, and again if implemented properly would be essentially transparent but will ensure integrity of electronic records. Are there any other comments or questions? John?

MR. KELSEY: So I'm going to be doing the opposite of what Bill was doing. He was talking to you about something that mostly, you know, we're experts in and most of you aren't. And I'm going to talk about something that I'm not an expert in and you guys are.

So it’s going to be a little different. This is like the student getting to teach the class but the teacher’s in it, or something.

So this is a talk on equipment requirements that support required auditing steps. So I want to be clear up front. What we’re talking about here is requirements on the equipment to make sure that they can support these auditing steps that address known attacks, known threats. So we know that there are threats to voting systems that can only be addressed procedurally, right? That’s pretty obvious. Everybody knows that’s involved in elections. And with this nice requirement for software independence from the previous TGDC meeting, software independence means that the voting system, that an attack that is involved in just tampering with the software on the voting machine can be detected, that it’s possible to see from the behavior of the voting machine that there’s an attack going on or an attack has happened.

And what we want to do here is talk about what is required, what procedures have to be supported so that it will be detected with high probability. So basically

this amounts to requirements on the equipment, kind of requirements on what the equipment does, on the documentation, and on how the equipment is tested. And at a high level, all of this is going to apply to the innovation class, but of course we don't really know much about what that's going to look like yet, so it's hard to nail any of that down.

Part of the threats we're addressing, mostly what we're talking about here is threats that involve tampering with the software on the voting machines. So we want to say, given that we have these voting machines that have that paper record that the voters can verify, what could happen if somebody tampered with the software, and then what defenses are there to make sure that that would be detected. And then, like I said, we want to make sure that those defenses can be used by election officials, given what the equipment does.

So at a high level, what we're really doing is, there are two different kinds of attacks that we're worried about that involve tampering of the software. One kind messes up the agreement between records that we have. So, for example, if the voting machine just

silently changes a vote -- so you think you voted for Smith and it records a vote for Jones electronically and prints a vote for Smith -- there'll be a discrepancy in the records. The electronic records in the voting machine will say one thing, the paper records will say something else. And if you check those records against each other, you'll catch the attack.

The other kind of attack that we worry about involves the presentation of the choices to the voter and the machine behavior. So, for example, if the machine introduces errors that kind of favor one candidate over the other or one question or one outcome over another, or in the case of observational testing, if the machine sometimes just tells you that you're voting for Smith and then prints a vote for Jones on the paper record and records it electronically, those are things we want to make sure we can catch.

And so there are two different kinds of classes of attacks and two different kinds of classes of auditing steps that would be supported.

Let me just skip ahead to the diagram here because it's a lot nicer, if I can get it up here. It doesn't

quite fit on the screen. This is a picture I put in to try to explain some of what we're doing here. The idea here is these three blued areas are sets of records that are produced by the voting system, by different parts of the voting system. And I'm not sure this is exactly the right way to refer to these, of right, proper, technical terms, but the idea is there's this process where voters check in, there's something entered in the poll book. And at least normally it's a manual process. So you've got this sort of poll book audit, this ballot accounting that needs to be able to be done so that you can verify that the number of ballots that were cast is not way higher than the number of voters that came in. That would be an obvious problem. And that each kind of vote, each kind of ballot, you know how many of them were given out, how many were received.

You have this requirement for this check that is a hand audit that we normally talk about with paper records where you're just looking at the voter-verifiable paper records, looking at the electronic records that are kind of the summary or the outcome of the election for that machine or that precinct or

whatever, and checking them and making sure that they're the same, that they agree. And there's also this kind of check here where you're going to make sure that the electronic summary of votes wound up correctly in the final election report. That's part of the canvassing process, I think, to look at that (indiscernible) make sure that the available security is there, that all the voting equipment provides all the information necessary in these reports to make these audits as easy as possible, and that we get the advantages of security, in particular going from here to here because we digitally signed this in this nice way.

These are all kind of, I think, motherhood and apple pie requirements. I don't think there's anything that's very controversial in these first three at all. So poll book audit -- and really what we want to do is make sure that the voting system provides enough information that we can catch it if, for example, a voting machine -- say that you have a voting machine with the VVPAT. If it were to wait until a quiet time when nobody was watching and then electronically record and print out three or four extra votes, you'd want to

make sure that that got caught. And this is how you catch that. And so you just want to make sure that the different summary records or the different records from the system give you enough information that you would reliable catch that, or that you could if you did these auditing steps.

And so the picture here is you just want to make sure that ideally this election report that comes out of this tallying process gives you this breakdown by polling place of how many of each kind of ballots were cast. And you can check that in a fairly straightforward way. If there's other information that needs to be included to make this work out, I'd like to get some feedback on that.

The other thing that's going to happen is we want to make sure that it's possible to take these electronic summary records and make sure that all of the same information is included from everywhere. So, for example, all this information about how many ballots of each type has to be available from the set of paper records, the set of electronic records from each machine, and the final report so that you can make sure

that they’re all in agreement when you do each of the auditing steps.

Similar things with the hand audit, essentially I was kind of surprised that this -- and the ESI report on the Kihoga (phonetic spelling) County recounts, there actually were some pretty surprising problems as far as not having all of the information necessary on the paper records to unambiguously figure out which paper record went with which machine and with which electronic record. And so we want to make sure that that’s required, that every paper -- that if you have a paper roll, that each paper roll has to have the identity of the machine on it, which election , which ballot styles are used, all that stuff. And then you have to deal with marking the write-ins and provisionals, because those have to be handled differently.

And the final election report also needs to show -- you break this information down. It needs to be possible for the tallying process to provide you with a breakdown by voting machine or by polling place or precinct, depending on at what level you’re going to do this count. As a rule, what you want to make the count

reasonable is the smallest set of paper records at a time that you can count. You don't want to require people to hand recount everything from the entire polling place. So it's all pretty straightforward.

The last bit is where we get into some of the cryptography. And Bill's already kind of talked about that. The idea is you want to be able to reconcile that the final election report included the totals for each machine and all the information you need to verity that it was included correctly. So the idea here is these electronic summaries are digitally signed, and they're digitally signed in a way that is bound to a specific election and to a specific machine.

So once these are produced, it's kind of committed to by the machine and the machine loses the ability -- even if you were take the machine apart and get the key out you couldn't go back and produce and backdate your records. It's kind of a nice feature of this. And so when we electronic summaries of the votes cast on the voting machine and they're digitally signed and we have this final election report, if we wanted to in principle we could put these out in the public, we could post

these on the web or something and everybody could check that this summary was actually included in here, that each summary for each machine was included in the final report. There are some privacy issues you'd have to deal with there involving provisional ballots and write-ins, because those won't have been resolved yet. At the point where the machine commits to its totals, it won't know how to resolve the provisional ballots. So if they're included on the normal electronic record you have to deal with that, and there are some ways of doing that where you could aggregate those into a different category on the final report.

But the (indiscernible) here is that this part right here, I believe becomes much stronger because of the cryptography, because now there's no question of -- you can look at this electronic summary, you can print it out, and you can verify that it was included correctly in the final report. And you can verify the signature here and here. And it's kind of a nice, you know, adds some verification that nothing's happened in between the time when that was committed to and the time when it was counted.

So as a summary, I don't think anything here is very difficult or surprising. And I think mostly this is done. We just want to make sure that it's written down and that it's required. All the data needed for these auditing steps that we know address specific attacks needs to be included in the outputs and the reports in the voting system. We want the electronic records to be digitally signed because that adds security at essentially no cost. It's just using the existing tool exactly the way it's designed to be used. I believe these requirements will have no impact or very little impact on the cost of the voting equipment or operating it, other than what's required to get the crypto module in there. But everything else I think it's just you change a little bit of software to make sure you can generate all the right reports.

So you have any comments or questions on this part?

MR. MILLER: This is Paul Miller. I have a question.

MR. KELSEY: Yes.

MR. MILLER: You made the comment about unambiguously identifying the paper tape as part of the

VVPAT.

MR. KELSEY: Yes.

MR. MILLER: And you eluded to Kihoga County. In terms of my analysis of Kihoga County, at least some of that, my understanding comes from broken paper tapes that they didn't get the second half of the tape together with the first half, and that they switched printer modules from one machine to another machine. Are we contemplating some kind of requirement that the machine be able to fence when a new paper roll has been inserted and prints the identifying information at that point in time?

MR. KELSEY: Yes. I know we've talked about that. There are places where this touches on reliability requirements, that I think are dealt with in core requirements as far as like having the paper tear easily, being able to change the paper rolls without causing problems. But in the paper records requirement that we've been working on, one requirement that we know has to be there is that if you change paper rolls, the machine has to know that you've changed paper rolls and be able to print identifiers, machine identifiers and

everything on that paper roll. Dealing with the special cases where you have a jam or the printer fails or something, that probably needs more working out in what we've written. But I think those are all really important issues because that's the place where you're going to get these breaks.

MR. MILLER: Okay. Thank you.

MR. GANNON: Where in the VVSG is all of this auditing information requirements going to come in?

UNIDENTIFIED INDIVIDUAL: I believe -- John (indiscernible) talked about this. I think there is some question of whether it winds up in the security section or in a later section. And I forget where it is in the current outline. There's an entry for it although it's not filled in yet because this is still in the process of being edited. Actually the big concern we have is this is something where we need a lot of input from election officials, because they've actually done these audits and they'll be able to point out things we're missing.

MR. GANNON: I'm sorry. I didn't identify myself. I'm Patrick Gannon. I have a follow-up question.

MR. KELSEY: Yes.

MR. GANNON: For electronic records, how is that being addressed. I mean, so far the only thing we have in there right now is interoperability and some high-level requirements on interoperability of common data formats.

MR. KELSEY: Right.

MR. GANNON: But where you have electronic records, what are going to be the audit requirements around that?

MR. KELSEY: Well, the electronic records, we have a chapter on electronic records that has been sent out to the STS but hasn’t been put out here. It’s still being worked on. One of the requirements we have there is that the electronic records have to be produced in a completely specified format, so that if you need to you can write your own software. You don’t have to just depend on the vendor software to give you all the information. There’s been some thought about using like the ML (indiscernible) and there are some issues with that. I think Dave Flater can probably address those. But there’s a lot of detail there and I’m sure we can get that out to you if you’re interested.

(**END OF AUDIOTAPE 2, SIDE A**)

* * * * *

(**START OF AUDIOTAPE 2, SIDE B**)

MR. GAYLE: -- to perform their job it seems like TGDC is starting to drift somewhat into the other half of election conduct, half of which is equipment and the other half is administration. And in looking at your chart and looking at some of the procedures you're suggesting, all seem to be far beyond the equipment. They have to do with the conduct of the election administration, which seems to be purely an Election Assistance Commission issue and not a TGDC issue.

MR. KELSEY: Do you want to say something or do you want me to?

MR. WAGNER: David Wagner. I'd say that's a very fair concern. I don't see that as an issue here. My sense is STS has asked NIST to look at, to go understand what are the auditing procedures that are typically used by election officials around the country and to develop requirements to ensure the equipment can support those, how election officials are using equipment. So this is not at all, not by any means, mandating the procedures

that election officials would use or how election officials have to do the audit. Rather, it's ensuing the machines provide the information that election officials would need to be able to do those audits and to make it easier to do those audits. But whether those audits are done and how they're done is entirely up to the officials. It's not something the standard would regulate or require, or have an impact on.

MR. GAYLE: Well, will you include this in TGDC report to the EAC? Since it's not a standard that can be tested to or certified to, it's up to state law and --

MR. WAGNER: David Wagner again. Maybe the way to explain this is that this survey of audit procedures is the background that will inform the drafting of requirements for equipment specifications. And what those equipment specifications might say are things like, the machine ID must be printed on any VVPAT record. And that was then informed by the research, the survey of audit procedures which came out of that survey where it was discovered, oh, election officials are having a hard time using these VVPAT records to do

audits because it didn’t have a machine identifier printed. So that is a requirement that we can make on the equipment that is testable and can support the election officials’ needs.

MR. GAYLE: And that makes sense to me in terms of helping ensure that the equipment provides the availability of information to do audits. But to go ahead and say, these are the kind of audits that you should do, seems like we’re beyond the equipment then or the availability of information from the equipment.

MR. CHAIRMAN: This is Bill Jeffrey. The next iteration of the VVSG will not include procedural issues that are done at the state and local level and (indiscernible). As David said, it would only ensure that however you do it, hopefully somewhere we have captured all of the data necessary so that your implementation of your audits, you will have all the information reliably with the fullest integrity. And so in reality, the requirements will encompass things that vendors will need to do that Nebraska won’t use. You may use a subset of it, Colorado may use a separate subset of things. But it’s a way to try to -- what he

was describing was trying to get an understanding of how any audit is done so that all of the relevant data or however you do it is captured. But it will not tell you how the state of Nebraska would ever do an audit.

MR. GAYLE: And that was my concern, because the EAC will be coming out with election management guidelines sometime later this year, maybe about the same time that this hits. And the two things need to be compatible. If you're going to set some standards or guidelines for how to conduct precinct audits, I think we would be way --

MR. CHAIRMAN: Absolutely not. This is Bill Jeffrey again. This is requirements on the hardware, not on the procedures. Whatever procedures are generated, it will hopefully have all of the data they need. And what we're trying to do is capture it to make sure that anything you could possibly want in your audit is put into the requirements and put in such a way that it is secure and integrity is assured.

MR. GAYLE: Thank you.

MS. PURCELL: If I could -- Helen Purcell, Maricopa County, Arizona. If I might, Mr. Secretary, having just

gone through a hand audit in the last general election, it would be impossible to do that if the equipment did not give you, as they stated, the precinct number and various identifiers that you had to have in order to do the hand audit. Whether it's on electronic voting machines or it's on your optical scan machines, and so forth, you have to be able to identify that to complete your audit. And it's a really important thing, I think, that equipment has to be like that. And we also see in the current session of Congress there are a number of bills that will require states to do audits of some type, mostly by hand.

MR. GAYLE: And I agree. I guess if you took -- this is John Gayle, Secretary of State, Nebraska. If you took the example of the chip that's going to be taken from the machine and then transported to a central tabulating office, that chip needs to be encrypted to identify the machine so it can be received and identified at the counting office. But the issue of whether two people accompany it or four people accompany it or what kind of car they drive seems like that is not an issue for TGDC. That's why I distinguish between the

two areas.

MR. KELSEY: Yes, I have to say I'm very deeply aware of my ignorance in the depths of election procedures, so I'm not going to try to write a procedures manual. I wouldn't be qualified to do that. So no question about that.

So if I can, I'd like to go ahead and talk about the more complicated procedures. So we talked about the ones that are just checking between records. These are already done, and we're just kind of saying well, the equipment has to give you all the information necessary to do them.

And the second set here, we're talking about things that either aren't done or they've just been done a little bit like parallel testing. And a requirement here is to verify that the machine is behaving correctly in ways that wouldn't -- that it's not carrying out some attack that wouldn't leave a discrepancy between records. So let me talk about this. Even though you've got a voter-verifiable paper record, machines can still certainly misbehave. One of the obvious ways which you hope that voters will catch in most cases is that it

could indicate a vote for Smith to you on the screen, and it could print a vote for Jones on the paper and also record a vote for Jones. There is now no discrepancy between records, it's just that the voter gets a chance to notice that. And there's an obvious problem, because if you're blind then of course you're not able to notice that, you can't look at the paper. And you need either an additional procedural defense or an additional technical defense to make sure that blind voters can't have their votes stolen.

Another issue that you have is that the voting machine could introduce sort of differential errors, errors that favor one side over the other. There is, I think, a known set of attacks on optical scan systems where if you misprint the ballots a little bit you can cause the scanner to catch the same vote for one candidate and not for the other. So the same sort of thing, there also have been problems with (indiscernible) where the screens misalign by mistake and you get these sort of errors. Well, in software you could certainly simulate the errors. You could make the errors benefit one candidate over the other. And so

that’s something that you would want to be able to catch.

There are other things that you can do as an attacker trying to attack an election, even though there’s a paper record and it’s being audited. And the kind of nice thing about this is that mostly these threats are easier to detect, because the voting machine has to misbehave in the sight of the voter. So there’s a good chance that the voter, especially a voter who is pretty aware of what the ballot’s supposed to look like, or somebody who’s working in the election or something is likely to notice there’s something odd going here and maybe complain. The problem you have is that when you just have a few people complaining, it’s not actually clear what you do next. There’s no clear place where you can check two different sets of records or have some procedure where at the end of it you know unambiguously what’s happened, that you’re being attacked. And I guess that’s a pretty common situation.

The other issue is that the blind voters and a whole set of voters who aren’t able to verify the printed record need some sort of additional defense, or

they don’t have security against software attacks. One comment I’ll make about this is we don’t have nearly as much experience operationally with doing this kind of thing. We have some experience with the states doing parallel testing -- I know California has done that -- but not a lot. And the observational testing is a defense we’ll suggest here isn’t something that’s been done before, as far as I know, formally, although I think it’s done informally.

We talk about observational testing. This is something that’s come out of the discussions about how to implement the full resolution that we had on software independence. We said that essentially, if I can summarize it, that we want software independence and we need it to work for blind voters, too, and it needs to work for everybody. So the threat that you have is a voting machine -- if there’s tampered software on the voting machine, the software could use the fact that you’re using an audio ballot or a screen magnifier or something like that as a clue that you probably won’t be checking the paper. And so it could change your vote on both the paper and electronic record, and if you can’t

check it there is no way for that attack to be detected because the records will all agree.

So there's sort of a simple procedure to address this that gives you some reassurance that this is not happening, which is to have a small number of authorized voters volunteer to use the audio ballot or the screen magnifier or whatever, and to carefully check the printed record. And the goal here is, if you kind of think about the attacks, the attack program can no longer reliably just change the printed record and know that it can get away with it. So 100 people in a state checking this are very likely to catch any kind of an attack that changes a large fraction of the votes.

And this is kind of a nice thing, just because the actual requirement on the equipment is really minimal. The requirement on the equipment is on the mechanism by which you authorize the voter to vote on the voting machine. It just has to be something where you don't just hand blind voters a different kind of authorization or something. It has to be possible for anybody to use the audio ballot or the screen magnifier. And I think that was already something we wanted to do.

Parallel testing is more problematic. It's a kind of powerful defense against these attacks where the voting machine misbehaves and tries to confuse the voter or introduces errors in favor of one candidate or something. But if you look at this, kind of the threat here is the voting machine is doing something, is misbehaving in some way that would only be detected if you watched it carefully on election day. And the assumption here is this isn't something that's caught by normal testing. So we're in the realm of malice here, we're in the realm of somebody putting software on the machine that is actually going to wait until election day and then misbehave only on that day. And so the kind of defense which I think was proposed originally by Mike Shamos is to do some testing on a few machines on election day and see if they misbehave. And then of course the requirement is that this has to look to the voting machine just like a real voting, like a real election.

So you can develop a lot of requirements here, but at a high level if you want the parallel testing to work, what has to happen is you have to be able to

isolate the voting machine so that it can't get any communication from outside, so the person running the attack can't tell it okay, you're being tested, don't misbehave, and the voting machine mustn't detect that it's being tested. And those are of course two high level -- actually doing, to say anything directly about the equipment. But we can go a little further down and we can say, you know, if you want to make sure that you can isolate the voting machine, that means that the voting machine can't be talking to other devices in the room, it can't be on a network. And that causes problems, because then that limits the set of possible designs.

So you kind of have some ideas of what you might have to do in order to support this. But the requirements to make sure that you can do parallel testing are to actually impose some real constraints on the design of the voting machines, and things like not being able to network them, things like the way that you do the authorization for voters to vote, it has to be something that the testing team can completely take over and use so that there's no way for the voting machine to

detect it's being tested.

This is something where we're still trying to figure out what makes sense. And we need feedback and we need discussion in the STS about this I think, about does it make sense to require support for parallel testing and how much.

The last piece of this is much simpler. If you have a ballot marker that doesn't record, that doesn't have any memory, you can do something a lot more like observational testing. You can just have the voter -- or you can just have somebody go in and, during the election, cast one test ballot on the thing, get a printed ballot, and use procedural mechanisms to make sure that that ballot that's been printed out is correct, that that isn't included in the final total or anything. And the only requirement on the equipment there is just on the authentication mechanism again, to make sure like the poll worker doesn't have some way that they can tell the voting machine or the ballot marker that it's being tested. And I think that's pretty straightforward.

And that's it for this set of procedures to address

these presentation attacks. The observational testing is straightforward and powerful. I don't know if it resolves all the problems with that, but it's at least something that's pretty straightforward and doesn't impose a lot of requirements. The parallel testing requirements are something where I think we need more discussion, because we need to see if it make sense to impose these requirements.

So is there discussion or questions?

MR. CHAIRMAN: Thank you, John. The other John. Sorry.

MR. GAYLE: Thank you, Dr. Jeffrey. John Gayle, Secretary of State. It seems to create and fabricate all of these imaginative defenses to what seems to be an issue with source code initially. If we're talking about malicious attacks and not random bugs, from what I read in the past minutes it sounded like it is difficult to review source code for, let's call it a large operation or a large system. But that what we're talking about in terms of elections is we're talking about maybe a megabyte of code, is what I read in the minutes. I just take it from the minutes. In other

words, a small amount of information, a small amount of code. Why isn't it possible, if we're going through the testing of the source code as part of the certification of equipment, why does it sound like there's such an immense likelihood that you're going to have malicious errors, virus in that code, which now we're constructing a lot of defenses to deal with? Does that make sense?

MR. KELSEY: Yes, I understand your question.

UNIDENTIFIED INDIVIDUAL: Perhaps I could speak to that? Yes, good question. I think that there's layers of defense here and various kinds of threats. So the source code review will be imperfect. The source code is just too complicated to hope to catch all bugs there. But I think the primary concern with parallel testing is related to the set of validation issue, too. I mean, the source code may have been manipulated as well. What you have on that machine may not be what you thought you had on that machine. So the question is, is the machine behaving inappropriately for some other reason other than what the source code may have said.

MR. GAYLE: John Gayle, Secretary of State. Well, I guess I haven't seen any evidence that any of these

things we're talking about have occurred in any equipment anywhere in any system. So we're really constructing an issue here that's how many angels on the head of a pin. How many ways can you protect against an imaginary foe, the imaginary foe being malicious construction of the source code by some evil person? It just seems to me if we're going to spend all this money on all of these back-up ways of auditing against source code intrusion, why don't we just focus our attention on preventing source code intrusion and not all of the variables to prevent consequences?

MR. CHAIRMAN: David?

MR. WAGNER: David Wagner here. This is a very long subject, and we spent a long time discussing it during software dependence. And we could discuss it again, to bring it back to John Kelsey's talk here, there are many states and there are places that want to do various kinds of testing of their equipment, including parallel testing, including observational testing, including other kinds of testing. So from the point of view of the work that John Kelsey's doing, I view this is as saying, if it is true that many places

want to do this kind of testing, then it's important that the equipment be able to support that kind of testing.

Now, maybe we could have a discussion about the general security issues in general. I don't know if you want to do that now or you want to do that some other time.

MR. GAYLE: John Gayle, Secretary of State. Well, I guess I'm just wondering because of the cost of the testing, the certification of vendors who are going to pass that on to all of my counties and every other county in every other state, we re building in so many redundancies here? We try to create zero-error perfection, which we've never had in 200 years of our democracy. Is this kind of a new standard we're setting here with these guidelines, zero error, and we're going to have everything tested to the point with so many redundancies and audits that nobody maybe can afford it to buy the equipment, but it's going to be a perfect election?

MR. CHAIRMAN: John, could you -- this is Bill Jeffrey. For clarification, when you say things are

expensive, like parallel testing is expensive, can you say where that expense is? Is it in the up-front hardware costs which would impact the states, or is it in the actual implementation of the test, which is a procedure that may or may not be done by the states? I mean, where is that cost captured?

MR. KELSEY: Well, the cost that I know of, first of all it imposes restrictions on the design because in order to be able to do this thing where you cordon off the voting machine on election day ideally, you've constrained your design because now the machine can't be talking to all the other machines, they can't all be on a network or something. You also impose a lot of costs which, I think if there's anybody here who's been involved in parallel testing in an actual state, it would be interesting to hear from them. But you have cost in the sense that you now have to have a testing team go out and do the parallel testing on election day.

MR. CHAIRMAN: That's my question, is if a state -- since we're not mandating again procedures, if a state chose not to implement parallel testing, what is the cost penalty because they had to buy equipment from the

vendors?

MR. KELSEY: I think the only cost there is it constrains the design. So the vendors will have fewer choices when they're designing the next generation of voting machines. I don't know how to put a dollar cost on that at all. I have no idea at all.

MR. MILLER: This is Paul Miller with the Secretary of State in Washington. First of all, a comment that I have done some parallel testing in the state of Washington. Second, I am concerned somewhat about the constraint on the design. I know of a couple, well particularly the (indiscernible) does network their devices within the polling place and is able to use a number code as the ballot token instead of having a device, a donkel (phonetic spelling) or a switch or whatever, a card, excuse me, to go around and use the device. At this point I think we should take a careful look at that issue to see whether or not the benefits of separating machines so that they can't be networked -- would this also include (indiscernible) machines daisy-chain their power cords as well? Would you be including that sort of a design in this factoring as well?

MR. KELSEY: I don’t believe so unless there’s communication possible over that line. At some point you’re going to start worrying about what you call like subliminal channels where one machine can kind of subtly tell something to the other machine. But I don’t think that’s a big issue that we’re considering right now.

MR. MILLER: And in the heart system where it’s a closed loop, and in order to operate the individuals machines, if you’re going to do parallel monitoring you still have to have a loop with a controller device that’s connected to it. And if you randomly select -- I’m not sure how the equipment -- I’m not sure how within that closed loop it would be able to communicate that this is a test versus this is actually in production.

MR. KELSEY: I suspect that if you were trying to do this, and this is more a guess because I certainly haven’t tried to run something like this, I suspect what you would do is test the entire loop. So you can imagine the testing team bringing out additional, a second set of machines and controller and then just test one of the ones that were already there. That’s the way

I suspect you would do it, but I'm talking outside my area of expertise.

MR. MILLER: Okay. No, I understand. I just want to take a careful look at that before we write something specific that might put a constraint there that's not necessary.

MR. KELSEY: Right.

MR. RIVEST: This is Ron Rivest. I'd like to follow up with what Paul is saying. I think this is a place where (indiscernible) the election officials is really important. And this is a procedure which is optional certainly by the states. It's expensive when it's done to do parallel testing when you've got software independence the motivation for that is perhaps decreased and then (indiscernible) validation may decrease that as well. So I think this is language that could be written in there if the states felt that was important to them. But I think as security devices go, it's certainly marginal compared to, say, having the software independence techniques that we have already.

MR. CHAIRMAN: Yes.

MR. MILLER: Thank you.

MR. WAGNER: David Wagner. I would second that. To mention how is parallel testing -- my understanding of how parallel testing is done in California, for instance, is that you set up a mock precinct. So for those that use precinct-based networks, I don't think that has to be a barrier to parallel testing. But to get to the broader point here, I certainly would be very reluctant to suggest requirements that would constrain the design of these machines in a way that, for instance, prohibits a precinct-based network just on the basis of parallel testing. So I think we should be careful here before drafting any requirements that constrain the design of the machines. And I think particular parallel testing, as John identified is a tricky one, and the TGDC should provide input to the STS on this particular issue about what, if anything, deserves to be in the standard.

MR. CHAIRMAN: This is Bill Jeffrey. Just trying to get the sense of the discussion. Given the fact that the software independent verification covers the vast majority of what we're talking about, and given some of the down sides, is there a sense from the TGDC that --

is there value to even continuing to try to draft and discuss the parallel testing options? Or is that something that we should recommend to the STS Subcommittee that they kind of move on to other issues?

UNIDENTIFIED INDIVIDUAL: As I said, I think the election officials need to give their input here, but from a security viewpoint I think that having a back-off on this would be a little be a little bit appropriate where the requirement could say something as simple as, the manufacturer shall describe in its view what a parallel testing procedure might look at, what's possible on the machine, and how it could be exercised.

MR. GAYLE: This is John Gayle, Secretary of State. Well, it would seem to me that if it was presented as suggestions as opposed to requirements, it could be helpful to -- obviously there are many sizes of different counties and election centers. So some can afford to spend more money to do more things than others. And if these are suggested ideas, I think they'll be received favorably. But to try to say one system fits all isn't going to work, particularly if we can feel some sense of reliance upon the software

independence goal.

MR. KELSEY: I guess the question that would be interesting to ask everybody is, are there other mechanisms or procedures that we know of, that anybody knows of, that address this issue of the voting machines misbehaving in some fairly subtle way that's hard to detect, that's not detected in the paper record versus electronic record. And I think one obvious thing is just to note complaints, but I you couldn't put that in an equipment standard at all. And also, you guys all know a lot more about that than we do.

MS. PURCELL: Helen Purcell. I think probably most of the jurisdictions have observers that go around during the day and they certainly discover any kind of error that might possibly affect the voting that day. And I don't see that that's going to be a problem.

MR. KELSEY: Okay. So if voters complain, people at the polling place at the time would know that and would write it down. And then I guess the question is, how is that addressed later? That's a harder problem.

MS. PURCELL: Well, you not only have the voter complaining to the people at the polling place, but you

also have observers, at least what we do of our political parties and so forth who are observing elections. So they're going to get that information to you and it's certainly going to be taken into consideration. In my jurisdiction we have hotlines that the polling places and the troubleshooters are in touch with us all day long, so we know of anything that occurs that day and can solve the problem then.

MR. KELSEY: Okay. That's helpful.

MR. MILLER: Yes. This is Paul Miller. I would concur with what Helen just said, that that is the way counties manage their system, is using troubleshooters and hotlines from the polling places. I guess one thing I think you're trying to get at, and I don't know how to get at it yet is the distinction between what is in fact a hardware user interface issue and what is in fact malicious. And let me offer one example. There's a lot of reports of people saying they touched one vote, one candidate and they got another. And I know most counties, or the counties I'm familiar with, if they get a report from the polling place they simply treat that as the machine was not calibrated correctly. If they

get that complaint they should shut down the machine and bring out a troubleshooter who either replaces the machine or recalibrates it. They usually bring in a new machine. And they don't distinguish between if it really was miscalibrated or whether it was simply that the user didn't use the screen correctly. And I suspect frankly that if it's miscalibrated first voter in the morning, it was miscalibrated. If it was a voter in the middle of the day complaining, I suspect that it's usually the voter. And I don't know how to get at the question of distinguishing between what was a genuine hardware failure and what was a user error or malicious code.

MR. CHAIRMAN: The question to the STS Subcommittee, do you have sufficient guidance from this discussion as to how to move forward on -- you know, it's really no formal requirements on the parallel testing of potentially suggestions or guidance as to the vendors.

UNIDENTIFIED INDIVIDUAL: I know you have discussion at this point, but I think further input from election officials as to the desirability of support for

parallel testing would be helpful. As I said, I think it's got some marginality to it, and if there's demand for requirements that the machine support that for many states that would be good to know. If there's not much demand for it, we can back off on requiring any kind of hard constraints on the design to support that.

MR. GAYLE: Dr. Jeffrey, John Gayle, Secretary of State. Well, since the EAC hopefully will be issuing their Election Management Guidelines in the fall and we don't have them readily available to know whether these issues are going to be addressed, certainly I think this should either be postponed, taken off the table, or delayed indefinitely until we have the ability for the EAC to interface their guidelines with some of these issues. Because it seems to me that's more an election administration issue as Ms. Purcell and Mr. Miller have addressed.

MR. KELSEY: Okay. That's all I had.

MR. CHAIRMAN: Let me just make sure that we're consistent. Given the discussion that we've just had, essentially we would not anticipate a requirement at this point on the parallel testing, again subject to any

additional input that the STS can get from state election officials and any additional guidance from the EAC that may be coming down the pike. So with the exception of the issues on the parallel testing, do I hear a motion to adopt the rest of the preliminary draft Security and Transparency Sections that were consistent with the discussion? Is there a motion to essentially concur with the direction that they're headed, subtracting out the parallel testing part? Is there anyone who doesn't want to second that?

UNIDENTIFIED INDIVIDUAL: Yes, I move to second it.

MR. CHAIRMAN: Okay. So let me be clear again as to what we just did. What we want is the TGDC to formally concur with the direction that the Security and Transparency Subcommittee has just presented. The one change is the subtraction of the parallel testing as a formal requirement.

UNIDENTIFIED INDIVIDUAL: I just wanted to make sure I understood. I think from what we were discussing, my understanding was Ron's suggestion was that we might include documentational requirements that say that the vendor shall document how, if parallel

testing is to be done how it should be done, but that we wouldn't impose hardware requirements. Am I getting what you're saying correctly?

UNIDENTIFIED INDIVIDUAL: Yes, I would think that would be reasonable. I think that's consistent with what you said.

MR. CHAIRMAN: Yes.

UNIDENTIFIED INDIVIDUAL: No hardware requirements there, but if the machine does have parallel testing capability it should be documented.

MR. CHAIRMAN: So a formal resolution -- I will propose a formal resolution, and I apologize for probably not getting the English quite right, that we accept the direction that they've given with the change that there be no hardware requirements on the parallel testing, but if a vendor's machine has such that they should document how a state could use that for parallel testing. Is there a second to that motion? Okay. There's motion and it's been seconded. Any discussion or comments on that? Any objections to unanimous consent? Hearing no objections to unanimous consent, we've got that. And the parliamentarian wants -- so

with that I thank the Security and Transparency Subcommittee for getting us just about back on schedule and for teaching us about cryptography.

Are there any other questions or comments before we break for lunch? Okay. If not, let's get back on schedule such that we meet back here at 1:30. And again, thank you very much for the TGDC members and the EAC members. We have a room reserved right next door, dining rooms A and B for lunch.

(Lunch recess.)

UNIDENTIFIED INDIVIDUAL: -- to see who's on the phone connection. Do we have any members that are joining us for this afternoon?

MS. TURNER-BOWIE: Sharon Turner-Bowie.

UNIDENTIFIED INDIVIDUAL: Thank you, Sharon. Anyone else?

MR. CHAIRMAN: Okay. An official NIST time, it's now 1:30 so if everyone could take their seats.

UNIDENTIFIED INDIVIDUAL: Just one point of administrative matters. The signer is over on my right. If people want to make use of this, please move over to that side of the room. Thank you.

MR. CHAIRMAN: Okay. Well, good afternoon and welcome back to the Eighth meeting of the TGDC. I'll officially call this meeting back to order, and I will ask our new parliamentarian, Thelma Allen, to please call roll.

MS. ALLEN: Thank you, sir. Williams? Williams? Williams not responding. Berger? Berger? Berger not responding. Wagner?

MR. WAGNER: Here.

MS. ALLEN: Wagner is present. Paul Miller? Paul Miller? Paul Miller is not responding. Gayle?

MR. GAYLE: Present.

MS. ALLEN: Gayle is present. Mason?

MS. MASON: Present.

MS. ALLEN: Mason is here. Gannon?

MR. GANNON: Here.

MS. ALLEN: Gannon is here. Piece?

MR. PIERCE: Here.

MS. ALLEN: Pierce is here. Alice Miller? Alice Miller? Alice Miller is not responding. Purcell? Purcell? Purcell is not responding. Quisenberry?

MS. QUISENBERRY: Here.

MS. ALLEN: Quisenberry is present. Rivest?

MR. RIVEST: Here.

MS. ALLEN: Rivest is present. Schutzer?

MR. SCHUTZER: Here.

MS. ALLEN: Schutzer is present. Turner-Bowie?

MS. TURNER-BOWIE: Here (via teleconference).

MS. ALLEN: Turner-Bowie is present. Jeffrey?

MR. CHAIRMAN: Here.

MS. ALLEN: Jeffrey is present. We have ten members in attendance.

MR. CHAIRMAN: Thank you very much. And, by the way, that's also sufficient for a quorum. At this point I think it's Dr. Allen Goldfine and David Flater. You guys are up next, and to present the Core Requirements and Testing Subcommittee preliminary report.

MR. GOLDFINE: Thank you Dr. Jeffrey. It's Goldfine.

MR. CHAIRMAN: And you can say Bill.

MR. GOLDFINE: Okay. Great. We're all even then. This is the Core Requirements and Testing report. I'm going to do part -- well, let me get to the next slide. There are four basic topics we're going to be

discussing: electromagnetic compatibility requirements, quality assurance configuration management requirements, review of the CRT changes from the previous draft of several months ago, benchmarks. I'm doing the first half, Dave Flater is going to be doing the second half.

Most of what I'm going to be doing is more of a status report than anything else, talking about where we've been, what our overall goals are, how close we are to accomplishing those goals, what are the differences between now and this past December, and so on. We are leading up to one unresolved issue that I am going to toss over to the TGDC for resolution. And I've been told by my management to stand up here at the podium until a resolution is agreed to, or that we perceive a consensus, or something like that.

Okay. First of all, the topic that we now call electromagnetic compatibility requirements, basically revision of Sections 4.1.2.4 to 4.1.2.12 of Volume 1 of the 2005 VVSG. Also this would revise part of Section 6 in 2005, namely telecommunications, although from the point of view that we're looking at this, it's pretty much new as far as telecommunications. There really

weren't any telecommunications requirements in this area. And as part of the process there would also be some changes in testing descriptions, test protocols, and so on, a revision of Section 4.8 of the 2005.

Basically what we're talking about here is, again what used to be called electrical requirements, pretty much the ability or the resistance of voting equipment, electronic voting equipment, to be resistant to or resilient in the face of disturbances, interferences, power surges, that sort of a thing. It's very highly technical. We've talked about it at several CRT meetings. We've had discussions outside that on e-mail threads, and so on. And it's pretty much well on its way to being finished. We've divided the area into three sub-areas: conducted disturbances, basically emanating our of wires and tables and so on; radiated disturbances, you know, electromagnetic signals through the air; and the third area as I said before, telecommunications disturbances.

The conducted disturbances section is complete, or at least it's complete as of yesterday. Probably by Monday I'll be posting the latest draft set of

requirements on the web. Radiated disturbances is still being worked upon. As I indicated last time, we've enlisted the experts in this particular area from NIST Boulder. There were some delays in that, but now they're working hard to define appropriate revisions to the existing requirements. And the telecommunications disturbances, which are partly visible on the current draft, still some to be completed. We anticipate that everything should be finished in the sense of having a complete set of requirements for complete examination and integration and development of informative text and so on probably no later than early to mid-April.

But everything seems to be very straightforward here. We haven't perceived any major disagreements or lack of consensus within CRT. I encourage any of you who are interested in this subject, if you haven't already to take a look at the document. Well, the document that's in the hand-out and also whatever the revisions are that we continue to place on the web.

The other area I'm going to talk about is that of quality assurance and configuration management requirements. Our work in this is a response to first

of all TGDC Resolution 30-05, which sort of mandated that the sections in 2005 that dealt with quality assurance and configuration management be reconsidered, rethought, in an effort to provide additional, stronger if possible, tools to help ensure reliability of voting equipment. This was reaffirmed and extended at the December 2006 plannary where the TGDC did reach a consensus that yes, ISO-9000, 9001, that family of standards should provide the framework -- I'm trying to quote as best I can from the actual transcript -- should provide the framework for VVSG 2007 requirements. Of course, I guess we're not supposed to use 2007 anymore, but wherever I have 2007 in this presentation, make a global change to whatever is the current, politically correct word. And these revisions -- well, in this case it's more than a revision, it's a rewrite from scratch of the existing sections -- would be a replacement for Sections 8 and 9 of Volume 1 and Section 7 of Volume 2 of 2005.

Now, the draft VVSG 2007 requirements, I guess the word draft should be in there, do require that a vendor's quality assurance procedures be in conformance

with ISO-9000, 9001. And of course, in this area the devil is in the details. Saying conformance doesn't mean a whole lot. It really comes down to the particular procedures, the particular detailed requirements that are specified, first of all in the VVSG, and then how those requirements are adopted, rephrased, implemented, and so on, by the vendor. So what we have now in the draft requirements of the –

(**END OF AUDIOTAPE 2, SIDE B**)

* * * * *

(**START OF AUDIOTAPE 3, SIDE A**)

UNIDENTIFIED INDIVIDUAL: So the argument is completely circular. We cannot determine a benchmark this way. At some point we need to know really what benchmark is required.

MR. CHAIRMAN: Sorry. This is Bill. Could I ask for clarification, because we may be asking -- I mean, we're asking hard questions to people who give us numbers, like what's average volume and things. But is there any reason to believe that an error rate or that the number of errors would be greater on a smaller volume than a larger volume? And the reason I'm asking

that is if --my intuition would be that the bigger the volume, the more errors we'd likely have that you really don't want to specify sort of what's a typical but you'd like to look at what sort of an extreme. I mean, what's the 95% volume rate on which there probably is data?

UNIDENTIFIED INDIVIDUAL: Well, the idea was to derive a rate. And to derive the rate, the way the question was formed was, with regards to a typical election the thought was in a typical election where, so we believed, there'd be a way to find out what the volume was. And there would also be a way to come up with a figure for how many errors could have been tolerated before we ended up with an unacceptable result. From that you divide the errors by the volume, and you have a rate. But in fact raising the question in this way may have caused more problems than it solved.

Now, to continue with the feedback, Paul Miller on the last CRT telecon was essentially speaking on behalf of NASET and saying, what we would really like is to assign different weights to different kinds of failures, so that these kinds of failures that might possibly be

resolved in the loss of votes would be a write out. But other types of failures that we might be able to rectify on election day or by replacing a machine and recovering the votes later, we might be able to tolerate those. So if we can define these different categories of failures in an objectively determinable way -- that's what the test lab needs -- then we can assign different weights to them and possible have a more complex benchmark, but a benchmark which satisfies the needs.

Now, in fact 1990 voting system standards tried to do exactly this. Appendix 2 of the 1990 VSS, the voting system failure definition of scoring criteria defined the idea of a relevant failure versus an irrelevant failure, and also assigned different weights typically. Any old failure from which you could recover and continue with, I suppose, paper jams being in that category, would count as .2, whereas something that could potentially end up locking up votes so that you couldn't get them out of the equipment got a value of 1. Now, this system was removed in its entirety from the 2002 VSS. And as of the deadline for this presentation, Paul Miller was following up to find out why this

occurred and I haven’t communicated with him since. So if Paul is on the line --

UNIDENTIFIED INDIVIDUAL: He’s on an airplane.

UNIDENTIFIED INDIVIDUAL: He’s on an airplane. Darn. Well, this is where we are. Something was done in 1990 VSS that once again is starting to sound like a good idea. We want to know why it was taken out. Rick would also know this probably, but he hasn’t made it here either. So --

MR. CHAIRMAN: David, let me ask a question, and obviously we’re not going to wait for Paul’s plane to land. From your understanding of Appendix G of the 1990, would that methodology really resolve the issues?

UNIDENTIFIED INDIVIDUAL: There are some minor, resolvable incompatibilities with the test method that’s there now. I know exactly what I would do to fix them.

MR. CHAIRMAN: Okay.

UNIDENTIFIED INDIVIDUAL: But what I can’t do is tell the election officials what benchmark they want.

MR. CHAIRMAN: Right.

UNIDENTIFIED INDIVIDUAL: Now, based on the old standard, if we just assume that the old standard was

correct in every way, we could stick as close as possible to those old numbers. But this sort of gets us back into the old -- I mean, it was defined in terms of meantime between failure, which we already have the resolution to move away from. So we wanted to rebase this in terms of volume instead of time. So really the election officials do need to weigh in.

MR. WAGNER: David Wagner. If you go back to the slide with your summary of the NASET letter, I believe it was, I wonder if there's something, some partial things we could learn from that letter. I thought that was a very informative letter. One of them is this distinction between unrecoverable and recoverable failures. And I think that's an important distinction. There's a big difference between a machine crashes and that corrupts or deletes all the votes and now it's impossible to recover those votes, versus a machine crashes and, I still have all the previous votes and maybe I can't accept any new voters, but I haven't lost any prior votes. And I'm not sure whether I read -- you know, I'm just trying to read this on the fly, but I didn't see that distinction made in the current

definition of failure rate in the draft before us. So I wonder if just starting by making that distinction might allow us to make some progress. For instance, one possible direction one could propose would be a failure rate where failures lead to loss of votes. The acceptable rate for that might be 0 as listed up here. The rate of failures which are recoverable or don't lead to a loss of votes just lead to the loss of ability to service new voters. That might be some non-zero rate that's acceptable that could be specified.

I also thought maybe I'd just comment a little bit on your statement that because we don't know what practices and procedures will be used in the field, it's impossible circular problem. I'm not sure that needs to be such a roadblock. I think that first of all we can identify there are some failures that are unrecoverable, no matter what practices and procedures we use. Those, it's very clear cut what to do. Then I think from there one could look at what the documented practices and procedures in the use manual provided by the vendor are. And if use manual that's provided by the vendor supplies practices and procedures, tells you to use this system

in such a way it would lead you to recover, then I think it's fair to classify that as a recoverable failure. And it's true, maybe there's some gray area for failures where the manual doesn't say what to do and we don't know what practices and procedures would be used in the field. And I don't know what to tell you for the gray area, but we may be able to make some progress there.

UNIDENTIFIED INDIVIDUAL: If I could respond briefly. One additional complication with regards to recoverable failures I didn't go into. And this was among the questions that we asked. We didn't want to ask if we wanted to have different benchmarks for different types of equipment, because if you've got one optical scanner counting all the votes, you probably want that to be more reliable than one of the hundred DREs that you have, simply because the consequences are worse. That's all.

UNIDENTIFIED INDIVIDUAL: I think we're on to something here. If you have the right procedures and policies, depending on whether they have back-up equipment and so forth, you might even be able to work around non-recoverable failures. I agree that we have

to need some surrogate, because you're not going to be able to have all the policies and procedures that each municipality might have. But you won't necessarily want to have the one that the vendor provides also. So I would suggest though that maybe if you could some prototypical or average kind of policy and procedure, people could agree what is approximately close or representative. And you try to do the task around that, then you might be able to get to it. In other words, if people tend to have only one optical scanner, then you test it with only one. People tend to have only one DRE in a particular (indiscernible) test only one, you know, if you follow what I'm saying. How do you start and recover in a normal kind of procedure. It won't be the same exactly for everyone, but there might some prototypical kinds of policies and procedures you might want to discuss how this would work with that test. And it would be some amount of them of what the vendor recommends and what the practical people in the field who have modified that to --

UNIDENTIFIED INDIVIDUAL: But then a given optical scanner might be deployed in a precinct count

configuration or a central count configuration, the number you have might change. My intuition would be that this could be like asking about typical volume, that there is no typical, would be the answer.

MR. CHAIRMAN: This is Bill. Do you have any concrete recommendation for TGDC --

UNIDENTIFIED INDIVIDUAL: Well, actually I have to get through accuracy, too. And then I will not make a recommendation, but I will say something that will hopefully wrap things up by July or June.

MR. GAYLE: Dr. Jeffrey, I'd like to just ask a couple questions of David. And I guess I'm just thinking in terms of election officials dealing with whatever they have. That's the reality. And most election officials keep spare parts, spare ink, spare cartridges, things that they can deal with and replace. So if you're running out of ink or if an optical scanner is getting too much dust on the light so that it's not reading properly, they can step in and clean that off and recover the equipment to continue to count. So there are so many things that, in some ways you might call a failure, but it's a very recoverable issue on a

lot of levels, as long as there's some training and they have spare parts. And that's one whole level.

Then you have the level of well, maybe there's a bigger problem and you have to have a technician come in, but the machine is not going to be put back into storage because the technician's available and can address the issue. So I don't know if, when you talk about failures, outside of that context where every precinct usually will have two pieces of equipment anyway. And so even if one is irrevocably down, irretrievably down, it doesn't mean the election can't go on. It doesn't end the election because the other piece of equipment can be used, or you can do ballot-on-demand and print a paper ballot.

So when we talk about failure, are we talking about failure of the election, or are we talking about just an irrecoverable failure of the piece of equipment, no matter whether there's back-up equipment to step in its place or not? So I have trouble with this issue of what kind of failure are we measuring.

UNIDENTIFIED INDIVIDUAL: Perhaps I should have started with the definition of failure. Anticipated

events like running out of paper, running out of ink, having to sweep the dust off the sensor, these don't even register on the radar. These are not failures. These are expected maintenance chores. An unexpected thing like a paper jam is probably the least severe thing that qualifies as a failure, and it gets worse from there.

Now, the issue about recoverability, even in the old standards there was a requirement to the effect of not having a single point of failure and things like that. An argument could be made, or in fact we could make it so by adding unambiguous requirements that the notion that any equipment should fail in a way that makes any vote completely unrecoverable is already a nonconformity, regardless of the reliability benchmark. And that would take that out of the equation. And then we would simply be focusing on everything in the middle. If unrecoverable votes are completely banned, replacing the ink is completely irrelevant, then everything in the middle is a failure, and those are what we cannot for the sake of the reliability benchmark.

UNIDENTIFIED INDIVIDUAL: Why are you saying

unrecoverable?

UNIDENTIFIED INDIVIDUAL: Well it's completely banned. How about a scenario where feeding in optical ballots and one of them gets chewed up?

UNIDENTIFIED INDIVIDUAL: Well, I'm just repeating what NASET told me. No failures that lead to unrecoverable votes are acceptable. That's one thing I have in writing.

(Speakers not using microphone.)

UNIDENTIFIED INDIVIDUAL: This is one of those cases where --

(Speakers are not using microphone.)

UNIDENTIFIED INDIVIDUAL: Yes, what it says in the standard may be something that in the real, physical world may not be enforceable. But the consequences in the test lab are that if this happens when anyone's watching, then the equipment will not be certified.

MR. CHAIRMAN: David, since there isn't a concrete recommendation at this point, because it's (indiscernible) more work, I think that there's a general sense that dividing up the failures into the non-recoverable and recoverable, there's maybe something

in there that looks good. And I think intuitively that seems to make sense to a lot of people.

UNIDENTIFIED INDIVIDUAL: Yes.

UNIDENTIFIED INDIVIDUAL: And we needn't go as far as instituting the scoring system from 1990 if there was a reason for taking that out, which we're still waiting on. But certainly making this division is a simple enough thing to do and everyone seems to like it, so great. Can I move on to accuracy now? No objection?

Okay. I'm paraphrasing NASET on accuracy. Something that I found myself commenting on on many occasions talking about elections past and future is that the real requirement on the voting system is that it have one less error than the vote margin between first and second place. That's the real requirement.

Now, if we get beyond that, so what's the benchmark? The old standard said one in 10,000,000 ballot positions was allowed to be wrong. And as NASET discussed, this was a compromise based on testing. Of course, the cost of testing, you can't of course prove perfect accuracy in any finite-length test. And on the surface, there's no reason to change this benchmark.

But there is a need to review the test methods. As I had mentioned earlier, there was ambiguity with the metric as it was specified.

They also expressed some concern that the 1 in 10,000,000 ballot position's benchmark might be achievable for perfect test ballots, but maybe not for real ballots.

UNIDENTIFIED INDIVIDUAL: (Speaker not using microphone.)

UNIDENTIFIED INDIVIDUAL: So it depends I think somewhat upon the voting equipment first of all, so let me illustrate. If you're talking about a DRE piece of equipment, then short of it breaking down and failing, well it being compromised it's going to be accurate. If you're talking about an optical kind of a thing, then yes, you may have accuracy problems. But short of the illustration I gave where, you know, it just gets chewed up and it's not recoverable, you could design it so that you make as high an accuracy as you want and if that system is unable to, with that confidence provide you that output, then it kicks it out for a human being to look at. So, I mean, the accuracy could be somewhat

influenced by the manner in which the systems is designed and used, if you follow what I'm saying. So I think you just have to factor that in also. It is somewhat related to the procedures and selected guidelines. If you think there's problems in the accuracy not being as good as you'd like in the (indiscernible) you could actually compensate for some of that if you were to adjust it for yes, no, or maybe.

UNIDENTIFIED INDIVIDUAL: Yes. I talked a little bit about marginal marks in December. This is a tunable, it's a calibration item for optical scanners. And at that time the issues was that in fact you do want the capability for this system to reject ballots that contain marginal marks, because even though your calibration may be this way or that, as long as you have this maybe zone defined -- above here you're pretty confident that it's a yes, below here you're pretty confident it's a no, and the rest might be below the noise at some point. And that's what you want to kick out.

UNIDENTIFIED INDIVIDUAL: Right.

UNIDENTIFIED INDIVIDUAL: Certainly using that

practice will help you certainly in the precinct where the voter is standing there and can be asked to clarify. I don't know what you'd do in a central count case or in absentee ballots. You have to arbitrate somehow.

UNIDENTIFIED INDIVIDUAL: You have to arbitrate with a human to look at and determine what they voted.

UNIDENTIFIED INDIVIDUAL: So with regards to --

MS. PURCELL: If I could -- Helen Purcell. The accuracy also depends upon the end user of the product, and that's where you get into what you were talking about about your absentee or early ballots, because you're not at the polling place so you can't determine if there was something wrong with that ballot. For instance, if somebody instead of marking an arrow circles something but it doesn't go through the read path at all, the machine doesn't pick it up because the machine doesn't know it's there. So that's something -- if they do everything on the ballot that way, it comes out as a blank ballot, so you look at that. But if by some chance they didn't do everything on the ballot that way, there is some errant mark in there -- so that accuracy is going to depend on what that user does with

the ballot.

UNIDENTIFIED INDIVIDUAL: Yes. In this case, questions of deriving voter intent from ballots where they completely ignored the instructions is sort of out of scope. This discussion is about ballots that conform to the requirements. This is a properly marked ballot, and we want to know how often does the machine make an error on a properly marked ballot.

MS. PURCELL: So you're not looking at an error made by the voter?

UNIDENTIFIED INDIVIDUAL: No.

MS. PURCELL: But merely an error made by the machine --

UNIDENTIFIED INDIVIDUAL: Yes.

MS. PURCELL: -- in reading what the voter put on there?

UNIDENTIFIED INDIVIDUAL: Yes. So that rate should be low. So continuing with the discussion of the benchmark, based on the feedback received the vote margin criteria and yes, in real life, we would always like to have the number of errors be less than the vote margin. But since you might get a vote margin of one or

even zero, that's a perfectly possible if unlikely scenario. That doesn't help us to set a benchmark other than zero. And as I said before, zero is a possibility.

Now, there was some support given for the 1 in 10,000,000 ballot positions number, but then if we move forward from that as a starting point, the clarification that I discussed in the draft in December was moving from ballot positions as the basis for the metric to something called report total error rate. And this has to do with the fact that what you're getting out of the system is a report. And actually if you go back to the 1990 spec, there was a sort of equivocation that started even way back then between ballot positions and votes. And what you're seeing in the reports is not ballot positions, it's votes. And the benchmark was there in terms of ballot positions, but then the evaluation about what you do when you see errors was written in terms of votes and looking at the reports. So there's been a bit of confusion all along on that. And the draft currently addresses that using report total error rate as looking strictly at votes instead of ballot positions. Having made that alteration, it's worth revisiting the 1 in

10,000,000 number to ask, is it still appropriate, because it will have some impact.

Now, perhaps more worrisome was the comment about the achievability of the benchmark for “real ballots”. The implication was that for some category of real systems, only perfect test ballots are going to be able to accomplish 1 in 10,000,000, and that if you took a stack of real ballots for a real election, you won’t make that benchmark. If that’s the case, we have already discussed using volume testing with real people and real ballots. There’s been a lot of support for doing that as part of the test campaign. If that’s what we’re going to do, then the benchmark should be something that’s achievable in that context unless you want to disqualify everyone. We don’t have that figure, so once again we’re sort of asking, what error rates are being achieved in practice. And I believe there are some comments.

MR. CHAIRMAN: John and Whitney.

MR. GAYLE: Well, I’m just sitting here thinking about equipment that maybe has a maximum use in a precinct of maybe 1,5000 voters and maybe will be used

the maximum six times a year, and so maybe you're getting 10,000 real votes, real ballots cast on that equipment. If you have a ten-year lifetime, you're talking 100,000. So this 1 in 10,000,000 just doesn't even begin to make sense to me as an amateur in this business when in terms of the reality of the equipment it's dramatically less in terms of the expected use, ordinary use. And obviously you'd need a multiple of that of somewhat, but I don't see what the options are. If 1 in 10,000,000 makes sense, of course it doesn't make sense to me, but maybe it makes sense in terms of science. But it seems like you're testing equipment way too high a degree of perfection that's going to drive up costs and going to drive up the inherent ability of election officials to buy new equipment if we test this to perfection, which is what this sounds like to me, as opposed to the reality of how the equipment is going to be used. So I'll defer to Whitney.

MS. QUISENBERRY: Before you respond to that, my question is actually related. And I'll eliminate the parts that overlap. One of the questions I have, between sort of machine testing with perfect ballots and

volume testing, are we thinking about having both of those? Because one of the things I've seen done in other contexts is that you use a fairly stringent, perfect-world test which is cheaper to run because it's a sort of machine test, before you go into something that involves lots of people, which therefore becomes a more expensive test to run. So you can use that a gating for, you know, is it mechanically sound enough to go on with. And so at that point you tend to get requirements that are more stringent than real world, because then you're also going to go through a kind of real-world environment.

UNIDENTIFIED INDIVIDUAL: We're talking about a couple of different things. So maybe we can dissect it and (indiscernible) testing. I mean, the one thing for volume testing is to just find out if the system will hold up under a lot of documents feeding through it or a lot of votes, and what's going to happen to it as you start beating the system with a lot of volume. The other thing you did when you talk about with real people is you're also introducing -- taking the case of the optical scan -- the fact that people may not do perfect

circles and so forth. So I would suggest that maybe we could separate those two types of tests, you know. One is we can feed through lots of perfect ballots to just see how the equipment holds up under that kind of volume from a reliability sense. And another is we get some prototypical samples of what some real ballots are and see how well the equipment operates under variability and how real people do the ballots. But don't submit that to the volume test. We're really trying to just see the kind of thing you're talking about, the circle and not full circles and so forth, rather than the filling it in, and see how well it works there and come to some conclusions.

UNIDENTIFIED INDIVIDUAL: Stress testing is in fact a separate item. And you don't care necessarily -- I mean, you're not going to achieve the full volume desired when you've got human beings in the loop. So stress testing can be performed validly without people in the loop, and that's written into the language now. The language is rather general about the series of different types of testing that are done, but that is a separate type of testing from the -- it's really only

called volume testing because of the California volume reliability testing protocol, which is where this idea came from. In reality if we had to pick a better name for it, it would be something like real people testing, realistic election scenario testing.

MR. CHAIRMAN: David?

UNIDENTIFIED INDIVIDUAL: I want to second what Secretary Gayle said. And my impression, the 1 in 10,000,000, that 10,000,000 number is artificial and doesn't seem to have much bearing to the real world performance of these systems. And I don't think there's any reason we should feel constrained to stick with that number. And I think what you're proposing here, to base it on error rate when it is marked with real ballots under conditions where people are filling them out, that we hope will be representative of how it will actually be used in the field, I think it's a very positive direction. And yes, I agree this is the stumbling block, is we don't have that figure. But if that figure of what's achievable turns out to be a much different number from 1 in 10,000,000, even if it's one or two orders of magnitude different, I think we should just

accept that and that will be a very positive direction. So I know that's not very helpful other than to say that I think this is a great direction you're heading. And I don't know how to help you go there further.

UNIDENTIFIED INDIVIDUAL: At one time I recall we had a discussion about (indiscernible) the optical stuff some kind of a halfway solution. In other words, let's say I had a machine which was a DRE-kind of machine but it's not recording any votes, it's not storing anything electronically. I'm just using that to drive the printer, if you follow what I'm saying. So some of you come up there and see the ballot on the screen, make their choices, and that drives the printer which produces an absolutely valid, perfect kind of a ballot every time. That might be a device you might want to think about in terms of --

UNIDENTIFIED INDIVIDUAL: That's a class of devices called EBMs, Electronically Assisted Ballot Markers.

(Speaker not using microphone.)

UNIDENTIFIED INDIVIDUAL: But accuracy and everything else, you'd want to know that, you want to convey that to people in the testing.

UNIDENTIFIED INDIVIDUAL: Well, in fact what we want to do is set a performance benchmark and not pick winners among the designs. If we set a benchmark and some particular design can't meet it well that's too bad, but we want to set a benchmark that will be design agnostic.

UNIDENTIFIED INDIVIDUAL: I guess I wanted to support the idea that 1 in 10,000,000 seems awfully high to me. I mean, voters are the most notoriously inaccurate part of the system here anyway, and getting a voter to be accurate with a better than 1% error rate is probably impossible. Some of the studies seem like they're more like 3 to 5 or more percent is common. So if you have a system which is -- you know, 1% is inaccurate as a voter is, I think you're doing, you're (indiscernible) so 1 in 10,000 would certainly probably be fine. And if we (indiscernible) 1 in 100,000 as a target number I think we'd be in good shape.

UNIDENTIFIED INDIVIDUAL: Alan picks the numbers.

MR. CHAIRMAN: This is Bill Jeffrey. Secretary Gayle has echoed what David -- further echoed by Ron. I'll continue the reverberation. For the volume

testing, obviously cost to what we really need is sort of paramount. You want to get this as efficiently as possible but it should be realistic. And I think John's back-of-the-envelope calculation gave reasonable numbers. I think one could go and actually get that kind of data to look at what are reasonable volumes that exist out there. And you've got all the statistical powerhouse that you can with ITL to figure out sort a confidence (indiscernible) the testing to come out with some reasonable level of assurance that again 1 in 10,000,000 sort of doesn't pass (indiscernible) for volume testing, but something, maybe more than a few percent but less than that. I mean, it seems like there should be a way to do it. My guess is by going out and continuing to canvas people's opinions on the matter is not going to be as productive as actually doing the back-of-the-envelope calculation, coming up with what you think is reasonable, and justifying why you think that's a reasonable number, and then letting people debate the reasonableness of your assumptions. Otherwise you're going to continue to get circular arguments.

UNIDENTIFIED INDIVIDUAL: Well, not being an election administrator I don’t have a whole lot of confidence in my back-of-the-envelope estimation of the achievable error rate. The bottom line is there is --

MR. CHAIRMAN: But maybe parse the problem differently. Secretary Gayle, if you whipped out just the number of times the ballot’s going to be cast on that device or the number of times an optical scanner is likely to read that, that gives you some upper limit essentially for that. And again, those were numbers just from experience, but my guess is that there’s some compilation that you can get from some of these groups as to how many times a typical machine does see a ballot. I mean, you could probably parse the question into something answerable. I have an increase in specificity of what the question is, and that can then drive some of your assumptions.

UNIDENTIFIED INDIVIDUAL: Okay.

MR. GAYLE: Dr. Jeffrey, I think I’m going to have to clarify what I said, because I was thinking of precinct scanners and precinct counting. When you get into central scanning, the M650s, you’re talking about a

much faster processing and a much higher number of ballots that do get processed per piece of equipment. So I guess we do need to clarify, at least (indiscernible) optical scanning, are we talking about the really big ones or are we talking about the precinct ones?

UNIDENTIFIED INDIVIDUAL: I look forward to getting back in touch with Paul Miller who also should have perhaps some of these back-of-the-envelope figures. I would encourage everyone with an interest in this to participate in the next CRT teleconference. And let's reach closure on this to the best of our abilities. The bottom line is right now we don't have the number and we need all relevant input now, if not yesterday or last month. What's in the draft now? There's a number in the draft now, but if you want me to do a back-of-an-envelope justification for it, it might be doable. But it will be far better if we have absolutely everyone on board here. Everyone needs to know where the number came from. We obviously have a problem with the 10,000,000, okay? That's been in there since 2002 and people are still being surprised by it. So we don't

want to do that again.

MR. CHAIRMAN: And to clarify my suggestion, just reaching out and asking people for numbers, I'm not sure what you're really going to end up being able to do with that. You know, the number is 42.

UNIDENTIFIED INDIVIDUAL: I've got it that I've asked the wrong questions.

MR. CHAIRMAN: Well, what I'm suggesting is that you perhaps outline a specific methodology. You know, populate it with your numbers with the assumptions clearly delineated, and allow people to then take each assumption and argue what the range of those numbers might be that can get you there, as opposed to just the end state, being a number 1 in 10,000,000 or zero.

UNIDENTIFIED INDIVIDUAL: Here's a thought. There are equipment out there in the field that people are using right now. So supposing I were to take some of that equipment in the field. Under even ideal circumstances and that it's not out there, it's calibrated before it's done, and so forth, and I run that machine that's actually being used under some of these tests. And I find out what numbers they actually

are achieving, and that's an ideal situation for that machine. Now, it may not be satisfactory from one vote. It might really turn out to be 1 in 10,000, but that's at least a number you can have as a benchmark. And they should be at least as good as what's out in the field today. And then of course the new equipment comes, we start finding some equipment can actually produce superior numbers, then you might want to consider at some time changing that number. But at least it's saying what municipalities are using today, the new equipment (indiscernible) would be at least as good as that under the same kind of testing conditions. And maybe that's good enough. It's not where we'd like to be, but that's what's being used. So you might think about that as (indiscernible).

MR. CHAIRMAN: David?

UNIDENTIFIED INDIVIDUAL: And a possible constructive direction that could help you commit a number would be, just to elaborate on what Dan is saying, there are a number of states that have been doing audits of their voting equipment. And one possibility could be that it may be possible to gather

data on the results of those audits. For instance, Helen sent around a great document from her county reporting on results of the audit. And there were a couple of cases in there where you could identify how frequently errors in how the scanners interpreted the ballot occurred. I know that my state of California does audits, and there are many states that do audits. So it may be possible to get some data there that could help guide you as well.

UNIDENTIFIED INDIVIDUAL: I was going to follow up on what Dr. Jeffrey said. You may not be comfortable with a number, but would you be able to construct the formula? If the formula is 1 in “n”, and that “n” is derived by saying a calculation like what Secretary Gayle just did, then you’re really arguing about what the input to that formula is, not the formula itself. And it might be interesting to do both things - just think about how you would decide that is the formula, get input on whether that formula is a good formula, and then ask what the numbers that are input to that formula ought to be, which is a second issue.

UNIDENTIFIED INDIVIDUAL: Well, in terms of the

test method, what we discussed in December was we didn't want the formula to be based on time. The suggestion was to move towards a volume-based formula.

UNIDENTIFIED INDIVIDUAL: No, no. Correct. What I'm saying is Secretary Gayle sort of gave you an off-the-cuff volume formula, which was well, this many voters, this many elections, this many years of service. And if that's the right calculation, you know, "x" times "y" times "z", then the only question is what are those numbers. And out pops your 1 in what number.

UNIDENTIFIED INDIVIDUAL: Yes. I agree. That's just another way of deriving a number.

UNIDENTIFIED INDIVIDUAL: But what I'm saying is that in terms of what question you ask to get meaningful input, you could construct that formula, show an example, and then say, and what are the numbers here, is it ten years or is it 20 years in service, is it one election a year or is it six elections a year, is it five voters per election or is it 100,000 voters per election. Those are probably two outside extremes. So that might help you narrow in on the number, because the number is really a product of a number of other numbers.

UNIDENTIFIED INDIVIDUAL: I will accept all constructive input on what the right questions are that I should be asking.

MR. GAYLE: And you have to write quality assurance documentation with (indiscernible).

(Laughter.)

UNIDENTIFIED INDIVIDUAL: I'm sorry, I missed that. Okay. Well, if I'm done with this, then that leaves me 3½ to do the other half of my presentation.

MR. CHAIRMAN: Keep going.

UNIDENTIFIED INDIVIDUAL: Okay. Well in effect, what I have to report here in review of CRT changes, I presented a whole pile of new sections in December. I don't have any big, new sections to present this time. All I have is sort of mundane maintenance to the sections that I presented in December. And most of this is probably not worth walking through in detail, given our time situation. I will point out two significant issues. One already is with regards to the benchmarking test method and the benchmarks themselves that we just discussed. Essentially what happened there is everything we talked about in December was pasted into

the draft.

The other issue was about coding conventions. There was an interaction with Dr. Wagner about block structured exception handling and the impact with regards to systems using the C language. And we essentially discussed this at length and reached a compromise in which yes, in fact there is a way to retrofit things written in the C language that satisfies block structured exception handling and good structured programming principles. So the text that's in there has been modified. It's mostly introductory text and like one or two words in the requirements themselves. But mostly it's the introductory text saying hey, yes you can in fact do this. We had a long discussion about structured programming in general.

Other than that, everyone has the document titled Review of CRT Changes. This is essentially a change log against the sections that I presented in December. All of the references to Volume 3, Section anything are off by a few chapters because a bunch of chapters were inserted after this went to print. But there's some algorithm figuring out where they're pointing. What I

would say is perhaps -- I don't mean to push the agenda around, but perhaps the best use of time would be if people care to examine this four-page document over the break at some point, if there are any questions about it afterwards I could take those questions. But otherwise we can move on to the next -- would that be acceptable to everyone?

All right. Thank you all.

MR. CHAIRMAN: Okay. So thank you very much. So the preceding presentations on the Core Requirements and Testing Subcommittee, actually my notes respond to the eight relevant TGDC resolutions. And unless there are supplemental directions or corrections above and beyond what we've already discussed, do I hear a motion to adopt their preliminary draft Core Requirements and Testing sections consistent with the discussion? In other words, that they're basically on the right path except for all of the unknown numbers that all will end up to be 42.

UNIDENTIFIED INDIVIDUAL: (Indiscernible.)

MR. CHAIRMAN: Okay. Is there a -- second is moved. Seconded is -- any objection to unanimous

consent? In hearing none, they pass and we're only 22 seconds behind on the break. So let's catch that up and I'll start again at 3:30 on the dot. Thank you very much.

(Break.)

(**END OF AUDIOTAPE 3, SIDE A**)

* * * * *

(**START OF AUDIOTAPE 3, SIDE B**)

MR. CHAIRMAN: Okay, it's just about 3:30. If everybody could take their seats, and all the people ignoring me in the back, that includes you.

UNIDENTIFIED INDIVIDUAL: One logistic item while everybody's sitting down. For those that are planning in the audience to take the shuttle back to the Metro, the last shuttle is at 5:30. So we're planning to wrap up around 5:30, hopefully a little earlier, but you probably want to leave here if you're going to get the shuttle at around 5:25.

MR. CHAIRMAN: Yes, I will wait --

(Off the record.)

MR. CHAIRMAN: Okay. At this time I'd like to ask Sharon Leskowski to present Human Factors and Privacy

Subcommittee preliminary draft report.

MS. LESKOWSKI: Okay. Thank you very much. Good afternoon. So I don't know if I'll take the full two hours here, but you never know. There's always some interesting questions that come up.

Okay, quick overview. I'm going to talk about four topics. First, some changes and issues that have come up on the HFP Section, some issues that require further analysis, then I'll give a little tutorial on how we're developing benchmarks and what the status is, and our progress thus far. And you've already got a bit of a tutorial on a different kind of benchmark, courtesy of David Flater, so I guess you're primed. And then some of the next research steps which both apply for the VVSG and beyond for the testing methodology.

There are three significant changes from the December meeting. And I refer to the requirements using the Chapter 12 numbering from the version that you have in your binders. There were lots of other little editorial things that did not change content, so I'm not going over those.

First one, in VVSG '05 we required the availability

of different choice of font size and contrast on the accessible voting station. And because when we looked at what's currently commercially available, we realized that just about all the voting stations do allow this kind of adjustment. So in aligning with one of our very first resolutions at the first meeting to strive for universal design, we said well, we ought to just require that on all the voter-editable ballot devices that are for the visual. So we moved them to Section 2 of Chapter 12, so now we've got available font sizes under the control of the voter. You'll note that one thing that we also allow is that second sentence, the system shall allow the voter to adjust font size throughout the voting session while preserving the current ballot choices. And we also moved the high contrast for electronic displays to the usability section, and again that the voter can adjust this throughout the voting session.

So our suggestion is that we should remove the requirements. We'll put a pointer in there in both forward and back, to remove if from the accessible voting station because it's redundant. All the

usability requirements pertain to all voting systems. And we also looked through all the adjustable controls of the voting station and we updated for them to be available throughout the voting station. So we've got general adjustability for all the requirements when the voter can control or adjust some aspect of the voting station. And that can be done throughout the voting session without loss of information. So for the most part, that was already in there.

There are two that are new, because as I said we revisited and looked through all the controls that were possible. One was for the synchronized audio and video. The change there is that the voter can choose either audio or visual output or both, the idea being that if you're blind you just want to hear it, you want to preserve your privacy, you want to shut the video off. If you don't need the audio you don't want to listen to it. And if you have certain (indiscernible) disabilities you might want to hear and see at the same time. But we added the switch ,the system shall the voter to switch among the three modes, throughout the voting session.

Similarly, we did the same for voter control language that the voter can select among the available languages throughout the voting session while preserving the current ballot choices. So now we've got the same parallel construction for all the controls.

Any questions?

The second issue, this was discussed in our previous meeting, we were given the safety requirement from (indiscernible) I think, and --

MR. RIVEST: Can I ask a question about the previous slide? This is Ron Rivest.

MS. LESKOWSKI: Sure.

MR. RIVEST: Sorry to interrupt, but it just (indiscernible) relevant. Voter control language, this will allow the voter to select among the available languages throughout the voting session. So it means everything on the display is going to -- if the voter requests to see English for the first part and then French later on, when they go to the review screen what happens?

MS. LESKOWSKI: Well, they'll see, if they decide all of a sudden -- well, I'll use my example from the

hardware store where I accidentally hit Spanish and couldn’t read anything and couldn’t change it back. If they decide -- and for the review screen they can have it in either language, because now we’ve said it’s controllable throughout the voting session.

MR. RIVEST: So it’s not the language that they requested earlier, it’s whatever they current language is?

MS. LESKOWSKI: Right. So maybe they start out in English and they say, well, I’m just confused now, I really need to see the Spanish version. And if that’s at the review screen time, they can do it at that point.

MS. QUISENBERRY: (Indiscernible) do this, but -- this is Whitney. One of the things that we’ve heard and observed is that someone who is a two-language speaker might start out going through the candidate races happily, voting for president, senator and so on, and then get to a complicated ballot question and want to be able to read that in their second language and to be able to at that point switch. Of course, the names are always what they are, so that doesn’t change.

MS. LESKOWSKI: Okay. So in an earlier version of

the VVSG from the last meeting, we referred to OSHA as sort of the umbrella safety requirement. And we discussed this with several NIST people who are experts in OSHA regulation and UL 60950, and they explained to us that the OSHA reference is a regulation. What we really wanted was the actual safety standard itself, which indeed is UL 60950. And that would be the correct way to refer to the safety regulation.

I said that was the third issue. The second one was general adjustability throughout.

So there are several issues that we've looked at that require some further analysis. Some are thornier than others. Let me go through them. I'm going to actually give a little tutorial on the common-industry format for usability test reports in a moment.

We've required this usability testing by vendor but what would be very useful is, this is a very general test-reporting format that we refer to, and we'd like to specialize it. For example, in our benchmark research we developed a variation of a user-satisfaction questionnaire specifically for voting. So there's no reason why we can't provide that as the questionnaire

for them to use and save them a lot of headache in trying to figure out what kind of questionnaire do we develop, our own, or are there any standards out there. So we can do things like that. We'd at least like to do that for the general-usability test that a vendor would submit, but there are several others. And I think that's the longer term. Research (indiscernible) is providing some guidance on how to specialize the (indiscernible).

And just to make that make a little more sense I thought maybe it would be interesting to give just a very quick, two-slide tutorial on what is the common-industry format for usability test reports. It very simply describes how to report, not what to test but how to report, on the usability test. The focus is not for, in forming design but to give a point in time, what is the usability of a particular system. That is what we call summative usability testing. And the original purpose of the (indiscernible) was just to have a common format so different organizations could review and compare results. We think the same logic applies for looking at vendor reports as well. It's just easier to

read them if they all kind of have the same look and feel.

This is an ISO standard. I've just cut and pasted sort of bits and pieces to give you a little bit of intuition of what's there, for example, the test objectives. And in this case when I talk about specialization of the (indiscernible) we can give very specific test objectives here because it's for a voting system. Other things that get reported is number of participants, and there's some guidance as to the minimum number of participants that should be reported. And of course you've heard about our usability metrics (indiscernible) satisfaction, which comes from other standards.

So in general -- and I bring this up because I'm going to talk about it a little bit later when I talk about usability tests, and this gives me a little framework to talk about it within -- the intended users, the actual users in the demographics, the environment, the working conditions. So for our voting benchmark tests we didn't do what's called a think aloud because we're timing. We just want the voter to vote on the

machine. Often you make a decision in your tests of whether to provide assistance or not, at least eight participants, you've got to define effectiveness (indiscernible) satisfaction, measures of effectiveness can be completion rate, number of errors, etc. And as I said, I just bring this up because I didn't want the (indiscernible) to be mysterious. It's pretty straightforward reporting.

So that brings me to this research issue of performance metrics. That is, you'll see in 12-2-11 that we've got a section for benchmarks. We don't have any numbers in that, just like Dave Flater's talk. No numbers at this point, but we've made some progress. I'm going to talk about that in a moment. But that is an open issue right now.

The next issue is based on our discussion from the earlier meeting of end-to-end accessibility evaluation. And in the VVSG glossary there are two definitions for end-to-end: the security definition, which is supporting both voter verification and election verification, and a more generic end-to-end, covering the entire elections process from election definition through the reporting

final results. So when we say end-to-end accessibility evaluation, we're referring to this more generic process. And that's what's in the glossary now. If there's an issue with it, we can certainly alter it.

So when you do accessibility testing of the components in the standard itself, a lot of them are designed guidelines requirements. And even if you do some usability testing with a particular set of voters just on the voting station, that's not necessarily sufficient to ensure the entire voting process is accessible, that it does not violate any of the -- that is, the entire end-to-end process does not violate any of the VVSG requirements such as privacy, and it doesn't break anywhere.

So our goal here is to create a place in the standard for a test method to ensure that we've looked at how the whole process fits together. So basically what we're going to try to author is a fairly simple requirement for assistance to support end-to-end process accessibility, which will then be demonstrated by and end-to-end comprehensive accessibility evaluation. That doesn't necessarily have to be with users if you have a

knowledgeable accessibility expert. One could do a walk-through of the system if they're knowledgeable about what some of the pitfalls are. But that gets us into the test method definition domain now.

But the second part of the requirement would be that the vendor shall document the process by which the system supports the end-to-end accessibility, so that the test lab could use that documentation to actually confirm that that end-to-end accessibility does work.

Any questions about that?

Okay. The next issue that I just want to put out on the table -- we're going to have lots of time to discuss this in detail tomorrow under (indiscernible) and accessibility. But I just wanted to put it on people's radar screen so that you would know there is some very early draft wording. I don't know the outcome of our discussion tomorrow, so I don't know if this is going to make sense after that or not, but it's a starting point. The accessibility of paper-based vote verification: if the accessible voting station generates a paper record or some other durable, human-readable record for the purpose of allowing voters to verify

their ballot choices, then systems should provide a mechanism that can read that record and generate an audio representation of its content. The use of this mechanism should be accessible to voters with dexterity disabilities. We can also discuss tomorrow shalls versus shoulds, but I just wanted to get the wording on your radar screen.

Any questions?

In our travels through editing the HFP section, we note the VVSG ’05 dexterity requirement that if the voting station supports ballot submission for non-disabled voters, then it shall also provide features that enable voters who lack fine motor control or the use of their hands to perform the submission. This is also going to be talked about in some detail tomorrow. So we recognize privacy is an important part of accessibility, and for people with dexterity issues there’s been suggestions and some ability to use a privacy suite to preserve that. But this requirement goes beyond that and says it also requires, as a shall, independence. And this does have some implications for the software independence and accessibility for

electronic ballot markers and precinct count optical scanners, because I'm not aware of any commercial systems that do indeed address this. So again, I'm going to put this on your radar screen because this issue is going to come up again tomorrow in the (indiscernible) accessibility discussion. It probably also will require some discussion with the EAC as well, since this is in the current version, of the VVSG '05 version standard.

UNIDENTIFIED INDIVIDUAL: So you're requiring this of all voting machines, or are you just saying it would be some special machines that could do this?

MS. LESKOWSKI: This is for the accessible voting station.

UNIDENTIFIED INDIVIDUAL: Okay.

MS. LESKOWSKI: All right? Okay.

So that completes my discussion of the issues that we're currently chewing on. I want to talk now about where we are with our usability performance benchmark research. And we've completed the first phase, so our overall goal for the VVSG is to have quantitative performance benchmark requirements for usability with

conformance determined by running usability tests with typical voters.

So here are the steps that we need to do to get to that point. Develop a test protocol and metrics - I'm going to talk about that in a moment. The checkmark is there because we've done that. Show the test is valid - we believe we have shown that our test is valid. I'll go into some details as to what I mean by test validity. Show the test is reliable - this is the next stage of our testing that's going on right now. By that I mean that we can reproduce it and repeat it. That is, the same testers can -- let me see if I can get this straight -- can reproduce it with the same results and that it can be repeated by -- no. Let's see. I reversed it again. I always do this -- that the same set of testers under the same conditions can repeat it and get the same results statistically speaking, and that another test lab can perform that test and get the same results.

The next step is to test a number of commercial machines so we get an idea of what their performance baseline is for this test, for this specific test

protocol and for those metrics. So some of the issues that we're in the process of dealing with right now are, how do you do this cost effectively, because we want a large enough number of voters to ensure statistically significant results. And we want tight confidence intervals, so I might say, I ran this test with ten voters and eight out of ten completed the ballot without errors. I'm giving you a way to count here. So that's binary. Yes or no, did they have a perfect ballot or not. I'll talk a little bit more about different ways to count and different ways to define errors in a moment, but for the purposes of this --

So I ran it with ten voters and eight out of ten produced a perfect ballot with our test protocol and our test ballot. Or if I told you I ran it with 100 voters and 80 of them completed it successfully with a perfect ballot, you would have a lot more confidence, a tighter confidence interval within that larger number of users. So we're trying to find out, to balance enough voters so we get a good, tight confidence interval on our results, but not thousands of users to make the test too costly to run.

(Speaker not using microphone.)

MS. LESKOWSKI: Microphone, please?

UNIDENTIFIED INDIVIDUAL: There's another variable, too. I mean, if I did it with, let's say, a very homogeneous population --

MS. LESKOWSKI: I'm going to talk about that in a moment.

UNIDENTIFIED INDIVIDUAL: All right.

MS. LESKOWSKI: Okay. In fact, I'm going to make two statements about that. There's two aspects of that. I've been thinking about this for a long time. And there's different ways of counting, and that determines our metrics system, which ones do we want to use and what is the statistical treatment of these metrics to determine the competence in these (indiscernible) because in general we don't get normal distributions on these and we didn't expect to. So we have been working with some of the NIST top statisticians to work through that because I'm not a statistician. And at that point we can then, by looking at a performance baseline and how we calculate that performance baseline competence (indiscernible) we can then put in our benchmarks.

So this is sort of an informal description of our test protocol, because I thought I would try to make this more concrete for the committee. Basically we recruit participants with specified demographics. In this initial test round to determine validity we used a rather homogeneous group of people that we expected to get fairly good performance on, just to see if we could get errors and see if we can distinguish between different systems. And I'll talk about that a little bit more in a moment. But obviously for a larger test you want participants with demographics that are relatively representative. It doesn't have to be entirely representative because we're just testing something in a lab. We want enough to generate the different kinds of errors that one would expect to see.

We have a meeting complexity ballot, 20 contests and referendum. We asked vendors to implement that test ballot and to show off their system in the best light. And the test administrators follow a script. Participants are told exactly how to vote so we know their intention, make the ballot -- out the other end, when you cast your vote, make it look like this. There

are 28 entries. No assistance or training is given. We say, here's a voting machine. Whatever is typical, training materials provided around the machine, that's what they see. So we measure errors and time to vote, and basically those errors are differences from what we expected to see given how we told them to vote and whether they were able to cast or not cast the ballot.

And we have (indiscernible) questionnaire. It's a modified survey of user satisfaction or SUS questionnaire that's widely used in the industry. It's modified for voting. It's basically ten statements, it's a five-point (indiscernible) scale. They rate on a scale of one to five things like, I felt confident that I used this voting machine correctly, I think that I would need support to be able to use this voting machine, or I thought this voting machine was easier to use. And this has been validated in a number of different contexts.

And so --

MR. CHAIRMAN: Whitney?

MS. QUISENBERRY: Just to clarify, could you -- well I suppose I should just say the answer since I know

the answer. The ballot that you constructed, was that using real candidates and real parties?

MS. LESKOWSKI: No. We didn't want to -- we try not to bias things, so we use different colors for parties and we made up names that looked like real names but have no relation to any candidates. They're just out of the phone book. I know (indiscernible) actually used some software that generates random names, which would be another way to do that, but we only needed one test ballot.

MS. QUISENBERRY: That's one of the things we're reading in the literature, is that trying to use a real ballot you get people who say, but I didn't want to vote the way you instructed me.

MS. LESKOWSKI: Right. Exactly.

UNIDENTIFIED INDIVIDUAL: Was it just for candidates or did you include like referendums and that type of thing?

MS. LESKOWSKI: Yes. There were three --

UNIDENTIFIED INDIVIDUAL: There are six contests that are yes-, no-type --

MS. LESKOWSKI: Yes. Some were just yes or no, and

I think there were three actual wordy referenda. Okay. So we had 47 test participants, high school through college degrees, 21 though 30. We wanted people who would perform reasonably well, and we had two different types of voting systems because we wanted to ask the question, do we find errors in this rather homogeneous group. Because then for sure we're going to find all of the errors in -- and do we find most of the errors, because if we can find them with this group, that means we can test with smaller populations because for sure we're going to get the whole spectrum of errors. If they did perfectly, then we have to broaden it. And indeed we got the types of errors that we predicted across the board, all the different kinds of errors that you could possibly see.

Does the test detect difference between machines is another way to look at validity, because does the test measure what we want to measure. And we did find that there were differences between the two different types of machines that we used. Are those differences realistic differences? Well, the way to do that is to look at what other kinds of research results are out

there and are those similar. And also do expert usability review to say -- and we would have expected to see this. And do our expectations kind of sound -- and did we see errors that we expected, and do they also kind of look like what the general population, the general public would kind of expect to see also and that they hear about. And we did see that.

So the question is, were the differences statistically significant for the errors. And they were for these kinds of errors that we expected. Time on task did not show statistical significance. Could be for several reasons because the machines were very similar. If we got some radically different kinds of limitations we might see statistical significance. We're not terribly concerned about that because we do expect to see a lot more variability. And it may be the case that even with a lot more users and different machines you're still going to see such variability that you can't really show statistically significant differences. And time on task also depends a lot on the individual's circumstances and the users, etc.

Dan, did you have a question?

MR. RIVEST: (Indiscernible) just curious if you went into that. That's people, when they aren't familiar with something the first time they may have more errors. And sometimes when they get more familiar with that particular device and the way of interacting, they adapt to it and they do better. So --

MS. LESKOWSKI: Well, kinds of numbers of errors were very similar to what we're seeing in the research literature where they did training and they did a whole lot of other things.

MR. RIVEST: I mean, you didn't try like repetitive use to come back and do it again?

MS. QUISENBERRY: Was this between or within participant --

MS. LESKOWSKI: Between.

MS. QUISENBERRY: So each participant only voted on one of the systems?

MS. LEWKOWSKI: That's correct.

(Speaker not using microphone.)

MR. RIVEST: If I voted on a system that I'm familiar with, I might do better than a system I was unfamiliar with. But it doesn't mean that if I went

back to that system I was unfamiliar with another time or two which is might what happen in practice -- you might a new machine in, you might have a little difficulty the first election, but after some repetitive elections you might do just as well. In other words, you might be favoring the one that people are more familiar with rather than the new one. That's all I'm saying.

MS. LEWKOSKI: Most of our users didn't have a huge amount of experience with these. One was a DRE, one was an optical scan. People know how to fill in bubbles, so they were very different.

MR. CHAIRMAN: There's another question.

UNIDENTIFIED INDIVIDUAL: I'm curious about the age range. You said 21 to 30. That's our poorest age range for voting, first of all. And you might see different types of errors maybe for an older population.

MS. LESKOWSKI: I think we pretty much saw all the kinds of errors you would expect. You might see a worse error rate. You might see more errors with an older population.

UNIDENTIFIED INDIVIDUAL: Right.

MS. LESKOWSKI: I expect you would.

UNIDENTIFIED INDIVIDUAL: I have --

MS. LEWKOWSKI: And that's the typical voter --

(Multiple speakers, speakers not using microphone.)

MS. LESKOWSKI: No, no. I'm going to talk about the next stage.

MS. QUISENBERRY: But I have to say I was quite surprised. After 2000 there was a lot of speculation among the political science and human factors testing community about how many people you would need to be able to find a subtle error. And the way this ballot was constructed -- and Sharon, please correct me if I'm wrong, but the way the ballot and the instructions were constructed were to test different types of conditions, like one race has a lot of candidates and they're asked to vote for someone low on the list, for example. And so it's both a little frightening and a little encouraging that with a small group of relatively unchallenged -- of voters who we expect to perform well that we were nonetheless seeing those errors. Because it suggests that the threshold at which you begin to see them is not thousands of people, but dozens of people.

MS. LESKOWSKI: Our purpose here, just to

(indiscernible) was to test the protocol, not the machine. So the fact that we were able to do this with a small number of users of people that you would expect would do well, and you still measured the range of errors with this ballot, supports the validity of the test protocol.

Okay. And most people have what are considered good SUS scores, and in particular they were confident that they voted correctly. And we didn't see any (indiscernible) significant differences between systems. But we expected this to be the case that the important benchmark here is of course did they cast their vote as they intended. If they take a little bit longer one way or the other. Some voters were happier than others. That's not as critical, but with our thinking of using time on task and SUS scores to at least report them and put a lower bound that says, if a system scores worse than this, this system has big problems. And to use those benchmarks in that way.

So there's lots of different ways to count errors for the effect of this benchmark. You could just say did they do just the strict binary, did they fill it out

correctly in total or not. And a second binary was, did they cast it or not. And you could calculate through the success rate as the number correct over total number of participants. That tends to have very loose confidence intervals, so if we went with binary we'd probably have to test with a lot of users. But for our next experiments, we can calculate the errors any way we want because we'll have the data on how they performed. So we're just sort of outlining, what are the different ways. And we're going to look at these number of errors and our competence rates when we do our statistical analysis and pick what we think is the best way to count errors.

So you can count number of errors for each contest, you can look at each possible entry the voter could make. And either they should have voted for this and they didn't. That's an error. Or they should not have voted for this candidate but they did. That's another kind of error. And you can count up all those kinds of errors, or you can count and weight different kinds of errors as more serious than others. You can look at the number of individuals making a particular kind of error.

So in general, depending on how you're counting, you count those number of errors and you divide by the number of participants times the voting opportunities per participant if you're not doing binary, and you can get an error rate.

MR. RIVEST: Question.

MS. LESKOWSKI: Yes.

MR. RIVEST: Ron Rivest. Did you have write-in votes on this?

MS. LESKOWSKI: Yes, we did.

MR. RIVEST: How do you count errors for write-ins?

MS. LESKOWKSI: Well, we told them what to write in, so either it was correct or not.

(Multiple speakers.)

MS. LESKOWSKI: Yes. Well, not necessarily. For DRE it would be typed in.

MS. QUISENBERRY: Sharon?

MS. LESKOWSKI: Yes.

MS. QUISENBERRY: Now, this is me, my opinion here. But it seemed to me that one of the discussion points we might have is whether we want to create a benchmark for errors, or whether there might be three or four

different metrics. For instance, surely failure to cast needs to be treated specially and we want to see that number be very, very low. We might want to look at how many different people have errors. That is, is this concentrated in certain parts of the population or others. We might want to look at whether -- a number of people might have a few errors scattered across the ballot, but you also might have a few people who have a lot of errors. And you might want to look at the distribution of errors across the races. So you might have only two errors on any ballot, but they always occurred in two of the tasks on this ballot, which would indicate a problem. And I started thinking about what are the kinds of usability errors that are not about perfect world, but are about indicating the possibility that the usability of this system --

(Speakers not using microphone.)

MS. QUISENBERRY: -- such that it could change an election. Because that's really what we're after. It's not perfection (indiscernible). And so we might end up wanting to say, we're going to measure it, we're going to measure three or four different aspects of errors,

and it has to exceed the threshold in all of them. Because any one of them could indicate a kind of problem. And I've thrown this out to them and they sort of said ah.

MS. LESKOWSKI: Well, the problem is you do have to do a statistical analysis. You can't just say okay, we ran this test, that is the number you have to --

MS. QUISENBERRY: No, no. We might not have to choose between these benchmarks. We might be able to say, there are two or three of them that, when we look at the data, that are more likely to indicate a kind of problem. And we'd want to see a successful passing of a threshold in all of those.

UNIDENTIFIED INDIVIDUAL: Question. To that point, Whitney, are you still talking about test protocol? So you're still just testing, you're not talking about the voting public?

MS. QUISENBERRY: No. We're just talking about a test protocol. And I think we should say the same about this test protocol as we would about any test protocol. We just had a discussion about accuracy and perfectly-marked ballots versus real-world ballots. Like any

test, we are creating a little bit of a bubble and saying, this is exactly how we'll test. That will not map exactly to any ballot, necessarily any voter, necessarily any precinct, but it's a --

MS. LESKOWSKI: Well, we hope it's a prediction. It's a lab test. We hope it does give you a prediction of performance out in the field.

MS. QUISENBERRY: Right. And that's why (indiscernible) --

MS. LESKOWSKI: A controlled experiment.

MS. QUISENERRY: -- not to go out and say, what are the under vote errors, what are the under vote counts out in the field. It's what are the errors that occur in this test against different systems with appropriately-represented populations.

MS. LESKOWSKI: Because that would be the one I would be concerned about the most, is the under votes. Because people don't realize how many under votes there are.

UNIDENTIFIED INDIVIDUAL: Yes.

MS. QUISENBERRY: And as I understand it, part of the instructions for voting this ballot include

instructions to under vote.

MS. LESKOWSKI: We tried to be as comprehensive as we could with the different tasks for voting.

Okay. All right, so we're currently running some experiments to determine reliability. So test repeatability, can the test results be repeated with the same test administrators and the same kind of participant demographics, same participant. Can it be reproduced, can the test results be reproduced in different geographic regions-- this gets to part of your question -- with different test administrators.

So we're planning a series of tests. We're got a larger set of participants. We may go up to 400 participants with a mix of age range, female/male, different socioeconomic standards and geographic region. We're going to probably use Virginia, Maryland, D.C. since it's expensive to get test participants across the country. But that actually gives you urban and suburban and rural areas. So that actually does give us a much wider geographic area that we did for the validity tests.

We're going to do a series of tests to see if we do

get repeatability or reproducibility. And we may need to repeat them with some adjustments, depending on the earlier results. And then we'll also bring in a wider range of commercial systems, because we've got to figure out a good baseline. So from the performance gathered from a wider representative set rather than just two, we'll have to calculate sort of a baseline that most can reach but that is not so low that it's trivial to reach that baseline.

Now, let me point out that we are not talking about participants with disabilities here. These are people using not the accessible but the regular voting station. We're assuming that they are typical but they're not designated as having particular disabilities, because one would hope that our baseline for errors would be similar and that we could still use that. But we don't know what kind of variability, what our confidence intervals are going to look like, and what our rates are going to look like because we have to test with those specific users. And that's kind of a next stage of research for the future.

MS. QUISENBERRY: And more precisely, if you're

using the audio ballot, you're actually using a different (indiscernible).

MS. LESKOWSKI: Yes, it's a different system. We certainly would expect the time to be longer with the audio ballot.

Any questions? So hopefully we'll get a baseline very soon.

MR. CHAIRMAN: What is the timescale?

MS. LESKOWSKI: We do want to get something into this version of the VVSG '07 and we hope to complete the experiments by the end of April, beginning of May, somewhere in there. But then we'll have to do some analysis, so we're really pushing.

MR. CHAIRMAN: It's tight.

MS. LESKOWSKI: It's tight. We're really pushing.

MR. WAGNER: David Wagner. I'm not a usability expert, but let me just say this. Sharon, I think that you and NIST and HFP are just doing an absolutely phenomenal job here and thank you. That sounded great.

MS. LESKOWSKI: Oh, thank you very much.

Okay. Next research steps. We heard this morning from Donetta that they are going to be putting out some

guidance on ballot design. When that is finally accepted by the EAC, we hope there's time to look at it and to make sure -- to look it over to see if there's something that we can also reflect in the equipment standard. But we haven't see it yet, so we don't know.

We are currently doing some research on additional voting-specific plain language guidance. Not sure if that's going to make it into the standard, but it may be more appropriate in any case to make it as a guidance document for suggestions to the vendor for a wording that works better, and just make it as a good guide. And it's similar for color. We have some color requirements, but really there is research out there that we need to collect up and just say, these color combinations will work, this is based on best practice.

We also want to do a small analysis of looking at how and when to use icons and pictures appropriately so that we don't introduce bias, and that they are helpful and not a distraction to those with cognitive disabilities. We're also going to try to be working on again some guidance documents. So we've put into the documentation volume requirement that talks about that

the documentation should be usable, and so we thought about how do you write requirements for that. And we weren't sure how to do that, but we said ah, there's a lot of technical communication experts that do this all the time. And one thing they do is write style guides that say, this is a good way to make sure that this documentation is coherent and easy to read and look at. So we said well, we could do a template that in a sense would be a test method for saying, judging whether, as I said, the documentation was usable or not

And then I already talked about generating some accessibility performance benchmarks, but we're way off from that.

MS. QUISENBERRY: Sharon, thank you. I'd just like to throw in that one of the issues that has come up and that Commissioner Davidson referred to is the question of how and when the standard can be updated. And given the time constraints, I'm particularly concerned that we can add in accessibility benchmarks as they're developed, and that we don't either leave them out entirely because we don't make this deadline, or rush and create bad ones because we're rushing. So that's

when we're, it would be the same test protocol, the same basic requirements, but where the benchmark might not be done in time for July. Well, will it be done in time for July is probably a fairer statement. So as whoever the "we" is, as it is considered how this is updated, this is one of the issues I'd like to see kept in mind.

MR. CHAIRMAN: It's Bill Jeffrey. I have a question for Sharon and Whitney. On what kind of timescale do you think reasonable benchmarks might be defensible? Is it that it's August as opposed to July, or it's August but of 2011?

MS. QUISENBERRY: I think that depend in part on NIST procurement.

MR CHAIRMAN: We may have some influence on that.

MS. LESKOWSKI: Well, I'm thinking '08, not --

MS. QUISENBERRY: Early '08?

MS. LESKOWSKI: -- early '08.

MR. GANNON: This is Patrick Gannon. I'm not sure if my question ties into the testing as much as the experience from ongoing voting activities, especially with DREs, and how that experience is playing into the development of the VVSG 2007. And I was specifically

intrigued with the report that came out last month. I think David was a part of that, the (indiscernible) report on Sarasota. And it seemed to indicate that it wasn't a system issue but it was more how the ballot was actually set up. And the fact that there is now lawsuits and potential bills in Congress and so forth coming out of that, is there something that is already in or will be put into the VVSG that provides guidelines to say, you know, to do (indiscernible) means you lay it out, right?

MS. LESKOWSKI: It does appear that there was a usability problem. It's hard to tell for 100% sure, but there was a Dartmouth report also that suggested that. And as I said, there is valid design guidance coming from the EAC that may help, but I haven't seen it yet so I can't speak to that. That's not NIST work. We do have things like, you know, consistency, consistent wording

MR. CHAIRMAN: Commissioner Davidson?

MS. DAVIDSON: Well, the ballot design, we have a meeting in Kansas City. That is the main purpose for that, probably carrying that will April the 18th. So

that will be out at that time. I mean, that will be coming very shortly after that timeframe. But I do have a concern with your moving target. What you're doing is creating for manufacturers a continued change in standards and guidelines. And that is what we're tying to get away from, because that's where our cost comes in. Every time you change something for the manufacturers, if we have a July date and then we come back and we have an August date or we have a next year date, early '08, I don't know how they can meet that. And that just pushes -- that's my opinion, the EAC, but Donetta's.

(**END OF AUDIOTAPE 3, SIDE B**)

* * * * *

(**START OF AUDIOTAPE 4, SIDE A**)

MS. QUISENBERRY: -- acceptable benchmark against this test. It doesn't matter whether you're testing a paper ballot with an optical scanner or whether you're testing the audio ballot of a system that, in fact, the time might be different. Because if a ballot has a long referendum on it, that takes longer to read out loud, or if someone can listen with the tempo turned up, that

would take less time. But the accuracy- and error-related benchmarks I think should actually be the same across the board, no matter how you vote. So we might actually be able to solve this quite easily that way.

MS. LESKOWSKI: Yes. We need still a little experimentation, but that --

MS. QUISENBERRY: I would be nice to have --

MS. LESKOWSKI: But let me also point out that the benefit of having a performance benchmark with a test protocol is that the vendors can run this themselves once we put all the data out. They can run this. They can use the same protocol for any reporting they do. Also at the state level, back to the Florida ballot question, they can certainly use that test protocol with one of their own ballots just to sort of see what kind of errors are they getting.

MS. QUISENBERRY: Yes. The other thing, the example you brought up, before one of the election officials beats me to it, is that there’s that narrow line between the ballot layout capabilities of the equipment and actually laying out the ballot. Because there is some human variation in that it’s one of the

reasons, as I understand it, why we ask the vendors to lay out the ballot for this. There was no question that the test lab didn't do a good job. I know that this project has been going on. I've been waiting with baited breath to hear the results, because I think the group doing it is interesting. Creating some ballot layout guidance for the election officials to use, one of the things that we want to do is look at that when it comes out -- now we know it's April 18th -- that we want to be able to look back at it and say, are there things that they're suggesting as good practice that we could add to or amend our requirement to make sure that the systems support that or even encourage that.

UNIDENTIFIED INDIVIDUAL: Well, when you're talk about and when you test it, there's two very different types of devices, like an optical scanner in this hand and a DRE in that hand. Then let the vendors design it any way they want. But if I'm talking about like I'm benchmarking two DREs or two optical scans, then it could very well be that one vendor just is better at laying out the design. The equipment really is no better.

MS. QUISENBERRY: But ultimately someone has to lay out that ballot.

UNIDENTIFIED INDIVIDUAL: I understand. So wouldn't you want it something more representative of what ballot designs there are coming out of the field to practice for those things?

MS. QUISENBERRY: An election official might like to lay out the ballot for the test.

UNIDENTIFIED INDIVIDUAL: Well, what do you want, to be able to get some output from these results that might give you feedback on better ways to do ballot design?

MS. QUISENBERRY: That's something that the vendors might get out of seeing the results (indiscernible) but it's not the purpose of the test. And I think we need to be very careful about distinguishing between an evaluation that tests the performance of the system under certain circumstances and design guidance back to the vendors.

UNIDENTIFIED INDIVIDUAL: Well, let me put it a different way. Suppose you did a test with the same piece of equipment and you handled two different ballot

designs and you found you had more variability in the test results that way than from the equipment themselves --

MS. QUISENBERRY: I wouldn't be too --

UNIDENTIFIED INDIVIDUAL: -- which might be the case.

MS. QUISENBERRY: I would be particularly surprised by that result. They're certainly very easy to do bad design and it's hard to do good design, but ultimately we are not testing the ballot design capability of election officials, although we are trying to encourage systems that provide good design. There are also aspects, especially on the DREs, there are aspects of the systems that can't be changed by the election official. And we certainly want to make sure that those do not put them in a situation where they can't design a good ballot, where a usability problem is designed into the system. And I think it is a very difficult area to separate which is which, and I think that I've been very impressed with the process that you've gone through to really sort of sort through the problems very carefully, and to make sure that the test itself is not inducing

any more bias than any test inevitably induces. I'm sure that having an atomic clock that we've all agreed on.

MR. CHAIRMAN: Ron?

MR. RIVEST: Yes, a couple of things. First, just let me second David's compliments here on the work you've been doing. It looks great. I had a minor question, one you've probably thought about but I wanted to hear the answer. It seems when you introduce this general adjustability you introduce some hazards with the voter turning off the audio or changing languages to a language you can't read, or -- I don't know if this is a requirement, for getting reset between voting sessions with different voters. So I just wanted to inquire about that.

MS. LESKOWSKI: There's a reset back to --

MS. QUISENBERRY: It's in '05.

MR. RIVEST: So that's include as part of this, so any voter at any time can reset to a standard state? Is that the requirement?

MS. LESKOWSKI: Yes.

MS. QUISENBERRY: And furthermore that the machine

resets to a standard state between voters so that a voter doesn't come in and find the machine having been adjusted for a previous --

UNIDENTIFIED INDIVIDUAL: Or with the screen off or something, or you hear Chinese.

MR. CHAIRMAN: I have sort of a comment for the group. I'd actually like to follow up on Commissioner Davidson's comment. You know, I agree with her sentiments and I think it would be difficult for us to put out next iteration guidelines that starts going to the Standards Board, public comment and others with TBDs in there. And if there are possibilities of simply coming up with a consistent way of looking at it, and quite frankly that was a somewhat intuitively compelling concept that you proposed, and while the testing is done, I think what the testing would do at the end would either validate, it could be used to help validate the assumptions. And if there's an egregious error that arises in that, it may be easier to get forgiveness in correcting an egregious error than having to go back through the process entirely.

UNIDENTIFIED INDIVIDUAL: That's a well-taken

point. I know from being a person who works in the federal system occasionally that procurement and arranging the mechanics by which the work can be done can be a challenge. So I turn to you as the head of NIST to do anything you can to help smooth that process.

MR. CHAIRMAN: And I --

UNIDENTIFIED INDIVIDUAL: Assuming there are issues there, which --

MR. CHAIRMAN: Right. Yes, I will formally task my staff to talk to me immediately after this about whatever issues there may be on that. Commissioner Davidson?

MS. DAVIDSON: I have a question that I would like to ask. I didn't see anything on the presentation on a usability on paper ballots. And where it comes from is in the software independence resolution. There was a requirement for paper-based machine and, you know, I seem to think that we need some type of a study to go along with that, or have you already (indiscernible)?

UNIDENTIFIED INDIVIDUAL: I'm trying to understand your question. The Opti-Scan was in this validity was --

UNIDENTIFIED INDIVIDUAL: You mean with paper records?

UNIDENTIFIED INDIVIDUAL: (Indiscernible) voter verification paper trail, the ability to accurately verify. Is that what you mean?

MS. DAVIDSON: Well, you know, just a study on the usability of paper I think is real important, and I just didn't see that at I wasn't sure if it had been discussed or not.

MS. QUISENBERRY: Well, I guess there are a couple of ways to look at it. One is that if the usability test encompasses any system that might be certified, that would include manual and marked paper ballots, that would include electronically-marked paper ballots, that would include DREs with a VVPAT on it, that it would include the entire spectrum. So that's one answer. Sorry, I can't see you past the podium and I don't know if that's the question you were trying to ask.

MS. DAVIDSON: I just didn't see that type of a study being done to see how people react to the paper, you know. And that was one of my concerns. And I don't know if there is (indiscernible) or anything like that.

MS. QUISENBERRY: Oh, no. I mean, specifically looking at what happens when a voter is confronted with a paper audit trail? I don't think we have anything like that on the schedule.

MR. WAGNER: David Wagner. It seems to me that, if I understood correctly, what you are trying to accomplish in HFP is to design a protocol that you can use for performing usability testing of any system, whether it has a VVPAT or not. So does that -- I mean, it seems like that isn't the scope of the TGDC to do new research on how to design the best VVPAT or something like that necessarily. Is that correctly --

MS. DAVIDSON: Well, you know, when they're having so many difficulties in counting that paper --

MR. CHAIRMAN: Oh, oh, oh. Okay.

(Multiple speakers.)

MS. QUISENBRRY: Usability for election officials of the paper?

MS. DAVIDSON: Right. I mean, the election officials are complaining that this is very difficult. And I know there's been some discussion about bar codes, whether they should be used or should not be used, but

I'll tell you what, they do it by hand is a disaster.

MS. LESKOWSKI: Well, I think the issue for me in trying to design what would the study be, we know there's difficulties hand counting paper. There's no sense in running another study when we have a lot of data that already us that. There are different ways to help with that. So I've had trouble formulating what does it --

MS. DAVIDSON: I probably should come to the mic. But, you know, there's certain things that the manufacturers are doing right now. Some are using the bar code. Is that successful? You know, there are some studies there that maybe would create a difference in the minds of the TGDC members if there is something that would be more successful than obviously hand counting that ballot. What kind of ability, can they scan it, or whatever. But there are things out there right now, and so I just wondered if that was being thought of.

UNIDENTIFIED INDIVIDUAL: I think that thought is on something we're going to discuss --

MS. LESKOWSKI: I think Alan needs to identify who is talking --

(Multiple speakers.)

MR. SCHUTZER: Dan Schutzer. I think I --

MS. LESKOWSKI: We had -- and earlier was Whitney Quisenberry and Donetta Davidson.

MR. SCHUTZER: I think we -- well, I don't know if we addressed the whole thing, but I think that borders on something we wanted to discuss tomorrow on the VV&P AT and ways to improve it and research and so forth. So if you want to hold that thought, and if we don't address it at all, then make sure we change what we're talking about to address those issues. But I think that is addressed. Wouldn't you agree, John, that we're starting to border on that a little bit?

UNIDENTIFIED INDIVIDUAL: Yes.

MR. CHAIRMAN: Ron?

MR. RIVEST: Yes, I just wanted -- I think that's a great issue, the usability of the audit. The audit is very important for the integrity of the elections, and so being able to make sure that that's usable for the poll workers is very important. Bar codes is something we've discussed a lot in STS and so on, too. And if you have a bar code, it's something that the voter can't

check himself and you've got a real issue as to whether verifying the bar codes is doing what you want in terms of the integrity verification for the election. But there are approaches to working with bar codes and human readable, too. Lots of interesting approaches, and I agree that's a great area for research and further improvement. I'm not sure how much we can put into the standards in the timeframe we've got here, but hand counting in some sense is sort of the gold standard, as sloppy as it is, for looking at paper ballots.

MR. CHAIRMAN: Whitney?

MS. QUISENBERRY: I'm sorry.

MR. CHAIRMAN: So if I can add to what Ron said, let's just take this as a plea for input and suggestions and further comments on this issue, because I think it is a critical one.

MS. QUISENBERRY: Yes, I think one of the things that's a real challenge for those of us who are not election officials is that it's sort of easy for us to imagine the usability challenges in voting, because we are voters and because that task it fairly well documented and well understood. I have to say that if

security is (indiscernible) what you guys do is at least a magical art. And so help in understanding how to formulate a question that could be answered -- what is it we want to do? Do we want to do time tests of different types of audits, I mean --

MS. LESKOWSKI: What's been helpful to me is, is there some requirement that would go one way or the other that we could do some research that would inform us and tell us what that requirement it. And I --

MS. QUISENBERRY: So I think we all kind of understand the general problem, but not how to get down to something specific enough that we can charge somebody with doing the research.

MS. PURCELL: Helen Purcell. If I could, main thing is that we don't make the errors that we have made in the past. And part of that being we were given certain things that we had to accomplish by the 2006 election, both the election officials and in particular the manufacturers, given very little time to do that we were given DREs and in some states we were required with the DREs to have VVPATs. And with that we had things added to a DRE that gave us a big printer that had tape

in it that a lot of people couldn't handle, it couldn't be changed. So if we don't get into a scenario of going back to the same thing of giving us something else in a short period of time that we have to do and the manufacturers have to provide us, you know, anything we can do to avoid that.

MS. QUISENBERRY: One thing I would point out is that one of the things in here in the VVSG '05 we had requirements for usability --

(Off the record.)

MS. QUISENBERRY: -- for the generally usability and then three groups of testing with different interfaces for people with disabilities. We've added one in this which is testing with poll workers, so we would actually be doing usability tests of the setup and operations. And I think some of the things that we've heard about poll workers not being able to change the paper well, that those things would come out in that test. So that is the one thing -- it's not a test of the audit, but it is certainly a test of the during-election maintenance type and operation stuff. So we've gotten that piece in. How we get to the next phase,

which is the audit, is the one that I find a challenge.

MR. GAYLE: I always get concerned when we start talking about what I'll call the third rail for the election administrators. And that's when we start writing standards for them in terms of poll worker conduct and poll worker training. I think that's not our jurisdiction.

MS. QUISENBERRY: Absolutely. I was thinking more of things like, can you follow the instructions to open the thing, can someone given a set of instructions that say to change the paper do the following four steps, can they follow those steps. And those are manufacturer's instructions.

UNIDENTIFIED INDIVIDUAL: Okay. If I could respond a little bit to Secretary Gayle. The intent is not to come up with any requirements for audits or for how they have to be conducted. The intent really is to look at the paper itself that gets produced in VVPAT systems or OpScan or whatever, and what can be done to that paper or to the format of it, or to the format of, you know, beginning of day/end of day reports so that it is easier for poll workers to handle them, so that it is easier

for election officials to use an audit. But absolutely no requirements for how they should be used, it's just basically make it easier to use.

MR. CHAIRMAN: Are there any other comments or questions? Sharon, did you get the information that you need out of this session to continue to move forward?

MS. LESKOWSKI: Yes, I have.

MR. CHAIRMAN: Okay. If there are no other questions or comments, do I hear a motion to adopt the preliminary draft, Human Factors and Privacy Sections consistent with the discussion that we've had? There's been a motion. Is there a second? There's a motion and a second. Is there any objection to unanimous consent? Hearing no objection, this passes by unanimous consent.

That actually ends today's discussion almost an hour early. So thank you. Those of you who have to catch the bus to the Metro, you'll have no problems. For the rest of you, we reconvene tomorrow morning at 8:30. So at 8:30 back in the same room. Right, Alan?

UNIDENTIFIED INDIVIDUAL: Yes.

MR. CHAIRMAN: Same room.

UNIDENTIFIED INDIVIDUAL: You can leave your stuff

here and we'll lock up.

MR. CHAIRMAN: Okay. So again, thank you very much. I'd like to thank the EAC Commissioners for attending and providing valuable input. Meeting is adjourned for today.

(**END OF AUDIOTAPE 4, SIDE A**)

* * * * *

(**AUDIOTAPE 4, SIDE B, BLANK**)

* * * * *

CERTIFICATE OF AGENCY

I, Carol J. Schwartz, President of Carol J. Thomas Stenotype Reporting Services, Inc., do hereby certify we were authorized to transcribe the submitted cassette tapes, and that thereafter these proceedings were transcribed under our supervision, and I further certify that the forgoing transcription contains a full, true and correct transcription of the cassettes furnished, to the best of our ability.

_____

CAROL J. SCHWARTZ

PRESIDENT