11:35AM 1 Had they talked to each other between themselves they
11:35AM 2 would not have been able to figure anything out.
11:35AM 3 We tried to be as solid as possible in
11:35AM 4 designing these tests. For Murphy's and my tests we
11:35AM 5 consulted with some of the scientists at the research
11:35AM 6 unit to make sure we hadn't missed anything. As far as
11:35AM 7 we went, we were solid on that.
11:35AM 8 But is it totally blind? No. Would it be
11:35AM 9 ideal if they were totally blind? Yes. And we would do
11:35AM 10 that except it's extraordinarily impractical and
11:35AM 11 virtually impossible. People can figure out if they're
11:35AM 12 taking a test or not. They can call a contributor and
11:35AM 13 it's going to be no time they can figure it out.
11:35AM 14 Are the tests too easy owing to pristine
11:36AM 15 samples? In all the validity and proficiency tests I
11:36AM 16 know of today we have test fired samples or samples --
11:36AM 17 they're fairly pristine. That's a debatable and
11:36AM 18 theoretical issue.
11:36AM 19 On the one hand I have written on this back
11:36AM 20 in 2000. In theory if you have changing barrels and so
11:36AM 21 forth you could probably create a bit of a problem. In
11:36AM 22 reality with the kinds of examinations we do that are
11:36AM 23 subjective I don't think that's going to be a problem
11:36AM 24 simply because as specimens degrade you simply get to
11:36AM 25 the point where you are saying it's an inconclusive.

I can give you kind of an illustration. Imagine that you have a photograph on a screen that's very fuzzy and you have a bunch of expert observers in the audience. This photograph is of Tiger Woods but you can't tell but everyone in the audience is a family member and knows Tiger Woods very well.

Now you bring the image into better and better focus. You tell the audience members, "When you recognize the person raise your hand." Over time there's going to be people raise their hands. When it gets completely in focus they will all raise their hand.

But go the reverse and they will take their hand down, if you could work it that way. You're saying inconclusive. The imaged samples or samples from guns that have been changed with changed barrels, I really don't think that's going to cause a problem.

What it will do is reduce your sensitivity and increase your rate of inconclusives. At least in the criminal justice arena it's basically a neutral outcome.

What about the confirmation and results a lot of labs will confirm in proficiency tests? A second examiner confirms the results. That's true and we did it in the FBI lab. We were testing the system because that's the way it's done in casework. Validity testing

it's not done but it's done in proficiency testing quite often.

What about a lack of mandatory returns in proficiency tests? What if you get a proficiency test and you're a trainee and there's no requirement you return it? You think to yourself, "This is difficult. I don't know about this. I don't think I'm going to return it," and you don't.

That will skew the data. Can that result in a little bit of an understated error rate? Yes. And that criticism is valid.

What about the fact that creators and administrators of validity tests are not independent? I have mentioned internal FBI tests. We weren't an outside agency doing this. Can we be criticized for that? Yes. Frankly, I wish that Dr. Speakman over at Texas A&M would do these tests.

He's talked about this before. He came up with this and then he said something to the effect that the FBI turned him down. That's news to me. If I had heard about it I would have advocated it be done. I would love for outside independent experts like him to conduct tests like this.

It would give it at least the sheen of greater credibility. I don't think you can design a

11:39AM 1 better test than what we did but it would have his name
11:39AM 2 on it. The difficulty of that is, as a practical side
11:39AM 3 bar here, he comes from a different arena.
11:39AM 4 He's not understanding that -- I think he
11:39AM 5 would not understand that firearms examiners have a very
11:39AM 6 definite stake in this in terms of their careers. You
11:39AM 7 want to have those tests anonymous. Otherwise, they
11:39AM 8 might have this perverse incentive for inconclusive.
11:39AM 9 I think from reading his material he wants
11:39AM 10 it to be completely transparent. I would love for him
11:40AM 11 to discuss it with laboratories, SWGGUN and others to
11:40AM 12 create a test and I think something could be done. I
11:40AM 13 don't think you're going to get any different results
11:40AM 14 but at least it would have the appearance of outside
11:40AM 15 independent experts.
11:40AM 16 What about claims of overstated error
11:40AM 17 rates? That is to say, "Look. These rates are higher
11:40AM 18 than they really should be. They're really lower than
11:40AM 19 that." One criticism of proficiency tests is that
11:40AM 20 anyone can participate. Trainees, I just mentioned
11:40AM 21 that.
11:40AM 22 If anybody in this room wanted to take one
11:40AM 23 of those tests, they could do it if they pay the money.
11:40AM 24 That's my understanding. Yes, that tends to inflate the
11:40AM 25 error rate figures.

11:40AM Also in validity tests very often we use consecutively-made breech faces, barrels, screwdrivers. We're testing for the possibility of subclass markings. These tests are oftentimes much more severe at least in theory than in casework. That would tend to inflate the error rate.

11:41AM The last one is I think a very important factor. It's not talked about much. It has to do with what I mentioned earlier about the probative value of a class characteristic match. That is that when you are looking at a validity test specimen and proficiency test specimen very often you're looking at specimens with the same class characteristics. You're only discriminating between them based on their microscopic characteristics in most cases.

11:41AM That does not mimic casework at all. In casework you can have all kinds of class characteristics come into the laboratory. So we have two basic kinds of evidence here: microscopic detail and class characteristics that in reality have to be combined together to get the overall probative value of the entire thing.

11:41AM In most validity and proficiency tests you're only looking at one type. You're not discriminating on the basis of class characteristics

11:42AM because everything you're looking at has the same class characteristics. That means one whole classification or category of evidence is ignored. That would tend to also inflate the error rate.

Okay. I only put this up here to point out where I think we stand as a science and also to point out trade-offs involved. Let me give you a quick example.

If it was the 1980s when HIV first started getting into the news and about that time the HIV test was developed, the HIV test by all reports was an exceedingly good test. Its sensitivity or true positive rate was in the area of 99, 99 and a half percent. Its false positive rate was in the area of one-half to one percent.

Let's say you are somebody in the population and you were a hypochondriac, randomly select person, and you went in and said, "I want to have an HIV test to see if I have HIV." Low and behold, the test came out positive. Does this mean you have HIV? Does that mean it's a 99 and a half percent chance you have HIV?

No. The reality was the chances are about six percent. Why is that? That's because it was very rare in the population. You have to account for the

11:43AM 1 base rate or prior odds. At a certain time in the 80s
11:43AM 2 the incidence of HIV was one in 3000 persons selected
11:43AM 3 randomly.
11:43AM 4 So you have a positive result. You
11:43AM 5 probably did not have HIV. If you had a second test and
11:43AM 6 it was positive, you can work out the math. You
11:43AM 7 probably did.
11:43AM 8 The point is no test is perfect. Our tests
11:44AM 9 are not perfect. DNA tests are not perfect. No test is
11:44AM 10 perfect and there are trade-offs. And this receiver
11:44AM 11 operating characteristic curve shows the trade-offs. If
11:44AM 12 you look at the dotted line, that's the line of
11:44AM 13 uselessness. Your true positive rate equals your false
11:44AM 14 positive rate. Whatever you're doing is useless.
11:44AM 15 For all different kinds of tests you can
11:44AM 16 have a family of curves in the upper left-hand portion
11:44AM 17 of that graph. You see if you dialed the sensitivity
11:44AM 18 dial on the breathalyzer or whatever test you have, you
11:44AM 19 can change it so that as you increase your true positive
11:44AM 20 rate, you're also increasing your false positive rate.
11:44AM 21 Some tests you can decide where you want to
11:44AM 22 be on that curve. Where are we? Where is firearms and
11:44AM 23 toolmarks? Judging from the testing done so far our
11:44AM 24 sensitivity could be anywhere between 50 and maybe 80 or
11:45AM 25 90 percent. So maybe a data point up here.

11:45AM 1 What's the false positive rate? As we saw,
11:45AM 2 they're low. We can't pinpoint exactly what it is in
11:45AM 3 the aggregate but we have some boundary values. And so
11:45AM 4 I would think that our data points would be up in here.
11:45AM 5 Overall looking at this in the big context and in the
11:45AM 6 aggregate, our tests are quite, quite good.
11:45AM 7 You can go to the next one.
11:45AM 8 There's the issue of maintenance of
11:45AM 9 standards and controls of the Daubert guidelines. We do
11:45AM 10 that in this business a lot. And these representative
11:45AM 11 documents we have, each laboratory has technical
11:45AM 12 protocols. There are SWGGUN guidelines that are
11:46AM 13 transparent and online.
11:46AM 14 There's the AFTE Technical Procedures
11:46AM 15 Manual that at least examiners can go online and print
11:46AM 16 them out which I have done. There is the AFTE Theory of
11:46AM 17 Identification that even though we're a subjective
11:46AM 18 discipline, it is objectivity stated about how you go
11:46AM 19 about it.
11:46AM 20 There's the AFTE Glossary, the AFTE
11:46AM 21 Training Manual available online and there are
11:46AM 22 accrediting agency manuals. There are standards and
11:46AM 23 controls and it's getting more and more stringent.
11:46AM 24 Q. ASCLD is one of the agencies that is
11:46AM 25 accrediting forensic laboratories?

11:46AM A. Yes. Sounds like kind of a nasty acronym.
11:46AM Association -- gosh, I forgot. American Society of
11:46AM Crime Lab Directors, slash, Laboratory Accrediting
11:46AM Board.
11:46AM Q. And they will accredit firearms laboratories if
11:47AM they meet certain criteria. And there's a process; is
11:47AM that fair?
11:47AM A. Yes. Based on all this theory, practice, the
11:47AM probability logic of it, history and so forth and the
11:47AM results of these tests, how strong are identification
11:47AM and exclusion conclusions? What claims and report
11:47AM language do I think are warranted?
11:47AM Very quickly, "LR," that stands for "likely
11:47AM ratio." Don't worry about that right now. It's just a
11:47AM measure of probative value of a piece of evidence. This
11:47AM number is accurate. This number is accurate. This
11:47AM number is accurate and so is this one. The other ones
11:47AM between 50 and 5000 are there for illustration purposes.
11:47AM I made a few quick estimates. Don't put
11:47AM any stock into those. It's for illustration purposes.
11:48AM I will be the first to admit and to assert that many
11:48AM times in the past firearms examiners would assert
11:48AM absolute certainty.
11:48AM That always bothered me. It's just wrong.
11:48AM It's scientifically untenable. It's bad practice and

11:48AM 1 makes no sense. There is no scientific theory or
11:48AM 2 technique that is absolutely certain, not even Newton's
11:48AM 3 laws of motion. And to assert that is just wrong. It
11:48AM 4 stills bothers me a lot to this day if it ever happens.
11:48AM 5 It's not justified.
11:48AM 6 Practical certainty, reasonable degree of
11:48AM 7 scientific certainty or high confidence or strongly
11:48AM 8 indicative or some phraseology such as this evidence
11:48AM 9 strongly indicates that this barrel -- this bullet fired
11:48AM 10 from this barrel, something like that, those are all
11:48AM 11 perfectly valid.
11:48AM 12 I think they're very reflective of the
11:48AM 13 probative value of the kind of evidence we're talking
11:49AM 14 about. When I say "evidence" I mean findings.
11:49AM 15 What about association or consistent with
11:49AM 16 which kind of implies a likelihood ratio of 50 to 500,
11:49AM 17 along in there? Maybe. But that's weaker than what the
11:49AM 18 underlying logic and underlying evidence indicates. A
11:49AM 19 typical indication or exclusion is really stronger than
11:49AM 20 that.
11:49AM 21 What about more likely than not? I
11:49AM 22 noticed -- I think there was a judge ruled somewhere
11:49AM 23 that the firearms evidence, all the firearms examiner
11:49AM 24 could say is more likely than not. The problem with
11:49AM 25 that is you don't even have to go typically to a

microscopic examination to say that.

All you have to look at, as I was explaining before, is the class characteristics match on a bullet and be a lot stronger than more likely than not. Of no probative value or negative probative value. That's kind of absurd.

That in my view with the ones underlined are perfectly rational and supported and valid conclusions whose strength is reflective of the underlying practice. When I say "reflective" I mean accounting for all kinds of errors that can insinuate themselves into the process.

Q. In all of the research that's been done over the course of time in firearms examinations and you said that -- how long did you say we have been doing -- we being the experts in the field -- have been doing firearms examination comparisons?

A. I'm not an expert even though my Ph.D. is in history. I think 20s and 30s.

Q. Has there ever been any documentation of two guns, two firearms leaving exactly the same markings on evidence?

A. No. But to be fair, you're never going to find that. Even two bullets fired successfully are never going to be exactly identical. It's a little bit

like -- you can make the analogy with a bar code reader. You go to the grocery store and get the bar code with the same prices, the bar code reader reads them as the same price.

If you look close enough the codes are not identical. You have to crank up the magnification and see all kinds of differences. The level that the bar code is reading them it's doing its job with a certain level of performance.

Q. So you don't feel like absolute certainty is an appropriate conclusion. How do you -- how can you justify saying practical certainty or reasonable degree of scientific certainty?

A. You can always play with words. But it's a little bit like if I were to see a photograph, say, of my brother. Am I absolutely certain that's him? No. It's also possible I'm not looking at it right. It's possible there's a double. It's possible that something is just goofy.

I may be very, very confident, however, or practically certain that's my brother. I'm not saying that's a perfect analogy with what we do. We're not recognizing the image of something we have seen before. We're recognizing a certain degree of microscopic correspondence but you learn how to recognize that very

11:52AM readily.

11:52AM Q. Just a few specific questions, if I can, Dr. Bunch. Did you have an opportunity to review the motion filed in this case by the defense to exclude this evidence?

11:52AM A. Yes, I did go over it at least once, yes.

11:52AM Q. I know you probably didn't study it. There were a couple of things I wanted to address with you if I could. There appears to have been a complaint made that -- first let me ask you this: do you know Susan Allen, the firearms examiner in this case?

11:53AM A. No.

11:53AM Q. You have never worked with her; is that right?

11:53AM A. No. I met her today in the witness room. That's the first time.

11:53AM Q. There's a complaint made that she didn't test fire any other gun of the same make and model and compare those to either the test fires from the suspect weapon or the other evidence that was collected. Is that something typically done? Is there a protocol for that?

11:53AM A. No. It's really not necessary. I know sometimes critics will say you need to have an evidence line-up just like with eyewitnesses. Well, two things. One, it's really not necessary. You do that in training

11:53AM 1 and so forth and we do that practically in all the
11:53AM 2 validity and proficiency tests we have had.
11:53AM 3 That's how -- you're presented a whole
11:54AM 4 collection of specimens. There are the error rates.
11:54AM 5 Good or bad. Is it really necessary in a case? No.
11:54AM 6 The other thing is -- it's a dominating factor -- is
11:54AM 7 it's impractical.
11:54AM 8 You can't go out and find another gun --
11:54AM 9 it's not that easy to find another gun with the same
11:54AM 10 class characteristics to test fire with. Some labs have
11:54AM 11 huge firearms collections; some don't. Even the one at
11:54AM 12 the FBI is not going to have multiple samples of every
11:54AM 13 gun.
11:54AM 14 Q. If it did, there's no need in test firing
11:54AM 15 another gun anyway?
11:54AM 16 A. No.
11:54AM 17 Q. You mentioned earlier the CMS or consecutive --
11:54AM 18 A. Consecutively matching striations.
11:54AM 19 Q. Thank you. Can you explain what that is?
11:54AM 20 A. In the late 50s Al Biosotti asked himself, "You
11:55AM 21 know what? I wonder if there's a way we can do this
11:55AM 22 objective -- more objectively?" Like it would be the
11:55AM 23 analogue to a radiologist not reading an x-ray but
11:55AM 24 getting out a measuring stick measuring what he is
11:55AM 25 seeing looking at the gray scale.

He came up with the idea, "Let's look at how many percent matching lines there are and no matches, how many percent matching lines there are and no matches." That wasn't very discriminating at all. When he looked at consecutively matching striations or lines on bullets that was very discriminating. And he could develop kind of a decision rule.

I forget the exact numbers but he test fired a bunch of bullets from known Smith & Wesson handguns and from those that weren't from the same gun and said, "All right. We never saw any more than X number of consecutive matching striations from known non-matches. Therefore if it's more than that, that's safe to say it's an identification."

So that started -- there's been increasing interest in this for various reasons because it would be objective. Theoretically it would be objective. I wrote a paper about this critiquing it back in around 2000 and there is something to be said for it but there's a lot to be -- no good reason really to do it in my opinion.

It's another method. It has been validated. Nothing says it's superior or inferior to the traditional subjective method. My criticism of it was if you are going to start using numbers, there's a

whole lot of things you can do. You can treat it more like DNA.

We can do a lot of statistical analysis and so on and you're going to get into a whole realm that right now we may not be able to get into very well. It's going to be hard to explain that in a report and so forth. There's nothing to show that it's superior to the current method.

Moreover, as a practical matter it's a very small minority of examiners who use the method. I don't know of anybody who uses it in isolation. In other words, typically what's done is you're doing a standard examination coming to a conclusion and asking yourself, "Hum. Maybe I should count them just to say as a corroborating factor."

What would happen if that method were pitted across the traditional method in a head-to-head testing with all of those performance measures? I don't know but I suspect it would be very, very close. It's conceivable you'd have a slightly higher true positive rate with the traditional method and a slightly lower false positive rate. I don't know. I'm speculating.

Q. You're still using the comparison microscope as your instrument, correct?

A. Yes.

11:58AM Q. If I understand it correctly, what CMS is doing
11:58AM or is attempting to do at least is quantify the findings
11:58AM of the examiner.
11:58AM A. Not exactly. That's what I wish to be done and
11:58AM that's what was in my paper about creating likelihood
11:59AM ratios from this data. The way it's done now is you
11:59AM simply are counting consecutive matching striations and
11:59AM drawing a conclusion of is this identity or is this
11:59AM inconclusive.
11:59AM The other thing I might mention on one of
11:59AM the criticisms of it, it's not truly an objective
11:59AM technique, just like facial recognition software is not
11:59AM totally objective either. You have to use images. Is
11:59AM the photograph a little off.
11:59AM If you are putting into a system bullets,
11:59AM for example, you're imaging those, too. What's the
11:59AM competence of the person doing the imaging. In the case
11:59AM of CMS you're having to count striations. The
11:59AM subjective comes into how many do you count as to how
11:59AM many somebody else is counting.
11:59AM Through training it could get possibly
11:59AM fairly consistent. I don't know. I have not had that
11:59AM training. I know there's a subjective element there.
12:00PM Q. There aren't a lot of examiners doing the CMS
12:00PM or using that procedure at this point?

A. SWGGUN did a survey on that issue and the results came back that it's a very small group that uses it. And I don't know if there's anybody that uses it exclusively and not using the subjective technique. It's possible but if there are, they are very few in number.

Q. It's certainly not something that AFTE is requiring its members or even recommending at this point?

A. Oh, no, no.

Q. There's been an allegation made that because of modern manufacturing techniques individual marks are not being left on weapons and therefore transferred to the fired components, bullets or cartridge cases. Have you heard that argument?

A. Yeah, I've heard that argument. I don't really put a lot of stock into it. However, there needs to be continuing research done. Very often this is done by trainees who go out to the factories, collected consecutive specimens and so forth. It constantly needs to be checked but I don't think fundamentally there's going to be a big difference.

One of the reasons cited is, well, our cutting tools are much harder than they were in the past. They don't have to be sharpened nearly as often.

12:01PM This assumes it's the wear on the cutting tool causing the marks. That's not clear at all.

It could be suspended particles in the lubricant. It could be chip formation. There was one article where the owner of the factory speculated there's little microwelding that takes place on the cutting edge. This was an example where a cutting tool wasn't changed but rarely and yet from specimen to specimen, workpiece to workpiece the marks looked completely different. He was trying to figure out why.

We don't know exactly why. The vast majority of cases they tend to be different and it could be from, as I was just speculating, suspended particles in lubricant, microwelding, chip formation, what-have-you.

Q. Have there been any studies, validation studies or any kind of tests done to your knowledge to indicate -- or published research to indicate that tests have been done on guns manufactured under modern techniques and that they don't leave these marks?

A. Well, there have been -- validity tests and a lot of the proficiency tests are -- use new or fairly new firearms or tools. If your gun -- the problem you're referring to subclass markings. It's something we look at very carefully. If it's happening to a

12:03PM 1 significant degree we would see it in the error rates in
12:03PM 2 the proficiency tests and validity tests.
12:03PM 3 There are good reasons why -- although we
12:03PM 4 were very concerned scientifically about subclass marks,
12:03PM 5 there are good reasons to believe as a practical matter
12:03PM 6 it's not that serious a problem at all. I can speak to
12:03PM 7 that if you want or we can move on.
12:03PM 8 Q. I want to make sure that my question was clear.
12:03PM 9 It's a little difficult to word. Are you aware of any
12:03PM 10 research that proves the opposite of your professional
12:03PM 11 opinion and that is that modern manufacturing techniques
12:03PM 12 in fact have done away with the ability to individualize
12:03PM 13 firearms?
12:03PM 14 A. Oh, no. I would deny that vigorously.
12:03PM 15 Q. There are people out there making these
12:03PM 16 allegations but they have done no research to back it
12:03PM 17 up?
12:03PM 18 A. Yes. That's a situation where it's easy to
12:03PM 19 come up with theoretical reasons. Let's see the
12:04PM 20 empirical data. The data doesn't show that.
12:04PM 21 Q. Are you familiar with the National Academy of
12:04PM 22 Sciences' fairly recent report?
12:04PM 23 A. I'm familiar with it. It's been a little while
12:04PM 24 since I read through it completely but, yes, I'm roughly
12:04PM 25 familiar with it.

12:04PM Q. Are you familiar with the committee, the group of scientists who put that report out?

12:04PM A. I remember reading who they were and, as I recall, there were quite a few statisticians and some lawyers and some chemists but I don't recall exactly.

12:04PM Q. Were there any firearm examiners on there?

12:04PM A. I don't believe so, no.

12:04PM Q. Your understanding of the report, was there ever a recommendation made that we stop relying on firearms evidence or firearms examination?

12:04PM A. No, not at all. As it happens, I actually agree with the vast majority of what was contained in that report. I think it was a well-done report. There was a couple little things where I disagree. One is I don't think they took into account all the studies that's been done.

12:05PM Reading through it, it's like they weren't aware of the validity research that's been done. I think they were maybe looking at the presumptive check type of research which admittedly, while valuable, is not as valuable as the gold standard validity tests.

12:05PM It's like comparing a medicine in a randomly controlled trial to an observational study. I like to tell people that doesn't mean an observational study in medicine were useless. If it were, we would

all be still smoking cigarettes. An observational study showed cigarettes are risky, not a randomly controlled trial.

Presumptive checks that have been done are not useless but they're not as overall high quality and tell us as much as the validity tests do. The other thing is one of the problems I had with it is they had no one on there that I recall that was a scholar or scientist of expert judgment.

The sciences are vastly different. If you are talking to a physicist on the one hand or a chemist or a DNA analyst you're following specific protocols. A lot of times you're dealing with very homogeneous materials. That's different from a radiologist interpreting films or digital images.

It's different than a paleontologist who goes in the field and comes across a fossil and says that's a such and such animal from such and such period. Is that scientific? Maybe, maybe not. If he had been tested on his accuracy, yes, it is. And they didn't have those kinds of experts. They tended to think in the DNA paradigm that, well, if it's good science, then you have this, this and this.

I disagree with that. Very often -- this is not recognized a lot by laypersons. Very often

12:07PM science is a little more squishy than what people think. You have a lot of disagreements. You have parochialism. Physicist who looks at geology might say that's not a science. You're not doing high mathematics.

So they're tending to look at it through a certain kind of lens and I think ideally they would have had people on that panel who were experts in expert judgment.

Q. My understanding -- correct me if I'm wrong. The NAS report, their criticism of the field was generally we need more research. We would like better numbers but certainly nothing to say that it can't be relied upon or the work that you guys have been doing in the field is not reliable?

A. I agree with that. I don't think they took into account or saw some of the research, frankly. They did not say it was invalid or anything like that. They did say, and I agree with them, it would be very good if there was more and more research. I would like to see more and more research where we can narrow down the boundaries for error rates, for example.

Q. This particular case you did not examine the evidence; is that right?

A. That's correct.

Q. In fact you haven't even looked at Susan

Allen's report; is that right?

A. That's correct.

Q. Or the photomicrographs that she took?

A. Correct.

Q. You're not familiar with her lab or her qualifications; is that fair?

A. Fair.

Q. But you can state that the gold standard equipment is a comparison microscope?

A. Yes.

Q. Where the test fired piece of evidence, whether it's a cartridge case or a bullet, is compared to an unknown cartridge case or bullet?

A. Yes.

Q. And using subjective -- mostly subjective experience an examiner is going to come to a conclusion based on that examination?

A. Correct.

Q. Assuming that a lab follows certain protocol and record keeping and maintenance of evidence and marking of evidence, certainly you would expect those types of things to be done for quality control?

A. Sure, yes.

MS. MOSELEY: I'll pass the witness.

THE COURT: Let's take a one-hour lunch

break.

(The noon recess was taken after which the following proceedings were had:)

THE COURT: The State has passed the witness; is that correct?

MS. MOSELEY: Yes, Your Honor.

THE COURT: You may proceed.

CROSS-EXAMINATION

BY MR. ANTON:

Q. When you concluded your direct examination you were talking about the NAS studies. Let's start from that. You understood the purpose of the NAS study was for the academy to examine a number of disciplines to see if they were truly a science or if they were practicing a scientific method. Among the disciplines examined was firearm and toolmark examination. You understood that?

A. Yes.

Q. You understood -- you don't have any -- well, let me say the National Academy of Sciences, you understand that's a congressionally enacted group of scientists, many of whom have Nobel prizes. They're considered outstanding members of their field, correct?

A. Absolutely.

Q. Whether or not they are ballistics examiners,

they understand or do you agree that they understand the scientific process?

A. In general absolutely. It's not clear if they totally understood -- I think generally they did, the processes we use. Every science has a different method.

Q. The science we're talking about here, toolmarks, they described as toolmarks are generated when a hard object, a tool, comes into contact with a relatively softer object, right?

A. Yes.

Q. Now then, talking about science, would you agree with the definition that I posted that scientific knowledge equals scientific method or methodology. A conclusion will qualify as scientific knowledge if the proponent can demonstrate that it is the product of sound scientific methodology derived from scientific method.

A. Yes and no. I don't believe and most philosophers of science and sociologists of science and historians of science if I read them correctly would deny there is a scientific method.

Q. You disagree with the National Academy of Sciences' definition of what science is?

A. Yeah.

Q. Okay. The National Academy of Sciences when it

comes to defining science they're not as correct as you are?

A. It's not me. It's philosophers of science, sociologists of science, historians of science had a little different take on it when they actually studied what scientists do compared to what inside scientists do.

Q. I posted here the Daubert factors. You reviewed those. They're empirical testing. The theory or technique must be falsifiable, refutable and testable; B, subjected to peer review and publication; C, known or potential error rate; D, the existence and maintenance of standards and controls concerning its operation; and E, the degree to which the theory and technique is generally accepted by a relevant scientific community.

Do you agree with that?

A. Yes.

Q. Let's start out with a discussion of what is the relevant scientific community. If toolmarks are basically the examination of the surfaces when a hard object meets a soft object, then you would agree that mechanics and engineering regarding surface examination, that would be a relevant science, wouldn't it?

A. It could be, depending on what the conclusions

are, yes.

Q. So it's not science unless you agree with their conclusions?

A. No. I'm saying let's say take a metallurgist, for example. If a metallurgist is saying, "I think there's a lot of subclass marks on this. It only stands to reason. I think from the tribology of it, that is, the study of the lubricants and so forth I think these things in my experience have happened," that's all well and good.

I don't have a quarrel with that. What I have a quarrel with is a rational or kind of a unaided reason kinds of statements have not been empirically tested. They're coming up with things that haven't been tested and whereas we have.

Q. Well, to go back to my question, you agree that there is a science for examining marks left on a surface?

A. I would agree there is engineering fields for, for example, roughness on surfaces and so forth, yes, not necessarily -- that's not what we do.

Q. I understand. There's a science in what you do. I'm going to get into that. You also would agree for people that examine things under a microscope there's a science of microscopy, isn't there?

A. I don't know if there's a science of microscopy. There's techniques and practices of microscopy. I don't know if I call it a science.

Q. One of the things I thought I understood you to say when you are critiquing the National Academy of Sciences is that you thought it was unfair for them to apply something as rigid as, say, DNA type standards to your toolmarks examination because your field was a little more squishy.

A. I wasn't saying my field was squishy. I was saying science in the large is a little more squishy than what the popular imagination would have.

Q. Another criticism is that they have people that were doing the critique for the National Academy of Sciences, people like chemists and so on that, you know, you didn't feel was proper necessarily for them to be reviewing your work.

A. No, no. I didn't say that. I simply said they tend to use -- in fact in my view anybody can be a critic. I'm not the kind that's going to use an ad hominem attack about somebody simply because they're not in the field. No, I'm not going to say that.

My point was there's a tendency for people in various fields to think certain ways. Think of it as tunnel vision if you will and their way of thinking

about their science tends to reflect on what they think of other sciences or techniques.

Q. What is it you're saying about the people -- National Academy of Sciences get people that were looking at you from like chemistry backgrounds and you didn't think that was appropriate or it could have been better?

A. I think all kinds of backgrounds on the National Academy of Sciences board would have been good. But I think that should have included not just people familiar with DNA or chemistry. I think it should have included people who are scholars and scientists and study expert judgment.

That would have rounded out the committee so they could better judge all kinds of fields that use expert judgment such as ours and radiology and so forth.

Q. I'm not clear on this. There is a scientific field for expert judgment?

A. I don't know if it's a scientific field or scholarly field. I think there's been a lot of studies, interdisciplinary kinds and others, in recent years in the last decade or two on expert judgment, how do experts make their judgments and so forth.

Q. In theory the National Academy of Sciences was trying to ascertain if toolmarks examination was really

an objective science or it was fairly subjective, correct?

A. I don't know if they were trying to determine that. However, I certainly would not deny that it has definite subjective elements.

Q. What's the difference between being subjective and squishy?

A. Well, when I said "squishy," for sciences I mean by that a lot of times in the popular imagination people think, "Scientists said it. Therefore, it's an exact science.

Therefore, it's perfect," et cetera. There's a lot of debates within a lot of sciences. There is a lot of non-consensus on a lot of issues and in even the very hard sciences. They don't come to conclusions necessarily strictly on the basis of experiment or observation.

Believe it or not, the general consensus may come on the basis of authority. It's not this thing where you follow a cookbook scientific method and you get the answer and that's the end of it. It's not anything like that.

Q. Your primary complaint of the National Academy of Sciences' critique of your field is that you think they were using the wrong cookbook?

A. No, not at all. Well, yes, in certain terms, yes. I think their perspective was more from a DNA perspective. Had they had a paleontologist on the panel, radiologist, those areas that were using more subjective interpretations, I think that will help round it out and give it a little more balance.

I think from my impression of reading the report they did not look very carefully at all at the studies that have been done. They might have been -- they are under time pressures. They might have poo-poo'd some of the presumptive checks.

Q. The National Academy of Sciences, all of those Nobel prize winners, they didn't do it very well?

A. I don't know if there were any Nobel prize winners on this panel. If you know that, that's news to me.

Q. I think your criticism about subjectivity in terms of CMS, you said one of the reasons you don't like to -- you don't think it's necessary to use that kind of methodology is because it's just too subjective.

A. I'm sorry. Say that --

Q. At the end of your direct examination you were visiting with the prosecutor after you finished your lengthy PowerPoint presentation. You were discussing various aspects of matters not contained in the

PowerPoint. Do you recall that?

A. I recall a discussion about comparison with DNA and theoretical studies.

Q. Do you remember talking about CMS?

A. To the prosecutor?

Q. Yeah, yeah. Consecutive matching striation. Do you remember having that conversation and she asked you if it was necessary in your science -- if you thought it was necessary for a ballistic examiner to use CMS and you said no. That's really a pretty subjective science.

A. I think what my point was, it's a lot more subjective than what people realize in counting striations. If you are -- someone is trying to assert it's a purely objective method, it's not.

Q. How is it more subjective than the eyeballing comparisons you do in ballistics? What's the difference?

A. I'm not saying it is more subjective.

Q. They're both subjective?

A. Yes.

Q. Now, in terms of your radiology example, a person is a radiologist, they have got a degree in radiology, correct?

A. They have a M. D. degree and residency in

radiology.

Q. In terms of your association, AFTE, is it required that all -- anybody have a scientific degree?

A. I don't think it's required to be a member, no. If you are going to come into the field in the present day I think it is a strong recommendation. I think most laboratories require a science or engineering degree, yes.

Q. Science, including like chemistry?

A. Sure.

Q. Just any kind of scientific degree, nothing that necessarily relates to toolmark examination?

A. Correct.

Q. I think your point with the radiology analogy was to point out that you eyeball ballistics. You see the comparison much the same way the radiologists eyeball the reports and see whether or not there's an imperfection in the spine.

A. My point was the way they learn how to read or interpret images is very similar in principle or even identical in principle to the way we learn how to recognize sufficiently agreeing correspondence, yes.

Q. Of course, I think your primary point on direct examination was that your organization eyeballs these ballistics comparisons, that they're basically -- it's

virtually no error. You published all of those zero error points for us. When you guys eyeball something it's just virtually no error?

A. Well, how do you define "virtually no error"? What I would go to is the data and let it speak for itself.

Q. In radiology do you think there's that degree of concurrence? Do you think radiologists debate over the proper interpretation of a slide?

A. Sure.

Q. Your science is more exact and you have a better rate for eyeballing things than radiologists?

A. Let me put it this way. I recently read a book written in 2007 by Dr. Jerome Groopman, a Harvard physician. He wrote a book called How Doctors Think. There's a chapter on radiology, very informative chapter. And from reading that chapter I would conclude that, yes, our discipline is on average reaching more accurate results.

Q. Let's go through these five Daubert factors. Empirical testing. Do you agree or disagree that AFTE and its members have never provided any documented empirical testing where it is clearly shown the theory or technique must be or is falsifiable, refutable and testable? You agree with that?

A. Let me read it. Theory or technique must be falsifiable, refutable and testable. If I understand the meaning correctly, is that saying that there have not been any empirical testing to check for falsifiable --

Q. That's right.

A. No, I disagree with that.

Q. The National Academy of Sciences when they review your field, they made the observation that much forensic evidence, including toolmark identification, is introduced in criminal trials without any meaningful scientific validation determination of error rates or reliability testing to explain the limits of the discipline. Do you agree with that?

A. No. That's where I was pointing out in my view that was one of the areas where I disagreed with the NAS. In sum total I fundamentally agree with most of everything they say but I think they did not look at our validity studies and on the proficiency tests.

Q. We'll return to that in a little bit. The National Research Council on ballistic imaging suggests additional general research on the uniqueness and reproducibility of firearms-related toolmarks would have to be done if the basic premises of firearm identification are to be put on a more solid scientific

01:29PM 1 footing. You see that comment?
01:29PM 2 A. Yes.
01:29PM 3 Q. Do you disagree with that, too?
01:29PM 4 A. Not entirely. I think to put it on a more
01:29PM 5 solid scientific footing, I would like to see more
01:20PM 6 studies done, yes.
01:29PM 7 Q. One of the things that I noted in your direct
01:29PM 8 examination, while we're on this topic of testing, is
01:29PM 9 that you said that there are critics which you were
01:29PM 10 fairly dismissive of who have raised questions about
01:29PM 11 your testing procedures. But I think that your point
01:29PM 12 was they haven't really produced studies that dispute
01:29PM 13 your opinions. So you're not going to consider them at
01:29PM 14 this time; is that fair?
01:29PM 15 A. No, that's not fair. What I meant to say is
01:29PM 16 I'm not aware of any validity studies or proficiency
01:30PM 17 test results of the kind that we were doing which were
01:30PM 18 directly testing the science, directly testing our
01:30PM 19 propositions. I'm not aware of anybody having done
01:30PM 20 those that came up with serious error rates.
01:30PM 21 Q. In terms of proving that something is a science
01:30PM 22 and there is empirical testing regarding its
01:30PM 23 reliability, is that the job of the proponent of the
01:30PM 24 science to prove or is that the job of the critic?
01:30PM 25 A. Both.

Q. So since this time, since the National Academy of Sciences has come out with their report suggesting other methods and techniques should be considered how many of those have been implemented?

A. I'm sorry. Could you repeat that, the what recommendation?

Q. The National Academy of Sciences, I think, made some recommendations, did they not, their conclusion, about the types of testing that should be done or would be preferable? Do you recall that?

A. I would like to see that. I don't recall specifically what they said or recommended.

Q. Okay. I can read it to you or have you read it if you want. Overall the process --

MS. MOSELEY: Where are you reading from?

MR. ANTON: Page 5-21 of the report which was previously introduced in evidence.

Q. (By Mr. Anton) You can just read that summary assessment.

A. Okay. Okay. I kind of recall seeing this, yes.

Q. All right. Well, what steps have you taken to implement any of that?

A. First of all, they're talking about a lack of a precisely defined process. That's what I was talking

about when I am saying what they're looking for is a duplicate of DNA. There's a tendency to say it's not done the way DNA does it, it's not any good. There's a little bit of that here in the second paragraph.

They're basically talking about the subjectivity. Well, the scientists that were on that committee aren't used to that. They're not used to something like radiologists, paleontologists and so forth. The second paragraph they report Hamby, Brundage and Thorpe. I think they're not looking at the validity studies that we did within the FBI. They're looking at maybe some of the presumptive checks that I was talking about.

Q. Okay. Basically the National Academy of Sciences' critique of your field is basically these scientists didn't do enough research in your field, they're lazy and didn't take into account all the things you thought needed to be considered?

A. I think they could have done -- in that respect I think they could have done a better job.

Q. The National Academy of Sciences?

A. That's correct.

Q. You're familiar with Rule 702 posted up there?

A. Go ahead.

Q. Rule 702 I think you recognize incapsulates

some of the standards of the Daubert test but it provides for testimony essentially if the witness' insight will assist the jury as a finder of fact. Do you see that?

A. Yes.

Q. You understand there's a difference between expert testimony that will assist the jury and science. You know there's a distinction between the two, right?

A. You might want to fill me in on that to remind me.

Q. I could have Ms. Barnhill, the court reporter, explain to the jury what she's doing. They may not know how her machine works and would assist them in their understanding. I don't know that I would say she's in the science of court reporting.

A. Okay.

Q. Okay. You understand that for something to be called science it's supposed to meet the Daubert factors?

A. Right.

Q. Just because it doesn't meet Daubert doesn't mean it's not admissible. You understand if it doesn't meet Daubert it's not considered science?

A. It's not considered science by the Daubert justices, yes.

Q. I guess you're saying -- is the Daubert test wrong?

A. I'm not disagreeing with you.

Q. In terms of the first criteria for Daubert, the empirical testing, the National Academy of Sciences, National Research Council concluded that really you don't have empirical testing to qualify at this time as a science. Not that your testimony may not be helpful to the jury but you don't qualify as a science. You understand that?

A. I understand that's what they're asserting or may be what they're asserting, yes.

Q. You disagree with that?

A. Yes.

Q. National Academy of Sciences is wrong. Okay. This is what the National Research Council also stated. "Because not enough" --

THE COURT: Read slowly, please.

MR. ANTON: Yes. I know Bridget will throw something at me if I don't.

Q. (By Mr. Anton) "Because not enough is known about the variabilities among individual tools and guns, we are not able to specify how many points of similarity are necessary for a given level of confidence in the result."

"Sufficient studies have not been done to understand the reliability and repeatability of the methods. The committee agrees that class characteristics are helpful in narrowing the pool of tools that may have left a distinctive mark."

"Individual patterns from manufacture or from wear might in some cases be distinctive enough to suggest one particular source but additional studies should be performed to make the process of individualization more precise and repeatable."

You see that conclusion?

A. Yes.

Q. They're saying you haven't done enough studies.

A. May I comment on it?

Q. Sure.

A. It's the same thing I said before. Many of those on that committee were statisticians. They're used to doing things a certain way. That's why you get words like "repeatability, more precise, given level of confidence."

All of that is good and if you are going to have something like DNA, all of that is outstanding. That's the top of the bar. What would they say about paleontology? What would they say about radiology? By going by some of the things they say there's a lot of

stuff you're going to throw out as useless or unscientific.

Q. You're -- I think that's exactly the point. Your point here -- I think that you have been characterized as primarily the person that's been pushing back against the NAS report; is that fair?

A. No. The vast majority of the stuff in that report I agree with.

Q. You're the one that prepared the seminar for members of your field about how to withstand Daubert challenges?

A. That was long before the NAS report.

Q. So you're trying to tell the Court that even though you don't -- that the National Academy of Sciences' definition of science, you don't meet it, you should still be considered a science anyway?

A. I think had they looked into this more carefully of the studies done and had they had a more balanced panel, they wouldn't have come up with quite these harsh conclusions. They would have looked at the studies I showed you earlier and said, "It would be nice if there were more of them."

I would certainly argue that but those are testing directly the propositions. Most philosophers of science, sociologists, historians are going to say,

"What is it that really distinguishes a science from a non-science?" If there's any consensus at all on that, it would be what you talked about earlier, testability. Karl Popper's criteria was falsify reliability.

No one uses that anymore. Testability, yes; empirical testing. That's what those validity studies did. That's what the proficiency studies do. Though they're not designed to do that, they offer information on that, too.

Q. We will visit the testing field in a few minutes. Now, you characterize in terms of it's not science, it's useless. You know that's not what I said. I said it's still admissible possibly under 702 but it's not science. It's either science or it's not.

A. The Daubert justices have said here's how you determine that.

Q. Okay.

A. Separate from that outside the legal context, that's not what philosophers of science would say.

Q. The National Academy of Sciences, those scientists that Congress has seen fit to review the scientific fields, they have required, they have suggested that you don't have enough empirical testing but they're not qualified to define what science is.

A. I would say the best people to define what

01:39PM science is are not scientists but are philosophers of
01:39PM science and sociologists of science and historians of
01:39PM science, especially philosophers of science who it is
01:39PM their job to do that.
01:39PM Q. Basically you're saying that -- if we tried to
01:39PM bring a critique from somebody not a member of your
01:39PM organization that really that that -- you have to
01:39PM dismiss or discount their critique because they're not
01:39PM part of the team?
01:39PM A. No, I disagree with that. Can I explain?
01:39PM Q. Yeah.
01:39PM A. What I mean by that, I don't think that's a
01:39PM fair statement. You need to invite criticism from all
01:39PM quarters. Sometimes insiders are going to be the best
01:39PM critics but sometimes outsiders can see things insiders
01:40PM don't and they will have very good critical points. It
01:40PM just depends on the specifics.
01:40PM Q. Let's talk about that. You would agree that in
01:40PM discussion, these attacks, these Daubert challenges
01:40PM about the AFTE, that the courts have found that the
01:40PM lawyers that want to capitalize on the NAS report to
01:40PM challenge the disciplines including toolmark analysis
01:40PM need to look outside the field of forensics and identify
01:40PM potential experts in the academic community because you
01:40PM can't expect crime laboratory employees to help defense

attorneys challenge -- they won't help lawyers challenge methods that they and their colleagues use. That's right, isn't it?

A. I think there's a certain element of truth to that.

Q. When we try to bring an expert from outside the field to level a critique against it, suggest there might be some ways to improve it, are you going to accept or dismiss him because he's not part of the team?

A. I'm not going to dismiss anybody's critic out of hand, nor am I going to dismiss them simply because they have got some ad hominem attack. I think anything they say you have to look at the substance. I would remind you that I don't work for a crime lab.

Q. In terms of your comment about testing, the National Academy of Sciences addressed those tests that you were submitted. I think one of them specifically was the -- I think there were three authors. There's Brundage, Hamby and somebody else submitted to the National Academy of Sciences.

You talk about these validation studies that you have done. Do you recall what the National Academy of Sciences said about whether or not that's an appropriate critique, appropriate method of establishing empirical testability?

A. Didn't you just have that up on the slide?

Q. I have got it up there right now.

A. There was another one where I think it was talking about that study.

Q. What you are talking about -- we can look on the web and go to the SWGGUN website. We can see these capsule summaries, right?

A. I guess.

Q. The capsule summaries suggest a heavy reliance on the subjective findings of examiners rather than on the rigorous quantification and analysis of sources of variability. You agree with that statement?

A. The capsule summaries -- yes, I agree with that. Note what are they looking for again when they say analysis of sources of variability. They're looking at it in a DNA way, saying what would be ideal. You go in and look at all the variations on these striations, et cetera, et cetera. You measure them and then you compare the variations are the same, et cetera, et cetera.

Well, if that can be done that would be great. That would be similar to a radiologist saying, "I'm going to take this MRI and start measuring it and check the gray scale intensity and I won't say anything until I do all of that." That's ridiculous.

01:43PM Q. Basically those standards are too rigorous for you?

01:43PM A. They're not too rigorous. It's not something that lends itself, at least not yet, to this kind of analysis. Now, that's coming and there has been some work done in that area with machines using perfectly objective algorithms and measuring devices in all kinds of ways to come up with objective outcomes. That research has started and it's continuing.

01:43PM Q. So the Court will understand, make sure I understand, you had in your PowerPoint presentation a number of side by side images of toolmarks and you eyeball those and say a competent examiner would see that these two side by side comparisons match. You recall that?

01:44PM A. Yes.

01:44PM Q. When you have these studies what you have done is taken those same two images and passed them around to a bunch of other people in your field and you get the reports back and they all say they match?

01:44PM A. No. There were no images sent around. They had to look at the exact specimens, test specimens.

01:44PM Q. Now, if I understand, that is the method, right? You pass these specified pieces of evidence around. People look at them, eyeball them. They report

back and say, "Look the same to me"?

A. That's certainly not the way -- I can speak better to the ones done in the FBI. Each examiner got its own test packet that no one else looked at.

Q. Do you agree with the following critique by the National Research Council? "Most of these studies are limited in scale and have been conducted by firearms examiners and examiners in training in state and local law enforcement laboratories as adjuncts to their regular casework."

"The report concluded the validity of the fundamental assumptions of uniqueness and reproducibility of firearms-related toolmarks has not yet been fully demonstrated."

A. I partially agree with that. Depends on how far you go -- could you go back there just a second, please? Hasn't been fully demonstrated. It's a question of how high do you set the bar? If you set it really, really stringently, yeah, I agree with that.

Q. You don't think it's the appropriate place for the National Academy of Sciences to be the one that sets the bar?

A. No, I wouldn't say that necessarily. I think for the National Academy of Sciences to do a really fair job, there should have been other people on the

committee and they should have looked more fully at all the research.

I understand why that doesn't happen. There's not unlimited time or resources. But there were no firearms examiners on that committee which is not a terrible thing but at least they needed to be consulting a lot with them and they needed to be aware of all the studies done, not just some of them.

Q. You wrote back the National Academy of Sciences and you said, "Look. We have got all of these additional studies"?

A. I did not write back to them and say anything.

Q. You're saying, "They just didn't give us a fair shot," but really you haven't done anything to change their mind.

A. I think what they're saying -- they're not saying what we're doing is invalid. They're saying we need more work. I'm not disagreeing with that.

Q. And they're not only saying you need more work but they're saying until you do the work, although your opinions might be helpful and there might be a basis for eyeballing a comparison, it's not really science yet. Isn't that what they're saying?

A. I don't know if that's what they're saying or not. Whether or not they are or not, I would disagree