TRIAL COURT CAUSE NO. F08-45280

THE STATE OF TEXAS	)	IN THE 283rd JUDICIAL
	)	
VS.	)	DISTRICT COURT OF
	)	
THAI-AN HUU NGUYEN	)	DALLAS COUNTY, TEXAS

EXCERPT FROM REPORTER'S RECORD

(Testimony of Stephen Bunch)

On January 24, 2011, came on to be heard before the HONORABLE RICK MAGNIS, Judge of the 283rd Judicial District Court of Dallas County, Texas, the above entitled and numbered cause.

Proceedings reported by computerized stenotype machine; Reporter's Record produced by computer-assisted transcription.

APPEARANCES

MS. ANDREA MOSELEY
MR. JUSTIN McCANTS
Assistant District Attorneys
Frank Crowley Courts Building
133 N. Industrial Blvd.
Dallas, Texas 75207
(214) 653-3600

FOR THE STATE OF TEXAS

MR. BRUCE ANTON
MR. WARD MAEDGEN
Attorneys at Law
Dallas County, Texas

FOR THE DEFENDANT

CHRONOLOGICAL INDEX

January 24, 2011

(The following excerpt was had on January 24, 2011:)

MS. MOSELEY: State calls Dr. Stephen Bunch.

THE COURT: Raise your right hand.

(The witness was sworn by the Court.)

THE COURT: Have a seat.

STEPHEN BUNCH,

was called as a witness by the State, having been first duly sworn, testified as follows:

DIRECT EXAMINATION

BY MS. MOSELEY:

Q. Good morning. Will you please state your name for the record and spell your last name?

A. Stephen Bunch, B U N C H.

Q. Is it Dr. Bunch?

A. Yes.

Q. Did you bring with you or had you provided me with a current CV of yours?

A. I believe I did, yes.

Q. I'm showing you what's marked for identification as State's Exhibit Number 3. Will you take a look at that and see if that's your CV?

A. Yes, it is.

MS. MOSELEY: For purposes of this hearing

I offer State's Exhibit Number 3.

MR. ANTON: We have no objection for purposes of this hearing.

THE COURT: It's admitted.

(State's Exhibit 3 admitted into evidence and is attached to this transcript.)

Q. (By Ms. Moseley) If you would, Dr. Bunch, would you just briefly go through some of your educational background?

A. Well, I obtained a bachelor degree in mechanical engineering from the University of Missouri, a masters in history from the University of Missouri and Ph.D. in history from the University of Illinois/Champaign-Urbana. Then a year or two after that ended up with the FBI laboratory and then -- do you want me to explain that training?

Q. Sure.

A. Started out as a physical science technician, then became an examiner trainee. Went through about a two-year training program involving visits to factories, ammunition-making factories, firearms factories. You go through a very long training syllabus doing a lot of practical and theoretical exercises.

Attended a lot of schools, many of them put on by the FBI, go through a series of oral board

examinations and moot court examinations and then you're qualified and start -- there's also a lot of hands-on work with other examiners. After the final qualification you work on casework.

Q. When did you say you started your work as a firearms examiner?

A. I believe this would have been 1999, if I am not mistaken, and then I was an examiner through about 2002, got promoted to the chief of the unit, firearm and toolmark unit there at the laboratory.

Q. At the FBI laboratory?

A. Correct.

Q. So how many firearms examiners were you responsible for supervising?

A. It varied. We were being downsized with an area between eight to maybe twelve, something around there.

Q. As part of your job as a supervisor did you participate in casework -- in performing casework analysis of the evidence yourself?

A. No. I did proficiency. I kept up with the proficiency tests. The only part -- the only role I would sometimes have in actual casework would be confirming identifications, occasionally doing -- well, I would also do administrative reviews of case reports,

10:18AM 1 case notes and the reports.
10:18AM 2 Q. Your experience as a firearms examiner, how
10:18AM 3 many cases or how many -- I don't know how you all
10:18AM 4 define casework but how many examinations have you
10:18AM 5 done -- comparisons have you done?
10:18AM 6 A. I've been asked that before. I can't give an
10:18AM 7 exact answer. I'm estimating somewhere between 500 and
10:18AM 8 a thousand actual comparisons. I don't know how many
10:18AM 9 cases that was, a hundred, 150. I'm not sure.
10:18AM 10 Q. Have you testified in court about your findings
10:18AM 11 as a firearms examiner?
10:18AM 12 A. Yes.
10:19AM 13 Q. On few or many occasions?
10:19AM 14 A. May have been about ten, something like that.
10:19AM 15 Q. Have you been qualified as an expert in the
10:19AM 16 past?
10:19AM 17 A. Yes.
10:19AM 18 Q. Have you ever testified in Texas?
10:19AM 19 A. I don't believe so.
10:19AM 20 Q. Have you testified in other state courts?
10:19AM 21 A. Yes.
10:19AM 22 Q. Have you testified in federal court?
10:19AM 23 A. Yes.
10:19AM 24 Q. I guess as an FBI examiner most of your
10:19AM 25 testimony has been given in federal court; is that fair?

10:19AM A. Probably. I'd have to reconstruct the record but that's probably a fair estimate, yes.

Q. When you were -- I don't know what you call it -- chief of the unit --

A. Yes.

Q. -- at the FBI lab -- you mentioned proficiency tests. Did you do any other sort of research or anything at the lab yourself?

A. Oh, yes. Starting in about 1999, the FBI was doing the laboratory -- was doing a review of the Daubert decision and what did the lab need to do to prepare for it and so forth. I was selected out of the unit to be a representative of our unit in a larger group to start preparations, what needed to be done.

At that point I started doing research of all kinds in trying to get our examiners trained to deal with Daubert hearings and also to look for what areas of research needed to be done to help bolster the science or if the science needed work, where did it need work, identify those areas.

Q. Did you prepare a PowerPoint that will help explain the science behind firearms examination and some of the research that you have done?

A. Yes. This originally started out as a PowerPoint many, many years ago. I provided it to the

10:20AM scientific working group for firearms identification. They modified it heavily as a template which is even available online. And then I took that again in recent years and heavily modified it to my own taste for courtroom purposes and that's what the outcome is.

Q. Do you think that having this PowerPoint presentation will assist us in understanding your testimony?

A. Yes.

Q. I'm showing you what's marked State's Exhibit Number 1. Does that appear to be a printed copy of the slides from your PowerPoint?

A. Yes, it does.

MS. MOSELEY: Your Honor, I have provided a copy to defense counsel. I offer State's Exhibit Number 1 for the record.

MR. ANTON: No objection for purposes of the record, Your Honor.

THE COURT: It's admitted for record purposes.

(State's Exhibit 1 admitted for the record and is attached to this transcript.)

MS. MOSELEY: I think we may be able to have him use the clicker.

Q. (By Ms. Moseley) If you would just explain --

start with your first slide and tell us best you can about the science behind firearms examination, what we're looking at and how we look at it.

A. Very well.

THE WITNESS: Your Honor, if this goes too slow or fast, please speed me up. There may be things you're familiar with.

A. Basically this is just a summary I wrote, brief summary of some of the issues and validity having to do with the science. Ultimately what this comes down to is the central question.

What does it mean when you get a definitive match or exclusion conclusion? How strong is that conclusion? How does it link up with the underlying scientific findings to result in the strength of a certain or certain strength of conclusion?

Those are possibilities, all the way ranging from absolute certainty down to negative probative value. It's worse than nothing. And so that's ultimately what we want to answer.

Before doing that, need to go through some background information. Trying to go through this quickly. Everybody may know this already. This is a cartridge case with the various parts: the powder, the casings.

10:23AM 1 You see the headstamp, primering mixture,
10:23AM 2 cup, marks put on by firearms whether during firing or
10:23AM 3 not on many portions of the cartridge case. Also
10:23AM 4 there's markings placed on the bullet which you see the
10:24AM 5 projectile flying out the end.
10:24AM 6 The barrel itself -- this is a graphic that
10:24AM 7 shows where we get what are called class
10:24AM 8 characteristics. This is a typical barrel. You can see
10:24AM 9 it's got spiral grooves cut into it in this case. And
10:24AM 10 they leave impressions on the bullets that have been
10:24AM 11 fired from that barrel and that helps us in our
10:24AM 12 examination and gives us information.
10:24AM 13 This one you can see -- if you think of
10:24AM 14 your fingers, goes in there, move your fingers around.
10:24AM 15 Your hand goes in a counterclockwise direction called a
10:24AM 16 left twist. You can count the number of them. You can
10:24AM 17 measure the width of them.
10:24AM 18 This is typical photographs of bullets that
10:24AM 19 have been fired and you can see that those lands and
10:25AM 20 grooves that were in the bore of the barrel leave
10:25AM 21 impressions on the bullets and we call those land and
10:25AM 22 groove impressions, the groove impression being left by
10:25AM 23 the groove in the barrel.
10:25AM 24 That's a closer view of a land impression.
10:25AM 25 You can see it has a definite width. You can see more

10:25AM fine markings on the interior of that impression. That's really just scratch marks. We call them striations.

What we're looking at when we do an examination of a bullet is both of these areas, if you want to call it this -- both these areas of evidence in a sense, both the rifling, the width, how many impressions there are, what direction they're going and then also the microscopic detail within that impression.

Q. (By Ms. Moseley) If I could interrupt. When you talk about the lands and grooves and the right twist or the left twist you mentioned class characteristics. Can you tell us what's significant about class characteristics? What does that tell you?

A. Yes, I can tell you.

Q. Have we not gotten there yet?

A. If I can -- now, this is a photomicrograph showing what we were looking at. Those are from different barrels. You can make out -- if I can -- I didn't hit right on it. There's a vertical line there. The left side is one bullet. On the right side is another bullet.

You can see the very strong microscopic correspondence between the striations of both. And this is a textbook example of what you might find from

bullets fired from the same barrel.

Q. How do we get that photograph? Are we taking a picture of them sitting on a desk?

A. This was undoubtedly -- I didn't take it. It's through a standard comparison microscope.

Q. Is the comparison microscope the standard tool used by firearms examiners in this country as well as all over the world?

A. Yes.

Q. Thank you.

A. You may have to do that one manually. The arrow stayed on there. I don't know.

Q. I'll get it for you.

A. This is an example of bullets fired from different barrels. They do have the same class characteristics. You see in this case at least the width of the land impressions are the same, the barrel on the left and the bullet on the right.

However, the microscopic detail is different. The correspondence between the striations are not very good at all. That's what you expect, virtually always find bullets from different barrels.

Looking at cartridge cases this time, these are cartridge cases fired in the same firearm. Those marks are from the edge of what's called the firing pin

aperture. Those are scratch marks. You can see the hairline. On the left is one specimen; on the right is another specimen.

You can see the significant correspondence between both sides. These were fired in the same firearm. Looks like I went backwards. These are cartridge cases fired in different guns. It so happens these two specific guns were made consecutively, one after the other.

We often use consecutively made firearms as a severe test in validity tests because if you are going to find similar marks, that's where you would most likely find them, from consecutively made parts, in this case breech faces.

Nevertheless, in this case, as in the vast majority of cases, the microscopic correspondence between one and the other is not very good. No firearms examiner would -- no competent firearms examiner is going to look at this through the microscope and say they're fired from the same gun.

We don't just do guns. We do tools, including drill bits. This is just a sample. This is textbook. I don't mean to imply you're going to find this kind of quality or quantity of marks in all identifications, in all matches.

10:29AM 1 This is an example of the marks produced by
10:29AM 2 the same drill bit. Left side corresponds extremely
10:29AM 3 well to the right side. Two different drill bit
10:29AM 4 impressions is what we're looking at.
10:29AM 5 On the other hand, these are two different
10:29AM 6 impressions made by consecutively made drill bits and
10:29AM 7 again the correspondence is not very good. That's
10:30AM 8 typical of what you will find from different drill bits,
10:30AM 9 different tools even if they have been made one after
10:30AM 10 the other.
10:30AM 11 Firearm and toolmark identification rests
10:30AM 12 on two fundamental propositions. These are testable
10:30AM 13 propositions. If anything is going to have scientific
10:30AM 14 status it needs to be empirically testable and tested.
10:30AM 15 Oops, went the wrong way again.
10:30AM 16 Proposition one. This is the principal
10:30AM 17 claim of the science. Macroscopic and microscopic marks
10:30AM 18 imparted to objects by different tools will rarely, if
10:31AM 19 ever, display agreement sufficient to lead a qualified
10:31AM 20 examiner to conclude the objects were marked by the same
10:31AM 21 tool. That is, a qualified examiner will rarely, if
10:31AM 22 ever, commit a false positive error or
10:31AM 23 misidentification.
10:31AM 24 Q. When you say "a qualified examiner" -- you
10:31AM 25 briefly went through your background. Do you have to

get a college degree in firearms examination or how do you become a qualified firearms examiner?

A. In this day and age you need a college degree in science or engineering. That may not apply to the past but ideally you do. And then you go through a training program out of a laboratory like I did and usually it's going to be a year and a half to two and a half years in length, yes.

Q. Some sort of an apprenticeship where you are working along with someone who has been doing it longer than you have and they teach you what to look at and what you are looking for?

A. There is the apprenticeship aspect to it definitely but there's theoretical work as well.

Q. Meaning what?

A. Course work. You go to various courses, read a lot of material. You do your own work. You do your own experimentation and so forth.

Q. Okay.

A. I'm getting the hang of this.

Proposition two is an explanation for why proposition one works. And that is most manufacturing processes involve the transfer of rapidly changing or random microscopic marks onto work pieces such as barrel bores, breech faces, firing pins, screwdriver blades and

10:32AM 1 the working surfaces of other common tools.
10:32AM 2 This is caused principally by tool wear and
10:32AM 3 abrasive chip formation or by electrical/chemical
10:33AM 4 erosion. Microscopic marks on both tools and objects
10:33AM 5 then continue to change from further wear, corrosion or
10:33AM 6 abuse.
10:33AM 7 Class characteristics, these are general
10:33AM 8 and/or measurable features of a specimen which indicate
10:33AM 9 a restricted group source. They result from design
10:33AM 10 factors and are therefore determined prior to
10:33AM 11 manufacture.
10:33AM 12 We will see examples of those. I mentioned
10:33AM 13 these before. You have got the barrel bore on the left
10:33AM 14 is going to leave large class characteristic marks on
10:33AM 15 the bullet on the right. You see of those impressions
10:33AM 16 there's a certain number of them. They have a certain
10:33AM 17 direction of twist and they have a certain width.
10:33AM 18 Those are the rifling characteristics. The
10:33AM 19 rifling characteristics are class characteristics.
10:34AM 20 There are others as well. The diameter of the bullet or
10:34AM 21 the caliber is also a class characteristic. All of
10:34AM 22 those together comprise what we call general rifling
10:34AM 23 characteristics. That's for bullets.
10:34AM 24 You might have to do that one.
10:34AM 25 The definition of individual

10:34AM characteristics is as follows: observed microscopic marks that are restricted to objects marked by the same tool. That is, they're individual to the source tool. These characteristics can be used to definitively link a tool to a toolmark and result from manufacturing processes, wear, corrosion or abuse.

These are examples of individual characteristics from manufacture, the very fine marks put on from the finishing process, the edge of a knife blade, for example.

Example of individual characteristics from wear, that's pretty self-explanatory.

Subclass characteristics. These lie between class and individual characteristics. They are microscopic features produced during manufacture that are consistent among some items fabricated by the same tool. These are not determined prior to manufacture and are more restrictive than class characteristics. We can talk about these more later.

As far as the examination process goes, actually fairly quite straightforward. At level one -- this is the first run-through. You're looking at class characteristics. You can easily in some cases eliminate or exclude specimens from having the same origin at this level.

10:35AM Those two bullet specimens there, the diameters are different. Even from that alone we can see they weren't fired from the same gun unless there are extraordinary circumstances of some kind of sub caliber device put in it or something.

10:36AM The direction of twist is different and I'm guessing the lands and groove impressions are also different. When you see things like this they're clearly incompatible class characteristics and the only justifiable conclusion is to exclude.

10:36AM Q. That would be an elimination?

10:36AM A. Same thing. However, if you can't exclude on that basis you go to level two analysis. You place a specimen on your comparison microscope and see what you see and draw your conclusions therefrom. It may or may not end up back a definitive match conclusion.

10:36AM The range of conclusions are just pretty much what I said. You have a definitive match, definitive exclusion. Usually they're called identification and elimination respectfully. And we have inconclusions.

10:36AM The easiest way to think of inconclusive is if someone puts a thumbprint on a pane of glass and smears it all around. You're not going to be able to tell anything about it. We have analogues here, too.

10:37AM Q. It may be the markings are not clear enough to convince the examiner that it is an identification but they can't really do an elimination either because class characteristics are the same and maybe it was fired from the same gun but we can't tell?

10:37AM A. That's correct. And we'll see examples of that coming up.

10:37AM As far as the identification goes, pretty much what you said, if the class characteristics present are the same and if the quality, quantity and character of the two toolmarks display sufficient microscopic agreement and compatibility, then an identification conclusion is often warranted.

10:37AM We saw this photograph before. This is an excellent example, textbook example but that's marks produced by the same drill bit. That's what we call an identification or definitive match.

10:37AM An inconclusive -- this is more what you were speaking to. Assuming similarity of class characteristics, if two toolmarks are insufficient in their microscopic agreement or incompatible in their character, then an examiner may be unable to conclude either identity or exclusion. In this event an inconclusive result is appropriate.

10:38AM And here's an example of this. I want to

10:38AM spend a little bit of time with this one. This is land impressions on two different bullets. So that the -- let's assume for the moment that you go around the entire bullet and the class characteristics are the same.

But as you can see from -- this is actually an outstanding photograph to illustrate this. If you look at the microscopic correspondence, it's very, very poor. No competent examiner would ever in a million years say this is from the same gun definitively. So it's an inconclusive.

However -- I'll get to this a little bit. This is something that's not commonly understood at all or even resisted. If there is a rifling characteristics match and we say inconclusive, the implication is that it's neutral. It's just as likely to have come from that gun as not. That's not true.

You have to think of this a little bit as the old serology examinations. If you have AB negative blood -- if a blood stain is AB negative and somebody else has got AB negative blood what does that mean? Well, depends on the frequency which AB negative blood is in the relative population.

Let's throw out a number and say two percent. Two percent of all people in the population

have AB negative blood. The stain is AB negative. The defendant has AB negative blood. That's not conclusive evidence by any means but it has probative value and we can measure that probative value.

Similar thing happens with class characteristics. The vast majority of guns do not have these class characteristics. We're restricting greatly the number of guns that could have fired these two bullets.

The difference with the bullets and the blood is that with the blood we know the frequency. We know pretty much the significance of that evidence, the probative value. We don't know exactly the class characteristics. We don't have a database that enables us to tell that.

We know, however, that it's not one-to-one. Even the most common of guns probably don't share class characteristics with any more than ten percent of all the guns out there, if that. When you are talking about rare class characteristics we may be talking about one in a thousand or 10,000.

We say "inconclusive" partly because the judicial process is innocent until proven guilty. That's proper. Inconclusive implies neutrality. It really is not.

10:41AM Q. Is it fair to say that we're going to err on the side of the inconclusive unless there is really no question in the examiner's mind based on their experience and their viewing of the evidence that in fact it is a conclusive match?

A. Absolutely. You have to remember these days firearms examiners don't just do examinations in criminal cases. They do it for the military. The military may have a different need. When we're talking in criminal courts we're talking innocent until proven guilty is fundamental.

If there's any doubt whatsoever every examiner, every qualified examiner knows you truncate it down to inconclusive. You don't want to have any doubt.

Q. We're not going to say that we exclude it unless the class characteristics are different?

A. In the vast majority of cases that's true. There can be extraordinary cases where you can exclude on individual marks. Generally speaking, that's not a good experience and generally speaking I don't think that's done.

As far as eliminations or exclusions go, I just basically said that. If significant disagreement exists in class characteristics, then an elimination conclusion is appropriate. This is a consensus

10:42AM standard.

10:42AM If disagreement in individual characteristics of an exceptional nature exists and there are -- an elimination conclusion may be appropriate. There is a lot of conditions you need for that and there have been errors made because that standard was not adhered to.

10:42AM This is an example of the kind of elimination or exclusion I'm talking about. These are clear-cut differences in class characteristics. The firing pin shape is different on one than the other and the -- very clear-cut rectangular firing pin aperture impression on the right side, none on the left.

10:43AM Q. And different manufacturers have a rectangular versus a round firing pin impression?

10:43AM A. It doesn't necessarily have to have different manufacturers but it can be but certainly it's a different gun that fired these. There's a lot -- this is where we're starting to get into the meat of it.

10:43AM There are further considerations in looking at the validity questions about the validity of the science. What's the standard for identification, objective versus subjective processes and the significance of conclusions.

10:43AM What I am going to read in the next two

slides will make your eyes roll. I'll read them and give a quick explanation. The theory of identification enables opinions of common origin to be made when unique surface contours of two toolmarks are in sufficient agreement. That's from the AFTE standard published in 1992.

Q. What is AFTE?

A. Association of Firearms and Toolmarks Examiners.

Q. Is that -- can you tell us more than -- like what is this group?

A. That's the main and perhaps the only professional organization for firearm and toolmark examiners.

I could have inserted a slide in there. It's a lot of verbiage that goes on to final conclusion is agreement -- is significant when it exceeds the best agreement demonstrated between toolmarks known to have been produced by different tools and is consistent with the agreement demonstrated by toolmarks known to have been produced by the same tool.

All that's really saying is that we're doing something similar to, say, what a dentist or radiologist does. I talked to my dentist about this to make sure. When he looks at an x-ray film of teeth to

10:44AM 1 decide whether it's tooth decay how does he know? How
10:45AM 2 did he learn that? How can he distinguish between a
10:45AM 3 healthy tooth and a decayed tooth?
10:45AM 4 The answer is in dental school he has to
10:45AM 5 look -- he or she has to look at a lot of images of
10:45AM 6 decayed teeth and a lot of images of healthy teeth and
10:45AM 7 he learns to distinguish between those two classes of
10:45AM 8 things.
10:45AM 9 That's what we're doing. We look at a lot
10:45AM 10 of known non-matches. We learn what kind of agreement
10:45AM 11 exists there. We look at a lot of known matches in
10:45AM 12 training and learn what kind of agreement exists there.
10:45AM 13 The two classes of things are distinct but
10:45AM 14 you readily learn to distinguish between those two
10:45AM 15 classes. It's not -- it's the same thing in radiology.
10:45AM 16 I had a bad back in the late 80s, went to the
10:45AM 17 orthopedist and took an MRI and I went back and he said,
10:46AM 18 "You have a ruptured disk." "How do you know?" "I'll
10:46AM 19 show you on the MRI."
10:46AM 20 I look at it. I couldn't tell anything.
10:46AM 21 If I had asked him, "How do you know that," he'd say, "I
10:46AM 22 know it when I see it." How does he know that? Because
10:46AM 23 he's looked at a lot of images of ruptured disks and a
10:46AM 24 lot of images of healthy structures in the back.
10:46AM 25 Does that mean his pronouncements were

perfect? No, not necessarily. But that's how the process is and how you learn how to do these things. By definition this is a controversial area to some extent.

Let me define the objective process. A fully objective process in theory is one that can be repeated by different scientists in different locations using identical methods and that yields consistent results. That is, in our case inter-examiner and intra-examiner variability should be low.

You're going to get high consistency or high precision, though validity and error rates, that is accuracy, may or may not be low.

We can think of a lot of examples here. College football, you can think of the various algorithm-based power rankings. Another example would be some of the newer facial recognition software. I saw a demonstration of this on television where they take a photograph -- find a photograph of a person.

You will measure between the pupils of the eyes. You will measure to the nose and the corners of the lips, all kinds of measurements across the face. Compare that to other photographs and you can get a scoring system for how likely it is to be the same person. That would be a completely -- not a completely but a relatively objective process.

10:47AM THE COURT: Excuse me. Mark your place. We'll take a ten-minute break.

11:03AM (A brief recess was taken.)

11:03AM Q. (By Ms. Moseley) Are you the same Stephen Bunch that was testifying before the break?

11:03AM A. Yes.

11:03AM Q. I think we had left off with where you were discussing the difference between a subjective process or an objective process.

11:03AM A. Right. You want me to continue?

11:03AM Q. You mentioned the facial recognition technology. I think that's what we were talking about.

11:04AM A. Yes. And how that would be an example of an objective technique as opposed to a subjective technique which I can talk about next.

11:04AM Subjective processes involve more human judgment. This does not mean that subjective procedures are unreliable, inaccurate or unscientific. Expert judgment, variability and error insinuate themselves into every science and technique, whether a physician diagnosing a disease, a DNA expert choosing the relevant population or radiologist interpreting MRI images.

11:04AM College football, think of the sports writers and coaches polls. Only by head-to-head testing and comparing performance measures is it possible to

11:04AM 1 weigh the relative superiority or inferiority of
11:04AM 2 competing objective and subjective methods.
11:04AM 3 These are generally recognized as the five
11:05AM 4 prongs of Daubert. I'm going to talk about these as
11:05AM 5 quickly as I can although I'll spend more time on error
11:05AM 6 rate and testability. The definition is the ability of
11:05AM 7 a theory or technique to be checked by experiment or
11:05AM 8 observation. These tests, if well-designed, tend either
11:05AM 9 to support or discredit the theory or technique.
11:05AM 10 Q. In the field of firearms is that testability
11:05AM 11 and the testing an on-going process?
11:05AM 12 A. Yes.
11:05AM 13 Q. And you mentioned earlier in 98 or 99 you
11:05AM 14 started doing research and testing your theories; is
11:05AM 15 that right?
11:05AM 16 A. I felt and others felt at that time that we
11:05AM 17 needed -- we definitely needed more rigorous testing of
11:05AM 18 the science. And since this is basically a science of
11:06AM 19 expert judgment the very best gold standard kinds of
11:06AM 20 tests to do involving all types of sources of errors
11:06AM 21 would be black box validity tests and those have been
11:06AM 22 done some. And proficiency tests are also of that type,
11:06AM 23 although they have disadvantages. I'll talk about those
11:06AM 24 coming up.
11:06AM 25 Q. Okay.

11:06AM A. Generally what evidence exists to support firearm and toolmark analysis? The short answer is numerous research studies and tests have been published in peer-reviewed journals over the past 50 years that support validity.

These studies have generally been of five types: presumptive checks, black box validity tests, black box proficiency tests, validation of consecutive matching striations, or CMS, and machine-based imaging systems.

For a selected listing of these articles they're online. You can to go to www.swggun.org under the "testability of the scientific principle" heading and find brief synopses of all these kinds of tests. That's -- as a side note, we have been striving increasingly over the years for transparency as a good scientific practice and a lot of the things we have done are on this site.

Q. So anybody can go to this website and find these validation studies and the different tests that have been conducted?

A. You can find references to the studies. You can't find the entire study but, yes, you can find reference to the studies and a short synopsis of the contents of the study.

11:07AM Q. You mentioned SWGGUN.org. What is SWGGUN?

11:07AM A. That's the standard for scientific working group for firearms and toolmarks identification. The "gun" part of it at the end doesn't quite fit in with the firearms examination but that's -- it's a professional organization, international. I was on that group for many years. It does excellent work and they review all kinds of matters related to scientific validity procedures, quality assurances and so forth.

11:08AM Q. Are most of the members of SWGGUN also members of AFTE?

11:08AM A. Yes.

11:08AM Q. Two professional organizations but not connected to one another?

11:08AM A. I'm not aware of any formal connection. Now, there may be a representative from AFTE sitting in on SWGGUN, something like that, but I don't think there's any formal connection. Don't quote me on that. I can't say for sure.

11:08AM The second one is what evidence exists to discredit firearm and toolmark analysis? The answer is no well-designed, discrediting empirical studies to my knowledge exist. Critics have advanced theoretical and practical arguments that point to possible weaknesses.

11:08AM We have to take into account those

criticisms. You don't -- would not just brush them off. Some with merit, a few. There have been no empirical studies to my knowledge that have been conducted that show high error rates or a fundamental lack of examiner skills or that refute the discipline's basic claims.

Q. When you say "to your knowledge" how long have you been studying the field of firearms examination, conducting studies yourself and reading articles and preparing for hearings such as this one in the courtroom as well as your professional knowledge base?

A. Since the fall of 1996.

Q. You're not aware of any formal studies that have been published in peer-reviewed articles or journals indicating that the science is faulty?

A. Well, there have been articles critical of the science but I'm talking about empirical studies where you are getting down into the weeds and asking examiners to make conclusions and finding high error rates or something of that nature. No, I'm not aware of any of that kind of thing existing.

One of the second tenets there, definition of general acceptance. Forgive me if this is old hat. The approval by a particular authoritative body of a technique or methodology. In addition to the forensic science community, you have got numerous colleges and

universities have courses in firearm and toolmark analysis, funding of scientific research in the area of firearm and toolmark identification. It's been accepted in court testimony for almost 90 years.

That latter part I put in there even though that frankly is irrelevant to the validity of the science.

These are some of the academic programs with firearm and toolmark curriculum. Usually in many cases it's a single course in a forensic science graduate program. There's quite a few.

Peer review and publication. Critical evaluation of research by qualified readers or referees as a screening procedure prior to publication in a journal. Depending on what philosopher or historian or sociology science you talk to, if there's such a thing as a scientific method, this is one of the things involved in a scientific method is peer-reviewed publication.

These exists in our field. The AFTE Journal. The American Academy of Forensic Sciences has the Journal of Forensic Sciences. The International Association of Identification has the Journal of Forensic Identification and there are others. There's Canadian journals, UK journals and so forth, all

peer-reviewed.

We kind of get into the nuts and bolts, the critical elements here. Known or potential error rates. I'm defining type one and two errors. That's a statistician term really for a false positive error type one. This is when -- in our case would be the identification of a toolmark to a tool when the toolmark was not produced by the tool.

Conversely, a false negative error type two error is the elimination of a toolmark as having been produced by a tool when the toolmark was produced by that tool. These are the fundamental critical errors in science of medical -- say, medical diagnostic testing and all kinds of testing. These are the fundamental errors.

Okay. These are the definitions. I don't want to cause anybody's eyes to glaze over but this is important. The false positive rate is simply in a test -- say, a validity test is simply the number of false positive conclusions divided by the number of true negatives.

That is, if we had 50 comparisons that we were doing on the microscope and all 50 were true negatives, they weren't from the same gun, that's what I mean by N. That's 50 -- N equal 50. If you say on one

of those times that they were from the same gun, that's one over 50. That's a two percent error rate.

By the same token, false negative rate is the total number of false negatives in a test divided by the total number of true positives. The failure rate or what the National Academy of Sciences may call a global error rate is combining the two at the bottom there.

Now, there are other test performance parameters besides just the error rate. These are used a lot of times in medical diagnostic testing. I was gratified to see the National Academy of Sciences came out in one of their reports and even suggested that these be used in validity testings and so forth. I had thought about this many years. I was very happy to see that and I would like to see more of it.

But they are true positive rate, also called sensitivity. You can read the definitions. Specificity, or the true negative rate or true negative results; positive predictive value, simply the total positive conclusions divided by the sum of the total positive conclusions and the false positive conclusions.

Negative predictive value. And then to lump some of these together there's something called -- you can say accuracy which is probably of lesser importance.

If you are talking about the criminal judicial system the most critical in my opinion would be the false positive error rate, false positive. That would tend to falsely incriminate someone. That's the biggest sin of all. It's also not good to have a false negative error.

Now, what are the possible sources of type one and type two errors in casework and in testing? There's basically four of them. There can be poor quality assurance, for example, confusing test with evidence samples, mislabeling identifiers on specimens, errors in transferring data from notes to reports, rushing through an examination and so forth.

Secondly, there's the application of the too liberal identification or elimination standard, usually stemming from inadequate training or incompetency on the part of the examiner.

Third. Subclass marks on specimens. We can talk about that more later but this is for various reasons probably quite rare in normal casework and only affect type one errors.

Fourth. The analogue to what DNA talks about, the random match probability. There's the possibility or the theoretical possibility of simply having two things match up, look very similar when they

were from different guns.

The example would be supposedly all snowflakes are different. Well, okay. But that doesn't really get you anywhere because who says? Theoretically if you have a database full of billions of images of snowflakes and the algorithm compares each snowflake to one another, you could theoretically have and probably would have some of them that look really close even though they're different.

And you might have a professional snowflake observer or a regular layman say, "Yes, they're the same," when they're not. What's the chances of that? Small but it's there. It's the random match probability for us, too.

Now, of all these kinds of errors which are the most serious? In my opinion number one is the most serious by far. While DNA won't talk about this, when they talk about these quadrillions and one in trillions and all that, that's a theoretical number based on population genetics. That doesn't account for laboratory error.

In our validity testing and our proficiency testing we account for that and that's deliberate on our part because we're a more subjective type of examination. I think number one is probably by far the

most likely source of error.

Number two could be, lack of training. Trainees might make these kind of mistakes. Going to be far less a problem, I think. Number three and four I think are perhaps orders of magnitude of lesser importance as a practical matter.

Q. We didn't talk about this and I may be jumping ahead but if we can just real briefly. As a firearm examiner if I come to you from a police agency and I say, "Here's a weapon, a nine millimeter handgun and here's ten cartridge cases and four bullets that I found at a crime scene," and I give those to you and say, "Tell me what you can tell me about them," what do you do, assuming you're the firearms examiner that's assigned to handle the work? What do you do with that stuff?

A. Short answer is you go through the protocol and do an examination. I mentioned before the level one and the level two. You do some preliminary work. You try to be very careful, make sure you label everything accurately.

Q. Do you label on the cartridge case? Is that part of the protocol?

A. I don't think the protocol specifies it has to be on the specimen. You don't want to be committing

errors. I forget what the current quality assurance guidelines are. At least the package that contains the specimen needs to be labeled. I mean it's a good practice to label the specimen it's if possible and it's not destructive. I think that's what the vast majority of examiners do.

Q. At some point did you test fire the weapon?

A. Yes.

Q. What's the purpose there?

A. You have to test fire the weapon to get control samples, test fire samples to compare to the evidence samples.

Q. And I assume the reason you want to be very careful with labeling the evidence that was brought to you is so that you are not confusing your test fires with the unknown samples.

A. Exactly.

Q. You know then you're comparing unknown to known?

A. Exactly.

Q. And then I assume that notes are to be made regarding your findings and photographs and things like that. The field is recommending those types of things at this point, correct?

A. Yes. Photographs have been somewhat of a

contentious issue in the past because unlike perhaps -- I don't want to speak for fingerprint examiners. I don't know for sure. I have heard they make decisions based on photographs.

Firearms examiners do not do that. We don't make decisions based on photographs. We make decisions based on what we see with the microscope and naked eye. There's been a little bit of controversy -- how many photographs do you want to take and so forth? The trend has been toward greater transparency and to be able to allow another examiner to re-create your exam if they're doing a re-examination.

So therefore photographs are good as well as a narrative description of exactly what you are looking at and so forth. This doesn't have to do with decisions -- decision-making, drawing conclusions. It has to do with really documenting what happened so that it's a quality assurance principle.

We want to have everything transparent so if another examiner needs to examine it he can follow up and retrace our footsteps.

Q. Thank you. I think that covers the background. Are you ready?

A. Uh-huh.

I have talked about tests. These are the

tests that produce the error rates, two basic types. Proficiency tests. These are usually provided by the Collaborative Testing Service and they're designed as quality assurance devices. That doesn't mean they're useless, however, for validity purposes.

The second is what are called black box or what I call black box validity tests. These are designed to directly and robustly test the discipline's fundamental propositions, the propositions I mentioned in the beginning.

These are the -- we're getting into the results. For the proficiency tests between 78 and 91 these are the results. The firearms false positive rate was 1.05 percent, false negative rate 1.76 percent. You can read the rest of the numbers.

The sensitivity -- importantly, sensitivity was 93.8 percent. Down below sensitivity was 78.7 percent. This relates somewhat to the issue of inconclusive results. The highest I have ever seen -- highest figure I have ever seen on aggregate data is the toolmark false negative of 5.36 percent.

Q. That's not firearms. That's other types of tools?

A. Yes.

Q. But not firearms?

11:22AM A. Correct.

11:22AM Q. If you can just briefly tell us how does CTS perform its tests -- how do I get that? Who sends it out? What type of evidence is it and what are they testing?

11:23AM A. The Collaborative Testing Service, in my experience with the FBI they will send, for example, two or three, four or five bullets to the laboratory and usually each examiner would get a set and they simply ask, "What can you tell us about these?"

11:23AM Basically what you are doing is looking at all of them and saying, "Were any fired from the same gun? Were they definitely not or can we not tell?"

11:23AM Q. Then you send your results along with those samples back to CTS and they then look at all the numbers and decide -- is there a pass/fail?

11:23AM A. Well, if you say one bullet came from gun A and a second bullet came from gun A when they didn't, that's a false positive error. Yes, that gets incorporated into the reports they produce.

11:24AM Q. Do those get sent back to the laboratory as some sort of performance measure for the examiner?

11:24AM A. Oh, absolutely. I mean that's -- if there's a mistake made that's a big deal. Obviously, the examiner first has to write his report on this test and there's

usually -- a lot of labs have internal quality managers or quality assurance departments or units.

That's a big deal. If there's a fundamental mistake like that, meaning a false positive or negative, means something needs to be done. There's a procedure invoked. What are we going to do about this? What was the cause of the problem?

Q. Is the examiner notified that they had a false positive or false negative?

A. I'm sure they are. I never experienced this in our laboratory where we had that happen.

Q. Meaning nobody in your laboratory ever had a false positive or a false negative on the performance test?

A. Actually there was one before I arrived, one that I heard about. But in all the years I was there from 96 forward I never heard of a false positive or false negative error. There were sometimes issues over is that inconclusive or not but, no, we didn't have those errors.

I know for sure there's all kinds of written procedures about what happens if there is an error and that involves not only the laboratory but that could involve the accrediting agencies and so forth.

Q. Okay. Thank you.

A. My colleague Doug Murphy has calculated from the CTS reports the results from post 1992 up to 2005 and you see the results there. Sensitivity is still high. Very quickly just to review, sensitivity has to do with how many times you're capturing true positives.

If there's 100 comparisons that involve bullets from the same gun and you say they're from the same gun 90 times out of a hundred that's a 90 percent sensitivity. The toolmark false negative rates -- you're basically in the same range as before on here.

Now, these are a little different animal. These are black box validity tests. I'll go into it shortly, why they're a little bit different. These -- some of them, especially the ones we did in the FBI, are very tightly controlled. We were trying to eliminate as much as possible quality assurance errors, for example.

We didn't want to include examiners who we thought might possibly be incompetent which were none in the FBI. We were trying to keep it very controlled. Doug Murphy on the tests we did in 2003 went to extraordinary lengths to try to design it such that it would be as valid as possible.

And then several of these tests were done in the FBI kind of using our study as a template. There were others done that went outside to the broader

examiner community. These can't be as tightly controlled but the advantage is it reaches a lot more of a base, the numbers of examiners participating.

The Brundage study had to do with barrels. The Bunch and Murphy had to do with cartridge cases. The DeFrance study had to do with barrels. The Thompson and Wyant study had to do with knife edges. The Smith study had to do with cartridge cases and barrels or cartridges cases and bullets or breech faces and barrels.

The Orench study had to do with fracture matching. The Hamby, Brundage and Thorpe study had to do with barrels. The Lyons study, I believe, had to do with extractors which is a part of a gun. The Giroux study had to do with screwdrivers.

Now, if I might say one thing. All of these you see zero percent. Nobody should be foolish enough to think the true error rate is zero, that it's flawless. If anybody tells me a scientific test is flawless, I'm going to say they're wrong. There's no such thing as a perfect scientific test.

In those tests there were no false positive or false negative error which is a very good outcome. In one of the FBI tests, the one at the bottom, there was a 1.3 percent error rate. That resulted from a

single mis-elimination whereas apparently -- I talked to Giroux about this.

We can't identify who committed that error, deliberately so. We designed it that way deliberately and for good reason. But that was an example of someone apparently excluding on the basis of microscopic marks and not on the basis of class characteristics.

Q. That study was looking at what?

A. Screwdrivers.

Q. Screwdrivers, not firearms?

A. Correct.

Q. And all of the firearms tests other than -- I think you said Lyons was extractor marks.

A. Correct, I believe, if I am not mistaken.

Q. All the others are zero percent?

A. Yes. There were no fundamental errors, no.

Q. Do you have any idea how many examiners were participating in these studies?

A. Yes. In the Brundage study, the early one, I think there were 30. The latter one I think there were hundreds, if I am not mistaken. Don't quote me on that. There were many, many hundreds, I believe.

Q. Dr. Hamby was one of the authors of that study?

A. Correct. The very tightly controlled FBI studies -- I'm not sure about -- Thompson and Wyant

study was also many, many examiners. The FBI studies which is Bunch, Murphy, DeFrance, Smith, Orench and Giroux, those involved only the examiners in the FBI because we wanted to control it and limit it to examiners we knew were competent and qualified. That would be eight to ten examiners. I think maybe usually eight or nine.

Q. The ones we were talking about before, the proficiency tests by CTS, who is included in those studies or in those numbers?

A. Who are the participants?

Q. Right.

A. Virtually all examiners to my knowledge. There may be some, I suppose, that don't take them. It's also broader than that. Trainees will take them. That's one of the things I'll talk about, the criticisms of the various tests and pros and cons. If you paid the money you could take one.

So that may inflate the error rate somewhat. There was at least one instance where my colleague Doug Murphy went in and figured out the difference in error rates between including trainees and excluding trainees. The error rate dropped quite a bit when trainees were excluded.

This is where your eyes may glaze over but

11:31AM 1 we have to assess the error rate data because there's
11:31AM 2 been a lot of criticism of one type or another and it
11:31AM 3 needs to be taken into account. There are claims of
11:31AM 4 understated error rates from validity and proficiency
11:31AM 5 tests.
11:31AM 6 That is the claim that the true error rates
11:31AM 7 are higher than this because of these reasons. There
11:31AM 8 could be a perverse incentive for effecting
11:31AM 9 inconclusives. If somebody is taking a test why should
11:31AM 10 I say yes or no? If I say I don't know, then I can't
11:31AM 11 possibly be proven wrong.
11:31AM 12 There's a structural possibility of this
11:31AM 13 perverse incentive. Does that really obtain? I think
11:31AM 14 fundamentally the answer is no. Here's why. For
11:32AM 15 validity tests, at least the ones in the FBI, we design
11:32AM 16 them specifically to be anonymous and we told the
11:32AM 17 participants, "These are anonymous. If you make a
11:32AM 18 mistake, it can't be traced back to you."
11:32AM 19 Therefore, they had no incentive to do it
11:32AM 20 any other way than they would casework. For proficiency
11:32AM 21 tests outwardly you would think they would have an
11:32AM 22 incentive to effect an inconclusive. But when you look
11:32AM 23 at the facts that doesn't seem to be the case at all.
11:32AM 24 If anything, it's the opposite.
11:32AM 25 When I was unit chief in the FBI laboratory

what happened was the pressure is actually for the opposite because if you effect an inconclusive, that raises a yellow flag to the quality assurance authorities both within the bureau and outside that says, "I wonder if this person is trying to play it safe by effecting an inconclusive or if this person is incompetent and isn't seeing something we know to be true."

In reality many times you would see specimens on proficiency tests that looked like that previous slide I showed you of the inconclusive where that is a proper call. You don't want to do anything other than say inconclusive on that kind of situation.

Sometimes that would happen on proficiency tests and I told my examiners -- and I think this happened across the country -- just do it exactly the same way as you do casework. If it results in an inconclusive we simply have to explain that to the quality assurance authorities and if they want to look at the specimens they can.

That's what happened, I think, a couple of times, not that they came and looked at them, but they believed us. The pressure, though, is for an examiner to say, "Hum. Am I going to cause a possible problem by calling it inconclusive when it's probably from the same

11:33AM 1 gun or probably not from the same gun?"
11:34AM 2 The pressure tends to be to say yes or no
11:34AM 3 even though structurally and outwardly it would -- at
11:34AM 4 that point the incentive would be for an inconclusive.
11:34AM 5 That's a lot of words for that first one.
11:34AM 6 The issue of inconclusives, I think I just
11:34AM 7 talked about that. The lack of total blindness. Yes.
11:34AM 8 You do not have total blindness in these tests, whether
11:34AM 9 it's proficiency tests or validity tests. By that I
11:34AM 10 mean the participants know they're taking a test.
11:34AM 11 Just like in medical drug testing when you
11:34AM 12 sign a consent form you know you're taking a test. But
11:34AM 13 just like in the medical testing, many of ours were
11:34AM 14 double blind, especially the ones in the FBI. We
11:34AM 15 specifically design them to be this way.
11:34AM 16 They didn't know -- the participants did
11:34AM 17 not know the answers. They could not possibly find out
11:34AM 18 the answers and we didn't even give them the test
11:34AM 19 packets directly. They had to go to a separate room to
11:34AM 20 collect the test packets.
11:35AM 21 We didn't know which test packet they had.
11:35AM 22 We couldn't tell them the answers because we kept them
11:35AM 23 under lock and key and it was a question of numbers and
11:35AM 24 randomly generated numbers. There's no way we could
11:35AM 25 have possibly signalled to them any kind of answers.