Some Predictions About Future Trends In Forensic Science

D. Michael Risinger
John J. Gibbons Prof. of Law
Seton Hall University School of Law
Increased Emphasis on Research
Increased Emphasis on Expressions of Uncertainty
Controversy over Interpretation Practices
Increasing Emphasis on Transparency and Completeness in Report Writing, and Consultation on request with all sides.
Emphasis on Developing Methods for communicating results understandably to non-specialists without overstatement or understatement
New Approaches to Error Management
Increasing Emphasis on Masking Protocols
Increased Emphasis on Research
Doesn’t mean every bench analyst must do research, but does mean that every analyst must cooperate in research.
NIJ must show it can be an honest broker funding well-designed research that can show weaknesses as well as strengths.

So far, so good.
Mnookin et al (including forensics people like Barry Fisher Max Houk, Glenn Langenburg, Jay Siegel, Keith Inman, Norah Rudin, and David Stoney) and a few others of us.
The Need for a Research Culture in the Forensic Sciences

UCLA Law Review, 2011

with nice added commentary by Joe Bono

Google “Mnookin Research Culture” and download for free
Increased Emphasis on Expressions of Uncertainty
Metrology and the Guide to the Expression of Uncertainty in Measurement (GUM)
Controversy over Interpretation Practices
English/Continental Approaches (Bayesianism) vs. Typical American Approach (Direct expression of probability of common source attribution)
Colin Aitken & Franco Taroni
Statistics and the Evaluation of Evidence for Forensic Scientists
Increasing Emphasis on Transparency and Completeness in Report Writing,

and Consultation on Request with all sides.
The Role of the Forensic Science as an honest broker representing scientific results to the adversary system demands this.
Emphasis on Developing Methods for communicating results understandably to non-specialists without overstatement or understatement
Graphic/Geometric is better than formally quantified.

Attend a program on statistical concepts by Stephen Morgan

Read some of the literature on “words of estimative probability” and then judge the defensibility of various word scales in forensic disciplines
New Approaches to Error Management
Realization that error is not synonymous with fault, and that errors need to be frankly admitted to be managed (Medical Model)

James Reason, Human Error (1990)

Risinger—Whose Fault? (2011)
E-mail risingmi@shu.edu

and I will send you a copy
Increasing Emphasis on Masking Protocols
Standards for distinguishing domain-irrelevant information, plus masking protocols, that result in no domain irrelevant information at the interpretation stage, and domain relevant information in the least biasing order.
Go to JFS electronic version and search for “sequential unmasking”
Level I Propositions: Offence Level

Level II Propositions—Activity Level

Level III Propositions—Source Level
The result provides support for the prosecution’s proposition over the defense proposition.
There are ways in which I think attempting a demarcation between science and non-science is not helpful, and ways in which it is.
A place where the science/non-science demarcation is surprisingly unhelpful:

Understanding the notion of expertise.
Expertise is best understood by functional contrast with the ordinary fact witness.
Fact witnesses are interchangeable with jurors except for time and place.
Whenever a witness makes assertions that cannot be accounted for by interchangeability, the witness is performing some sort of expert function.
“Expert” is best thought of as an adjective applying to assertions, not to witnesses.
This universe of expert function can be divided into two domains:

Summarizational (Educational) and Translational
Summarizational (Educational) expertise can also be usefully divided into two species:

Everyday

Technical
Shoe Man
Translational—Takes factual information equally available to expert and jury, and turns it into a conclusion about another non-obvious factual proposition (which we unhelpfully call an "opinion" or sometimes a conclusion.)
Translational assertions convert facts equally available to the jury and the expert from a less usable to a more usable form, based on an asserted translation system possessed by the witness but not by the jury.
Translational expertise is most people’s normal model, dealing as it does in “conclusions” or “opinions”
Translational assertions convert facts equally available to the jury and the expert from a less usable to a more usable form, based on an asserted translation system possessed by the witness but not by the jury.
There are many potential species of translational system, the most important of which are:

1. experience based
2. research based
In reality, there tends to be an admixture of both in real world applications, but one or the other is usually clearly dominant.
In judging reliability of asserted expertise, the questions to be asked are always the same for all forms.
What is the case specific target issue to which the expertise is directed?

What is the case-specific claim of expertise?

What available information bears on a rational belief warrant in regard to the reliability of this specific claim of expertise?

What is the appropriate standard of certainty for such a belief warrant given the kind of case, the issue involved, the distribution of the burdens of production and persuasion in the case, and the standard of proof involved in regard to the issue upon which the expertise is proffered?
Notice that nothing has yet been said about science. Science is only one form of translational system (but an important one)
However, when answering question three (about the reasons to believe a claim of expertise), we must decide if we are dealing with science or not, because that shapes how we approach that question.
It’s not that belief in non-science experienced-based claims are not sometimes warranted, but that they will be warranted (if they are) in a different way than the products of science.
Minimum condition of reliability for experience-based expertise:

- Unambiguous feedback for correct results in normal practice
When this condition is not present, only testing of practitioners by the standards of science can supply the appropriate belief warrant.
So we must now deal with some sort of approach to the demarcation problem.
Frye v. United States
293 F. 1013
(D.C. Circuit, 1923)
‘The rule is that the opinions of experts or skilled witnesses are admissible in evidence in those cases in which the matter of inquiry is such that inexperienced persons are unlikely to prove capable of forming a correct judgment upon it, for the reason that the subject-matter so far partakes of a science, art, or trade as to require a previous habit or experience or study in it, in order to acquire a knowledge of it. When the question involved does not lie within the range of common experience or common knowledge, but requires special experience or special knowledge, then the opinions of witnesses skilled in that particular science, art, or trade to which the question relates are admissible in evidence.’
Just when a scientific principle or discovery crosses the line between the experimental and demonstrable stages is difficult to define. Somewhere in this twilight zone the evidential force of the principle must be recognized, and while courts will go a long way in admitting expert testimony deduced from a well-recognized scientific principle or discovery, the thing from which the deduction is made must be sufficiently established to have gained general acceptance in the particular field in which it belongs.
If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence, or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify in the form of an opinion or otherwise.
Daubert v. Merrell Dow Pharmaceuticals, Inc.

509 U.S. 579 (1993)
General Electric Co. v. Joiner

532 U.S. 136 (1997)
Kumho Tire Co. v. Carmichael

526 U.S. 137 (1999)
Most people, including most judges at all levels, have never bothered to actually read *Kumho Tire v. Carmichael*. 
The “Task at Hand” requirement of specificity
“AS WE SAID BEFORE,…THE QUESTION BEFORE THE TRIAL COURT WAS SPECIFIC NOT GENERAL. THE TRIAL COURT HAD TO DECIDE WHETHER THIS PARTICULAR EXPERT HAD SUFFICIENT SPECIALIZED KNOWLEDGE TO ASSIST THE JURORS “IN DECIDING THE PARTICULAR ISSUES OF THE CASE”
This prescribes the “task at hand” approach to expertise. The issue is not global reliability, but reliability for the application actually involved in the case.
A mandate ignored by virtually every lawyer and every federal judge in the country!
Most people, including most judges at all levels, have never bothered to actually read *Kumho Tire v. Carmichael*.
The “task specific” approach to judging the reliability of expert assertions is not only mandated by Kumho Tire, it is the only reasonable approach to the question of reliability, even though it requires....
WORK
ONE TASK OR MANY?
What is the case specific target issue to which the expertise is directed?

What is the case-specific claim of expertise?

What available information bears on a rational belief warrant in regard to the reliability of this specific claim of expertise?

What is the appropriate standard of certainty for such a belief warrant given the kind of case, the issue involved, the distribution of the burdens of production and persuasion in the case, and the standard of proof involved in regard to the issue upon which the expertise is proffered?
“CAN DOCUMENT EXAMINERS RELIABLY IDENTIFY HANDWRITING AS TO AUTHORSHIP OR ORIGIN?”

VS

“CAN DOCUMENT EXAMINERS RELIABLY DETERMINE IF A PARTICULAR PERSON WHOSE NAME IS REFLECTED BY A PUTATIVE SIGNATURE ACTUALLY SIGNED THE SIGNATURE?”

VS

“CAN A DOCUMENT EXAMINER USING THE TECHNIQUES USED, RELIABLY DETERMINE IF A PARTICULAR PERSON WHOSE NAME IS REFLECTED BY A PUTATIVE SIGNATURE ACTUALLY SIGNED THE SIGNATURE UNDER THE CONDITIONS OBTAINING IN THIS CASE?”

VS

“DID THE DOCUMENT EXAMINER IN THIS CASE ACCURATELY DETERMINE IF THE SIGNATURE AT ISSUE WAS A FORGERY?”
“CAN DOCUMENT EXAMINERS RELIABLY IDENTIFY HANDWRITING AS TO AUTHORSHIP OR ORIGIN?”

VS.

“CAN DOCUMENT EXAMINERS ACCURATELY DETERMINE THE TRUE WRITER OF A FORGED SIGNATURE FROM THE FEW LETTERS INVOLVED?”

VS.

“CAN DOCUMENT EXAMINERS ACCURATELY DETERMINE THE TRUE WRITER OF A FORGED SIGNATURE FROM THE FEW LETTERS IN THE SIGNATURE IN THIS CASE UNDER THE CIRCUMSTANCES INVOLVED IN THIS CASE?”
What is the case specific target issue to which the expertise is directed?

What is the case-specific claim of expertise?

What available information bears on a rational belief warrant in regard to the reliability of this specific claim of expertise?

What is the appropriate standard of certainty for such a belief warrant given the kind of case, the issue involved, the distribution of the burdens of production and persuasion in the case, and the standard of proof involved in regard to the issue upon which the expertise is proffered?
HAVE TESTS SHOWN THE RELIABILITY OF THE EXPERTISE IN REGARD TO THE PARTICULAR TASK BEING PERFORMED IN THE CASE BEFORE THE COURT?
HAVE PROCEDURES BEEN ADOPTED TO SCREEN OUT SUGGESTION AND EXPECTATION? (Blind Testing and Line-ups)
IF NOT, THE RESULTS WOULD NOT SEEM TO BE THE PRODUCT OF “RELIABLE METHODS” RELIABLY APPLIED, AS REQUIRED BY KUMHO TIRE AND NEW RULE 702
FRE 702 TESTIMONY BY EXPERTS
If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise, if

(1) the testimony is based upon sufficient facts or data,

2) the testimony is the product of reliable principles and methods, and

(3) the witness has applied the principles and methods reliably to the facts of the case.
What is the case specific target issue to which the expertise is directed?

What is the case-specific claim of expertise?

What available information bears on a rational belief warrant in regard to the reliability of this specific claim of expertise?

What is the appropriate standard of certainty for such a belief warrant given the kind of case, the issue involved, the distribution of the burdens of production and persuasion in the case, and the standard of proof involved in regard to the issue upon which the expertise is proffered?
LOWER STANDARD OF RELIABILITY FOR PROSECUTION-PROFFERED EXPERTISE THAN FOR OTHER EXPERTISE?
Land Valuation
Expert
(appraiser)
LOWER STANDARD OF RELIABILITY FOR PROSECUTION-PROFFERED EXPERTISE THAN FOR OTHER EXPERTISE?
Science is …

- An enterprise committed to cognitive rationality by ideology
- Not timebound in its search for answers
- Not (directly) concerned with normative questions beyond cognitive rationality
Science is also an enterprise whose individual practitioners may depart substantially from the ideal.

- a social enterprise dependent on a complex web of checks and balances for its success over time.
Fundamental Characteristics

- Formal Data
- Reliable Taxonomy
- Generalizations asserting regularities, based on data
- Testing of predictions from generalizations by reference to formal data.
- No Formal Data
- No Testing of Claims
- No Science
Consider fingerprint identification
(And other forensic identification specialties)
Additional Considerations:

- Masking (blinding)

- The astrology test
So emphasizing the demarcation between science and non-science too early in analysis may interfere in analyzing expertise in general, and in organizing our approach to the right questions to ask.
But some answer to the demarcation problem is necessary to make proper judgments about why to believe a claim of expertise.
So the boundary between science and non-science is both irrelevant (to some issues regarding expertise) and centrally relevant (to proper evaluations of belief warrants)
Determining task-specific error rates for forensic science processes is very important.
But if I personally could choose only one reform to undertake, it would not be to do research to tie down task specific error rates under test conditions, it would be to establish appropriate blind testing in actual practice.
Risinger, Saks, Thompson and Rosenthal,

*Observer Effects in Forensic Science: Hidden Problems of Expectation and Suggestion*,

90 Cal. L. Rev. 1

(2002)
Any Information not necessary to the exercise of one’s expertise will distort results
And the more that information engages emotions and desires, the stronger the distortion will be.
This is one of the best established and supported general propositions of modern cognitive psychology, and has given rise to a great improvement in the methodology of research and application in much of science.
Any process using a human as a perceptor, rater, or interpreter should be “as blind as possible for as long as possible”

-- Robert Rosenthal
No area of forensic science can fulfill this fundamental responsibility without a formal consideration of what constitutes domain-relevant and domain-irrelevant information.
To date, no area has instituted such a study.
Why?
Think of the response of surgeons in the 1870’s to Joseph Lister’s message about the need to control microbial infection during surgery.
The usual response by forensic practitioners is that their good faith and their training make them able to resist such distortions.
Thus the data and principles that have been shown universally applicable in every context in which they have been tested don’t apply to them.
Are Forensic Specialists immune?

Consider the case of Brandon Mayfield.
Fingerprints: Brandon Mayfield
Merely Anecdotal?

How about this:
Bias Among Forensic Document Examiners: A Need for Procedural Change

Larry S. Miller
12 J. Police Sci. and Admin
407
1984
12 document examiners were divided into two groups of six.

Each group was given materials from a check forgery case in which ground truth was known by reference to other evidence.
One group was given the “request” writing of only one person, which resembled the writing on the check pictorially.
They were further told that there were two witness who had watched the checks signed, and identified that suspect.
The second group was given the same request writing, plus request writing from two other persons, and told nothing else.
All six of the examiners in group 2 eliminated all three suspects as the writer of the checks.
Four examiners in group 1 concluded that the “suspect” had written the questioned signatures on the checks.
The fifth examiner reported an “inconclusive” but said that the request exemplars bore indications of disguise.
Procedural Bias in Forensic Examinations of Human Hair

Larry S. Miller

11 Law and Human Behavior 157 (1987)
56 hair identification tests were prepared.
Half the tests reflected the usual practice of presenting a known hair from a “suspect” and a single “questioned” hair from the crime scene, and asking if the two “matched”. (a “show-up”)

The other half of the tests presented five “known” hairs from “suspects” to be compared to the “questioned” hair from the crime scene, and asked if the hair from the scene matched any of the suspects. (a “line-up”)
In every test, the “crime scene” hair did not come from any of the “suspects,” though the hairs of all the “suspects” were selected to present characteristics not obviously dissimilar to the crime scene hair.
14 qualified examiners were given four tests each, two from each set of test designs.
Erroneous declarations of “match” were found in 3.8% of the responses to the “line-up” condition, but in 30.4% of the responses to the “show-up” condition.
Still not convinced?

Visual hair comparison already known to be too unreliable to draw any general conclusion based on studying it?

Consider the Dror et al Study (2006 Forensic Sci. Int. 74-78)
Five experienced fingerprint examiners were asked by a colleague to evaluate the Mayfield prints after it was known that the FBI had misidentified them.
In reality, they were given prints they themselves had found to match in actual cases.
<table>
<thead>
<tr>
<th>5 yrs earlier</th>
<th>Now</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive Ident</td>
<td>?</td>
</tr>
<tr>
<td>Positive Ident</td>
<td>?</td>
</tr>
<tr>
<td>Positive Ident</td>
<td>?</td>
</tr>
<tr>
<td>Positive Ident</td>
<td>?</td>
</tr>
<tr>
<td>Positive Ident</td>
<td>?</td>
</tr>
<tr>
<td>Positive Ident</td>
<td>?</td>
</tr>
</tbody>
</table>
Four of the five came now came to a different result.
One now said that the latent was too small and smudged to reach a conclusion
<table>
<thead>
<tr>
<th>5 yrs earlier</th>
<th>Now</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive Ident</td>
<td>too smudged to call</td>
</tr>
<tr>
<td>Positive Ident</td>
<td>?</td>
</tr>
<tr>
<td>Positive Ident</td>
<td>?</td>
</tr>
<tr>
<td>Positive Ident</td>
<td>?</td>
</tr>
<tr>
<td>Positive Ident</td>
<td>?</td>
</tr>
</tbody>
</table>
And three now concluded that the latent didn’t match the known, (when they had come to the opposite conclusion in the real case)
<table>
<thead>
<tr>
<th>5 yrs earlier</th>
<th>Now</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive Ident</td>
<td>too smudged to call</td>
</tr>
<tr>
<td>Positive Ident</td>
<td>exclusion</td>
</tr>
<tr>
<td>Positive Ident</td>
<td>?</td>
</tr>
<tr>
<td>Positive Ident</td>
<td>?</td>
</tr>
<tr>
<td>Positive Ident</td>
<td>?</td>
</tr>
<tr>
<td>5 yrs earlier</td>
<td>Now</td>
</tr>
<tr>
<td>---------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>Positive Ident</td>
<td>too smudged to call</td>
</tr>
<tr>
<td>Positive Ident</td>
<td>exclusion</td>
</tr>
<tr>
<td>Positive Ident</td>
<td>exclusion</td>
</tr>
<tr>
<td>Positive Ident</td>
<td>?</td>
</tr>
<tr>
<td>Positive Ident</td>
<td>?</td>
</tr>
<tr>
<td>5 yrs earlier</td>
<td>Now</td>
</tr>
<tr>
<td>------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>Positive Ident</td>
<td>too smudged to call</td>
</tr>
<tr>
<td>Positive Ident</td>
<td>exclusion</td>
</tr>
<tr>
<td>Positive Ident</td>
<td>exclusion</td>
</tr>
<tr>
<td>Positive Ident</td>
<td>exclusion</td>
</tr>
<tr>
<td>Positive Ident</td>
<td>?</td>
</tr>
<tr>
<td>5 yrs earlier</td>
<td>Now</td>
</tr>
<tr>
<td>---------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>Positive Ident</td>
<td>too smudged to call</td>
</tr>
<tr>
<td>Positive Ident</td>
<td>exclusion</td>
</tr>
<tr>
<td>Positive Ident</td>
<td>exclusion</td>
</tr>
<tr>
<td>Positive Ident</td>
<td>exclusion</td>
</tr>
<tr>
<td>Positive Ident</td>
<td>Positive Ident</td>
</tr>
</tbody>
</table>
Think the n is too small?

Think the malleator is too unusual?

How about a replication using more normal context cueing.
Itiel Dror & David Charlton,
Why Experts Make Errors

56 J. Forensic Identification
600
(2006)
6 experienced fingerprint examiners were given eight sets of two prints each by their supervisor.
All of the print pairs given each examiner were from previous cases where that examiner had declared that there was a sufficient basis to declare a match (four each) or an exclusion (four each)
In addition, each of these cases had been rated as to difficulty by the examiner when originally performing the comparison.
In four of the test cases presented (two of previous “match” [one hard, one easy] and two of previous “exclusion,” [one hard, one easy], no context information was provided, merely a request for comparison
In the other four cases (similarly distributed), not uncommon context information was given ("suspect has confessed, etc").
The test thus resulted in 48 decisions (6 examiners X 8 comparisons each)
Of those 48 decisions, 6 were inconsistent with the previously rendered decision in the actual case. (12.5%)
Two of the six examiners gave results completely consistent with their previous decisions. The other four did not.
Three of the four remaining examiners changed one decision each, and the other examiner changed three.
Four of the changes were in tests where context information was supplied, and two were in cases where no context information was supplied.
Five of the switches were in cases rated as difficult, but the one switch in an easy case (from match to exclusion) was in a case containing context information suggesting exclusion.
Things are complicated by the fact that even domain-relevant information can have a deleterious biasing effect.
For instance, seeing a known specimen before the questioned specimen is characterized can distort the characterization of the questioned specimen to fit the known

The best solution: Sequential Unmasking

(See Risinger et. al; Krane et al. Krane et al.; Krane et al.; Thompson et. al.)
Sequential Unmasking allows the complete filtering of domain-irrelevant biasing information and allows domain-relevant information to be presented in its least biasing order.
Just like in physics!
(see Donald Koetke
Senior Research Professor of
Particle Physics, Valparaiso
University,
“A Blind Analysis”
slide show available at
www.physics.valpo.edu/faculty/dkoetke/presentations/Blind_Analysis1.ppt)
A Blind Analysis

You are *not* allowed to peek!

Prof. Donald Koetke
Senior Research Professor of Physics
Valparaiso University
Outline

1. What is a “blind analysis”?  
   • What is a “bias”?  
   • Does “bias” = “systematic error”?  
   • What is the special (subtle) bias?  
2. What is the history of the blind analysis”?  
3. What are examples of “blind analysis” in physics, astronomy, astrophysics, and, … ?  
4. Some thoughts and reflections --
Why the NAS Report’s “wait for more research” approach is wrong:

There’s already plenty of research.
That’s what this presentation has been about. There’s already plenty of research both generally and specific to forensic science.
Are there alternatives? There may be various approaches to structuring examinations, and other debiasing techniques, that give some protection.
But these are always second best and supplemental to masking protocols to filter the biasing information in the first instance.
So again, as for awaiting more research
You don’t have to measure the rate of rainfall to know you should use an umbrella, especially if you want to be sure not to get wet at all.