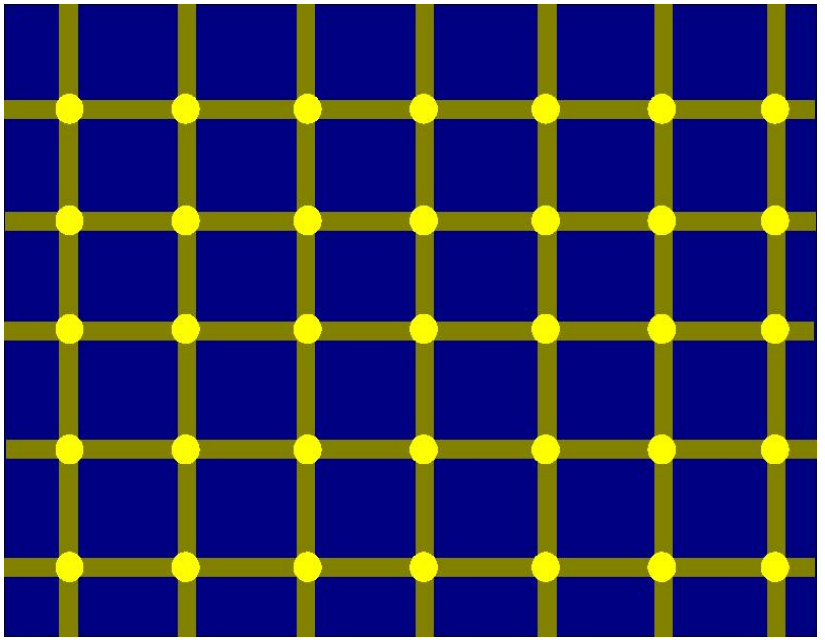
What you see is NOT What you get

HTCIA – Aug 23-25, 2009

Dr. Fred Cohen
President - California Sciences Institute
CEO – Fred Cohen & Associates

Outline

- Background of the speaker and subject
- The Big Picture
- Some things you may see – or not...
- Now you see it - now you don't
- What's the difference?
- I can see clearly now
- Questions / Comments?



Your speaker

- Education:
  - B.S. Electrical Engineering (C-MU '77)
  - M.S. Information Science (Pitt '81)
  - Ph.D. Electrical Engineering (USC '86)
- Experience:
  - >30 years of information protection R&D, design, engineering, testing, implementation, and operation
  - >20 years since first digital forensics case
- CEO - Fred Cohen & Associates
  - Enterprise information protection architecture
  - Digital forensics for high-valued legal cases

CalSci

- President – California Sciences Institute
  - Starting doctoral classes in 2009-10
- M.S. And Ph.D. Program in National Security
  - Technical aspects of these fields
- M.S. In Advanced Investigation
- Ph.D. In Digital Forensics
  - The first Ph.D. program in Digital Forensics in the United States
- calsci.org

What does he know about the subject?

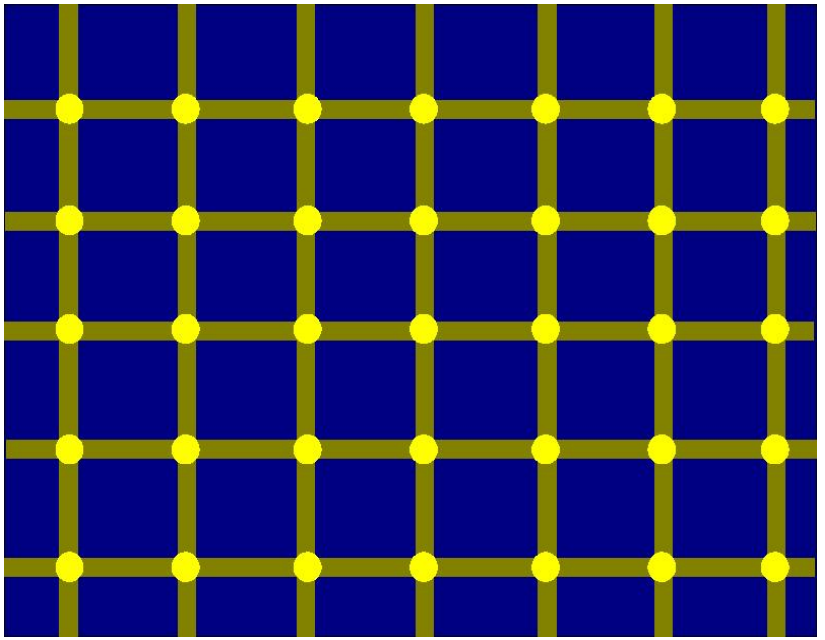
- Knowledge, skill, experience, training, or education
  - Federal Rules of Evidence 701-706
- Knowledge, Skills, and Experience:
  - Created commercial email servers and processing, industry analyst analyzing these mechanisms, operates email processing systems for high volumes of email, teaches, lectures, etc. in the area, POST certified trainer in these areas, admitted to testify as an expert in Federal, State, and Local Criminal and Civil digital forensics matters including related to emails, published refereed and other articles, etc.
- Education:
  - B.S., M.S., and Ph.D. in relevant field

Cognitive limits and digital data

- Several published journal articles on deception
- Two books on deception and information
- Teaches graduate courses on the subject
- 5 issued patents on deception technologies
- Deception ToolKit (used by thousands of sites)
- Magic tricks (including “The beat hand”)
- Deception detection in digital forensics cases
- Largest experiments on deception technologies
- Technical mechanisms in use in DoD systems

Outline

- Background of the speaker and subject
- The Big Picture
- Some things you may see – or not...
- Now you see it - now you don't
- What's the difference?
- I can see clearly now
- Questions / Comments?

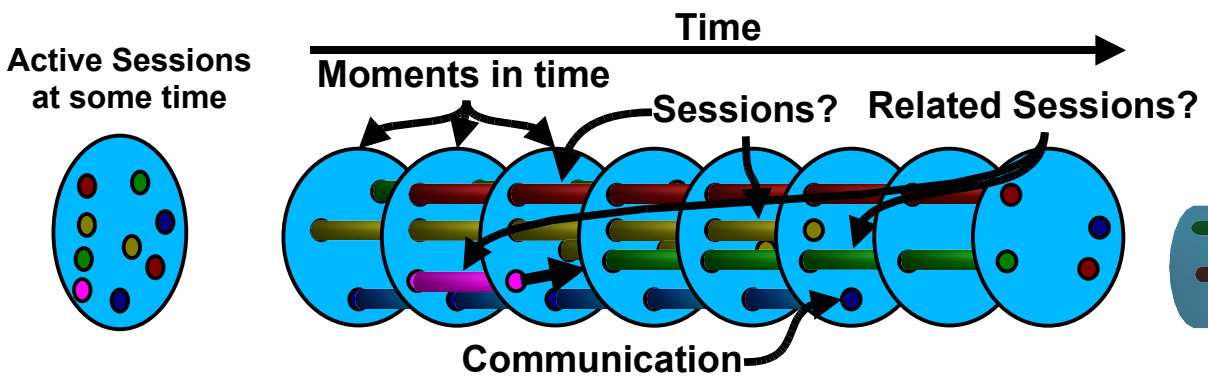


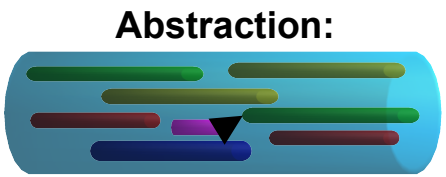
What is deception all about?

- People make cognitive errors (our nature)
- Intentional deception
  - Induce and/or suppress signals
  - Produce type 1,2,3 errors
  - Use feedback to support/verify (in active mode)
- But in forensics...
  - Before we can deal with intentional deception
  - We need to deal with self-deception
  - Our own cognitive errors in examining traces

Active deception and forensics

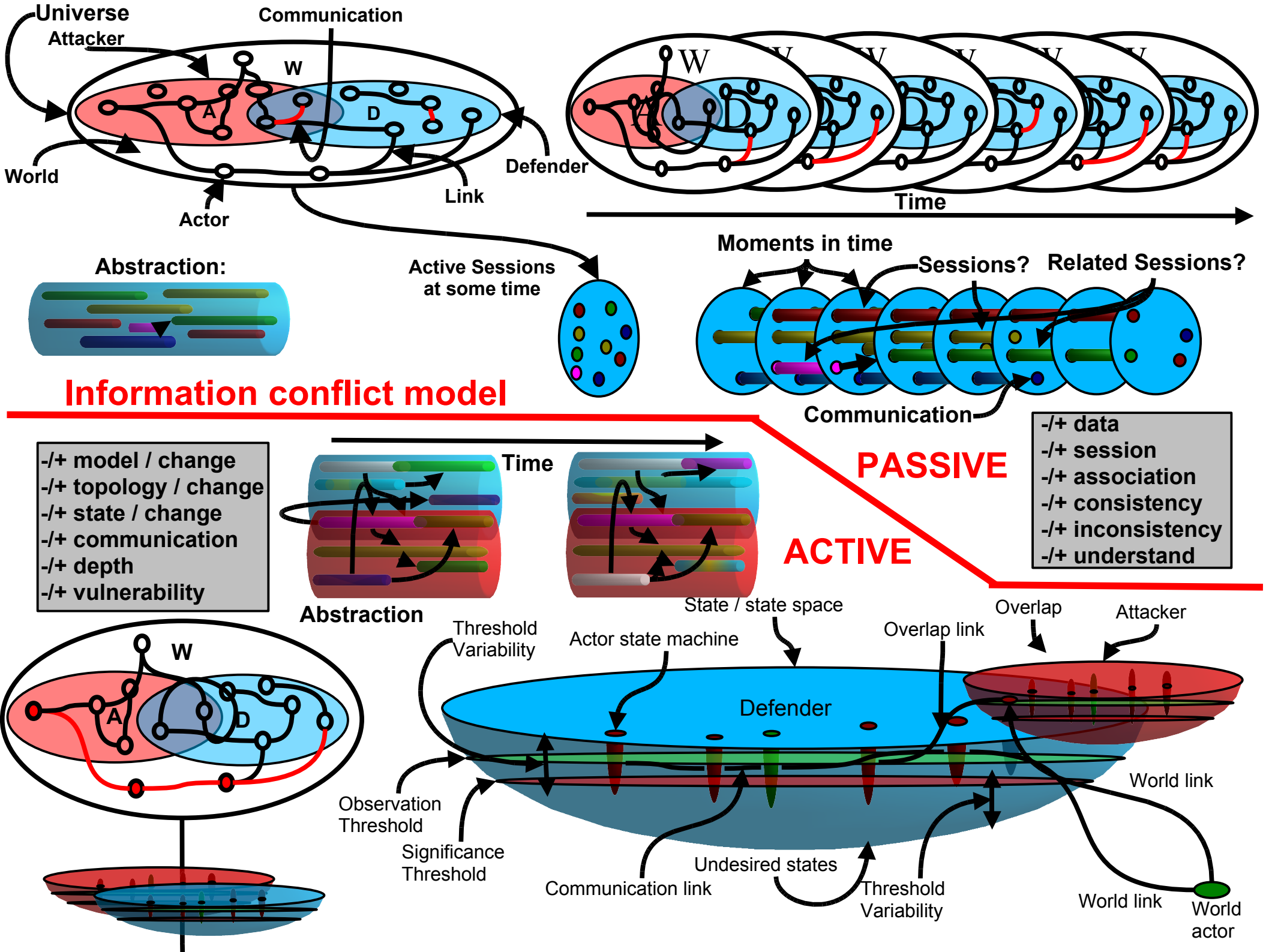
- We observe “flashes of light”
- Our brain turns them into the things we see
- When we use computers
  - Computers generate those flashes of light
  - From traces that are latent in nature
- An example from network sniffing





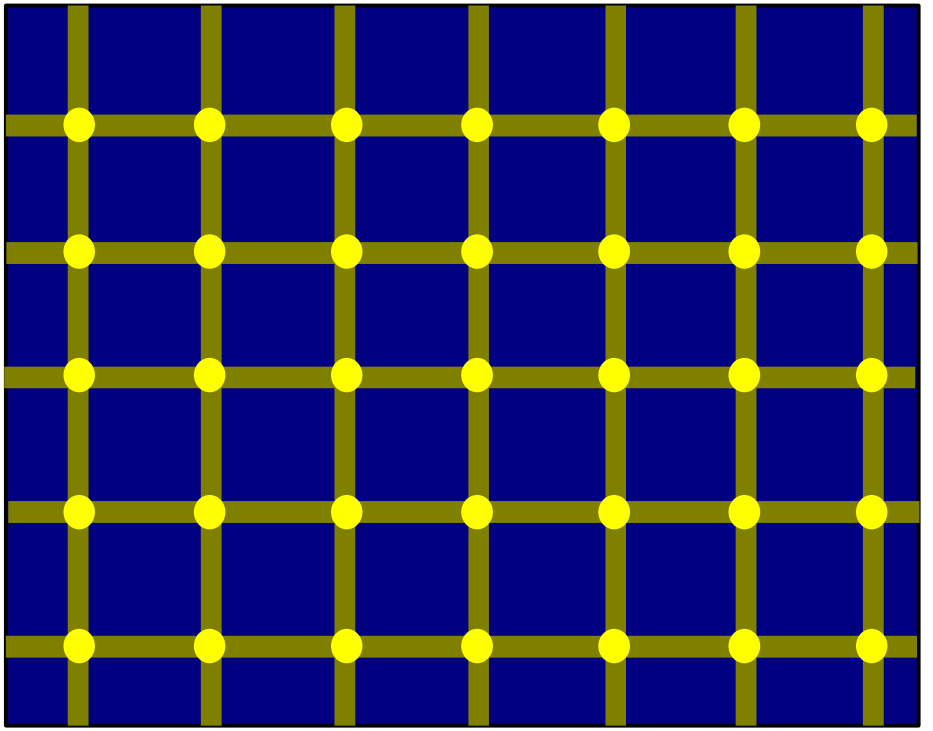
Cognitive Errors:

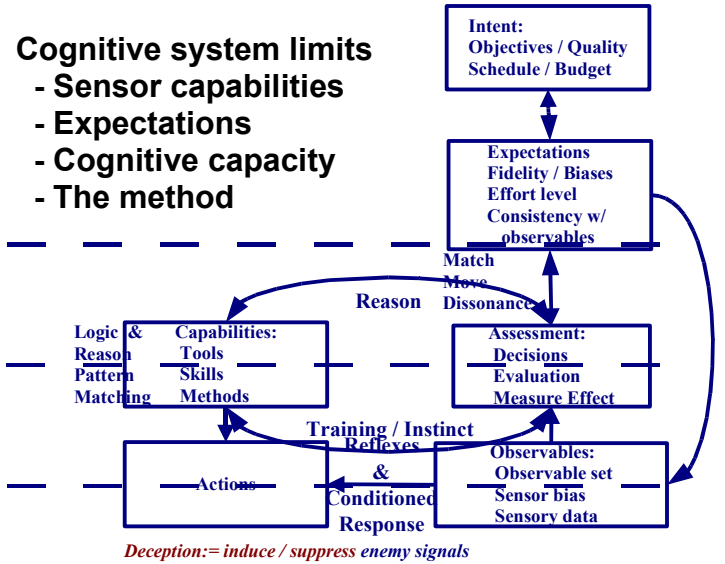
- -/+ data
- -/+ session
- -/+ association
- -/+ consistency
- -/+ inconsistency
- -/+ understand


Universe
Attacker
Communication
W
A
D
World
Actor
Link
Defender
Time
Abstraction:
Active Sessions
at some time
Moments in time
Sessions?
Related Sessions?
Communication
Information conflict model
-/+ model / change
-/+ topology / change
-/+ state / change
-/+ communication
-/+ depth
-/+ vulnerability
Time
Abstraction
PASSIVE
ACTIVE
-/+ data
-/+ session
-/+ association
-/+ consistency
-/+ inconsistency
-/+ understand
W
A
D
State / state space
Overlap
Attacker
Threshold
Variability
Actor state machine
Overlap link
Defender
Observation
Threshold
Significance
Threshold
Communication link
Undesired states
Threshold
Variability
World link
World link
World
actor

Cognitive mechanisms

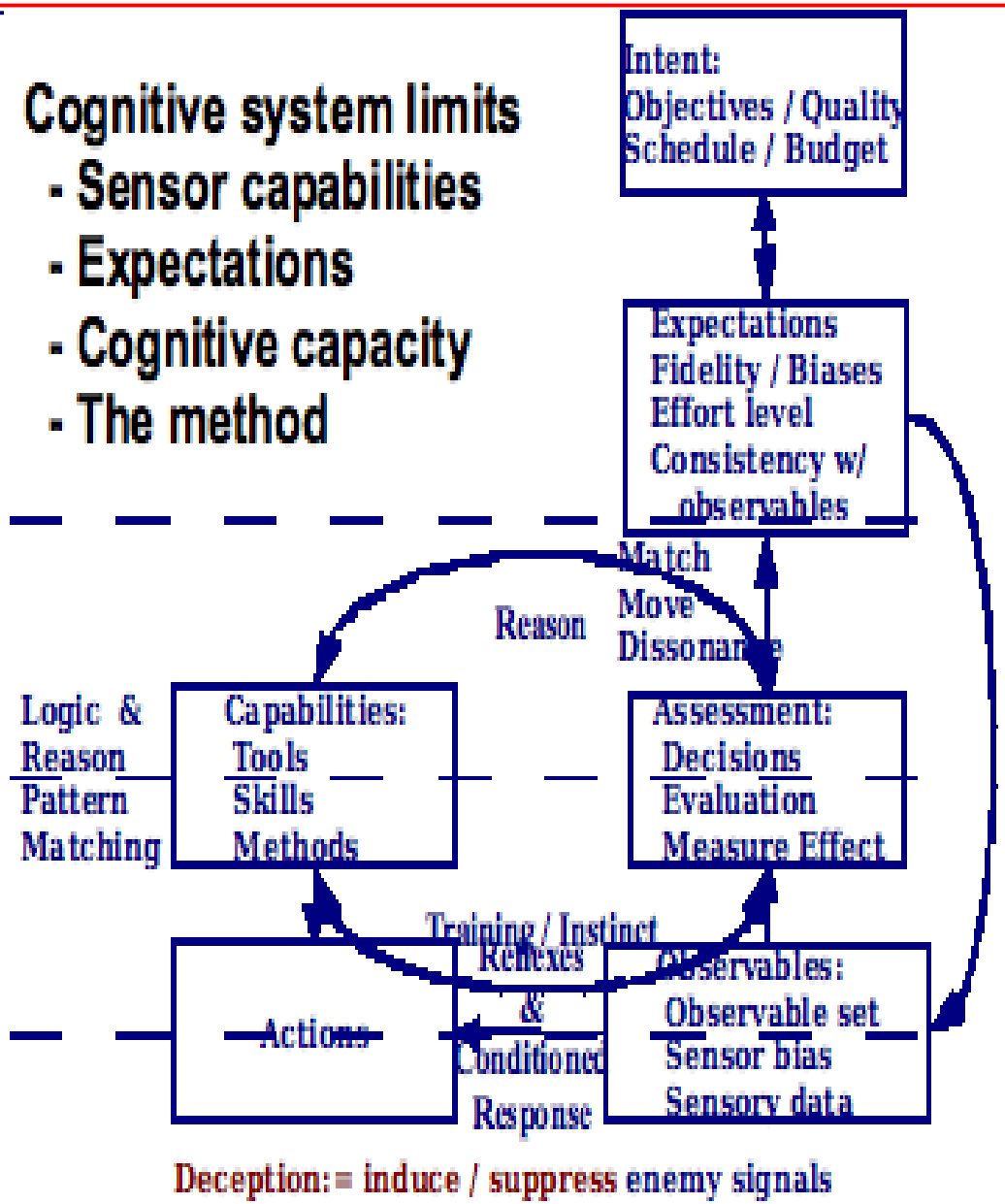
- The scintillating grid
  - A low-level deception
- Mechanisms are identified at several cognitive levels
- Lots of psychological research supports this
- These mechanisms produce errors in our interpretation





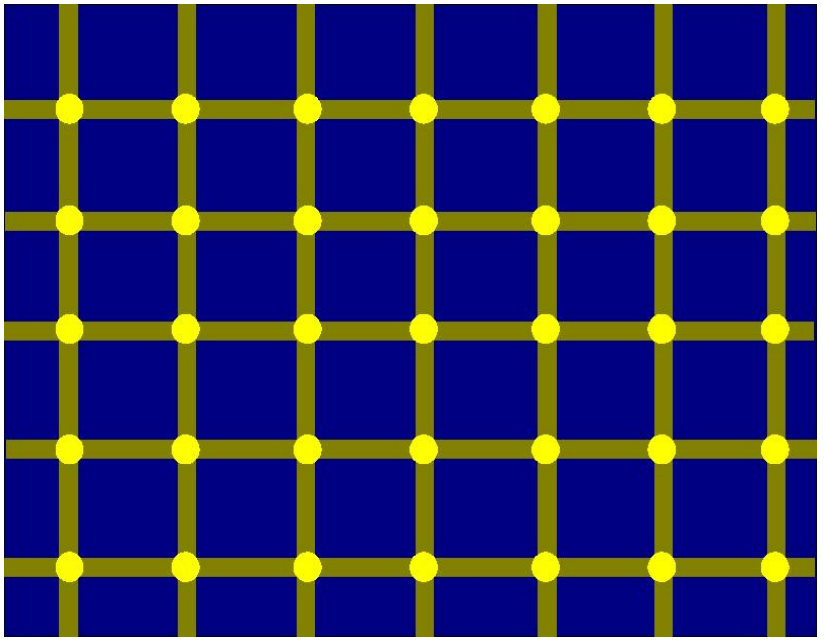
People and their limits

- People interpret what they see
- Based on their state of mind and their perspective
- They react and adapt and their adaptations change the way they perceive
- People see what they they expect to see



Outline

- Background of the speaker and subject
- The Big Picture
- Some things you may see – or not...
- Now you see it - now you don't
- What's the difference?
- I can see clearly now
- Questions / Comments?



How long has it been?

- 1O examples (this line has the I[st] three)
  - When a space is not a space?
  - When is a number not a number?
- Are these pictures identical? Similar? How?

There are a lot of possibilities

- Problem 1:
  - We see only traces – not complete histories
- Problem 2:
  - We only see reflections of the traces – not bits
- Problem 3:
  - We only see them in the presentation context(s)
- Problem 4:
  - Presentation contexts are not forensically sound
- Problem 5:
  - We don't know all the other problems

Traces are incomplete

- We essentially never have all of the potentially relevant traces
  - All of the traces from the machines we have (√)
  - All of the histories of the machines we have
  - All of the directly related machines
    - Their traces
    - Their histories
    - The linkages between them
  - All of the indirectly related machines
    - Their traces, histories, and linkages
  - All of the interdependencies of all of them

Evidence is latent

- We cannot see the bits
  - We use tools
    - Some are “forensically sound” in some sense
    - Most are not “forensically” sound in any sense
  - All we ever see is
    - The results of tools that perform functions that depend on the bits
    - Their presented reflections at the interface
    - The interface depending on the hardware we use
    - The flashes of light from them that hits our eyes
    - The interoperation of those flashes by our brains

Viewing context is limiting

- We can see traces in various formats
  - Hexdump
  - Editors
  - Tables and charts
  - Graphical depictions and roll-ups
  - GUI presentations by file type
  - Different layouts and colorizations
  - Database interfaces
  - Web interfaces
  - The list is small and finite...
- Otherwise, create tools for the need

Views not forensically sound

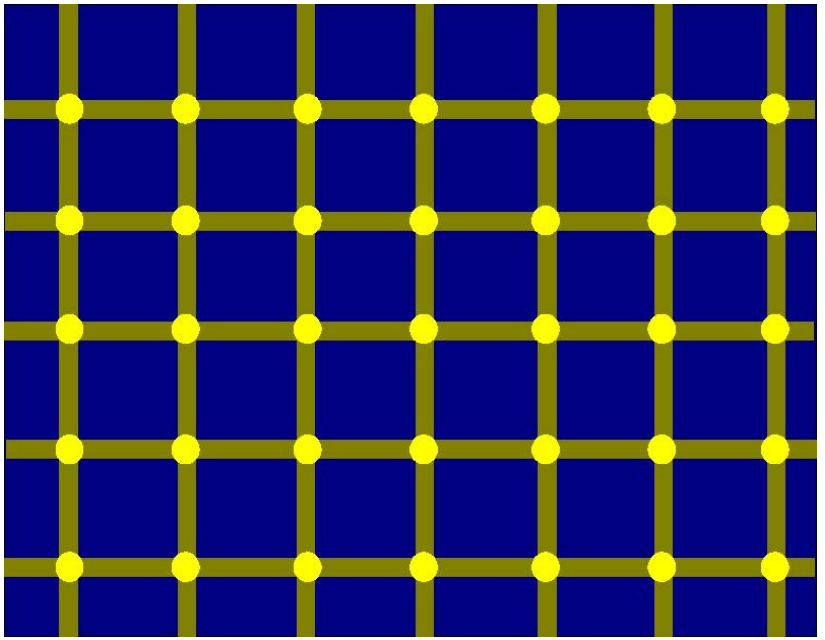
- What does that even mean?
  - No theoretical or scientific basis published for defining forensic soundness of presentations
  - Different contexts present different notions of soundness
  - Which do we show – or all of them?
    - The bits as 1s and 0s? Octal? Hex? Decimal?
    - Presentation based on the format we chose?
    - Presentation based on assumptions of syntax?
    - Mixes of images with text?
    - As they appear in one browser or another?
    - As we think they were seen at the time?

What are we missing?

- Lots of known unknowns
  - We are limited by time from even doing what we know how to do and have tools for
  - We know we don't have all the tools we could
  - We know we don't know how to use tools we don't have.
- Who knows how many unknown unknowns?
  - Actually – we do...
  - There are only so many possible traces
  - But the numbers are enormous
  - We can never be “thorough”!!!

Outline

- Background of the speaker and subject
- The Big Picture
- Some things you may see – or not...
- Now you see it - now you don't
- What's the difference?
- I can see clearly now
- Questions / Comments?



DFE is fragile

- The fragile nature of the original writing makes it easy to spoliate
  - How do you know what you are looking at is the real thing? Or an exact copy?
  - While in some cases you get a forensically sound image, in many cases, you do not
  - All you get are some traces
- How can you see the difference between a forgery and a real trace of claimed events?
  - The answer lies in the science of digital forensics

A scientific theory of DFE

- Science cannot tell us that what we see is what we think we see
- Science can only confirm or refute hypotheses
- The methodology for confirmation and refutation
  - Make testable claims (hypotheses or theories)
  - Predefine tests, outcomes, and interpretations
  - Test DFE against hypotheses or theories in an independently repeatable way
  - Use the predefined interpretations to confirm or refute the hypotheses or theories
- Refutation takes precedence over confirmation

Legal matters are competitive

- The other side may see things differently
  - If they look for things you don't look for...
  - If you look for things they don't look for...
- How do you know what you see is real/right?
  - You don't! You never can
  - But the jury is there to decide
  - If you say this looks like a chair and they look for themselves, you or the other side will win
- You have to present what you have found
  - And they will present what they have found

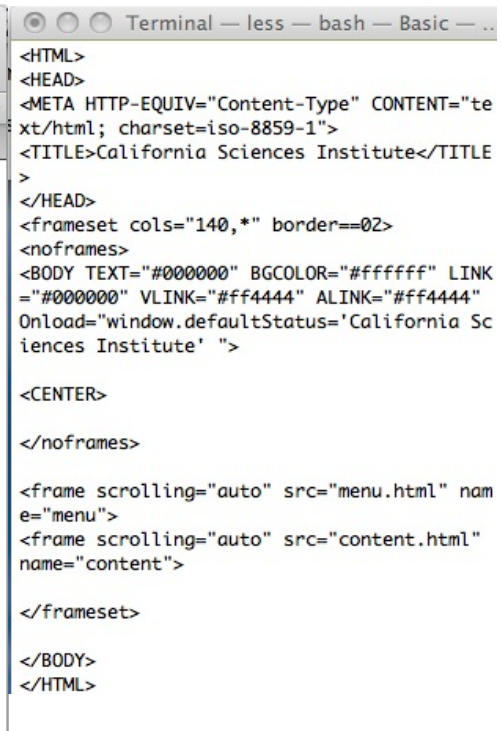
Presenting what you see

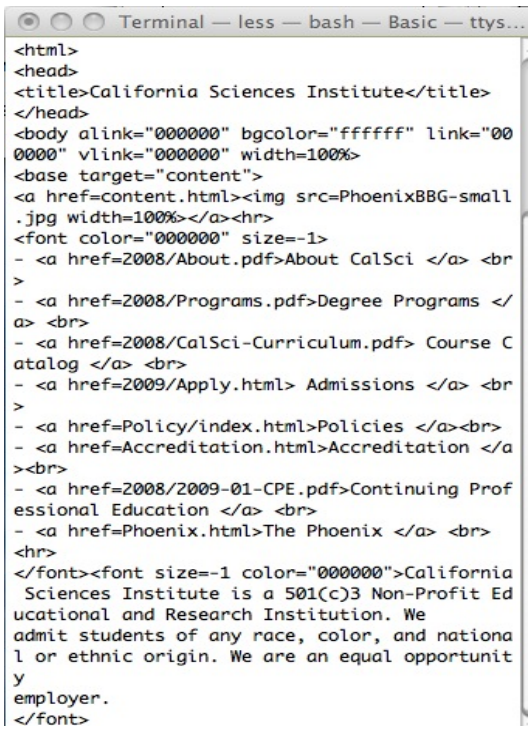
```
Return-path: <bounces-d09c3da279-5a62a0bae6@b.cts.vresp.com>
Received: from smtpin133-bge351000 ([10.150.68.133])
 by ms283.mac.com (Sun Java(tm) System Messaging Server 6.3-7.04 (built Sep 26
 2008; 64bit)) with ESMTP id <0KO000F4QC86ROP0@ms283.mac.com> for
 dr.cohen@mac.com; Fri, 07 Aug 2009 06:02:30 -0700 (PDT)
Original-recipient: rfc822;dr.cohen@mac.com
Received: from gc-mkt89.verticalresponse.com ([206.57.6.146])
 by smtpin133.mac.com
 (Sun Java(tm) System Messaging Server 6.3-8.01 (built Dec 16 2008; 32bit))
 with ESMTP id <0KO000A4MC824H41@smtpin133.mac.com> for dr.cohen@mac.com
 (ORCPT dr.cohen@mac.com); Fri, 07 Aug 2009 06:02:30 -0700 (PDT)
X-Brightmail-Tracker: AAAAAQAAAUA=
Return-path: <bounces-d09c3da279-5a62a0bae6@b.cts.vresp.com>
DKIM-Signature: v=1; a=rsa-sha1; d=vresp.com; s=dkim; c=simple/simple;
    q=dns/txt; i=@vresp.com; t=1249650150;
    h=From:Subject:Date:To:MIME-Version:Content-Type;
    bh=BZ5bJAvREVHNTo9/6eenQZ9ilWA=;
    b=aGEWXl/ivQYjY9elcPNUVu/QEDuCi6XypT01vdebhZqGPrjLkPkYZVNHXTza1EQG
    iuiybkowNXkLzXwd4b1Zsxd/74KKh0yeYgGT4SrYzOx3qVxaOJIeeevyjMnEd/M/
    rwuz0exyMv+x5l+WiecUiEf1TiGQHN7F3F58k24JVR8=;
DomainKey-Signature: q=dns; a=rsa-sha1; c=nofws;
    s=mkt; d=vresp.com;
    h=Received:From:Reply-To:To:Subject:Date:Message-ID:List-Unsubscribe:MIME-
Version:X-Company_ID:X-vrfbldomain:X-vrpod:X-CTS-Enabled:X-Campaign:Content-Type;
    b=fi/Njtsl+I/csxcUtaubvwN141DCIG0//Wa7/GaQQhopmP4jc3531ZBAhn+Nw1M2
    6Q1drGOYdEBVc3NUx+z0tIfh+s3Z73wvowVRI0b9uk+FnLU/qleddJudOrNsXq7n
    PSGW0KbGEKaCaJSBQZbdwPoZZpE59fsHxOJLEx0jE4c=
Received: from [10.4.7.56]
 ([10.4.7.56:54498] helo=mailer02.sf.verticalresponse.com)
    by pacifica.sf.verticalresponse.com
```

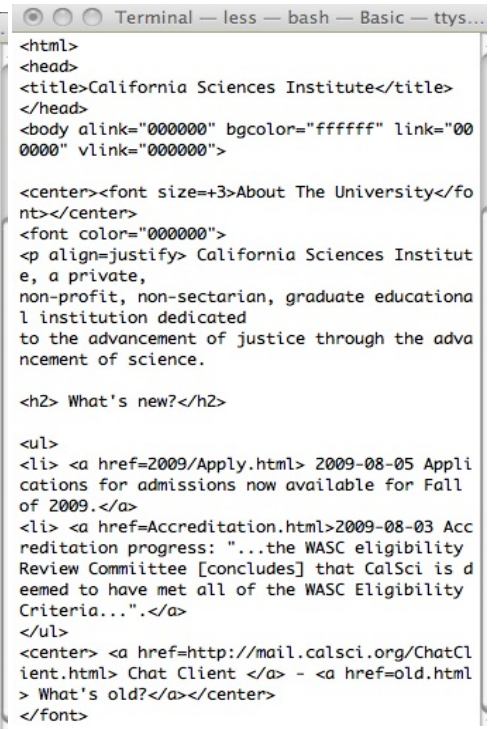
- The jury does not see as well as you do!
  - You have knowledge, experience, training, education, and skills that they lack
  - You need to come to be able to explain yourself to them so they see it your way

Presenting what you see

- The visualizations you use will prejudice them
  - The goal is to make the presentation probative as to the issues in the case while prejudicing the jury in favor of your view of the truth







⇠ is the visualization of ⇡ as presented by firefox
but the presentation of ⇡ is a visualization of bits as text

Presenting what you see

- If a visualization is misleading, how so?
  - They are almost all misleading to some extent
  - How do you explain interpretation to the jury?

Terminal — less — bash — Basic — ttys...

```
<html>
<head>
<title>California Sciences Institute</title>
</head>
<body alink="000000" bgcolor="ffffff" link="00
0000" vlink="000000">

<center><font size=+3>About The University</fo
nt></center>
<font color="000000">
<p align=justify> California Sciences Institut
e, a private,
non-profit, non-sectarian, graduate educationa
l institution dedicated
to the advancement of justice through the adva
ncement of science.

<h2> What's new?</h2>

<ul>
<li> <a href=2009/Apply.html> 2009-08-05 Appli
cations for admissions now available for Fall
of 2009.</a>
<li> <a href=Accreditation.html>2009-08-03 Acc
reditation progress: "...the WASC eligibility
Review Commiittee [concludes] that CalSci is d
eemed to have met all of the WASC Eligibility
Criteria...".</a>
</ul>
<center> <a href=http://mail.calsci.org/ChatCl
ient.html> Chat Client </a> - <a href=old.html
> What's old?</a></center>
</font>
```

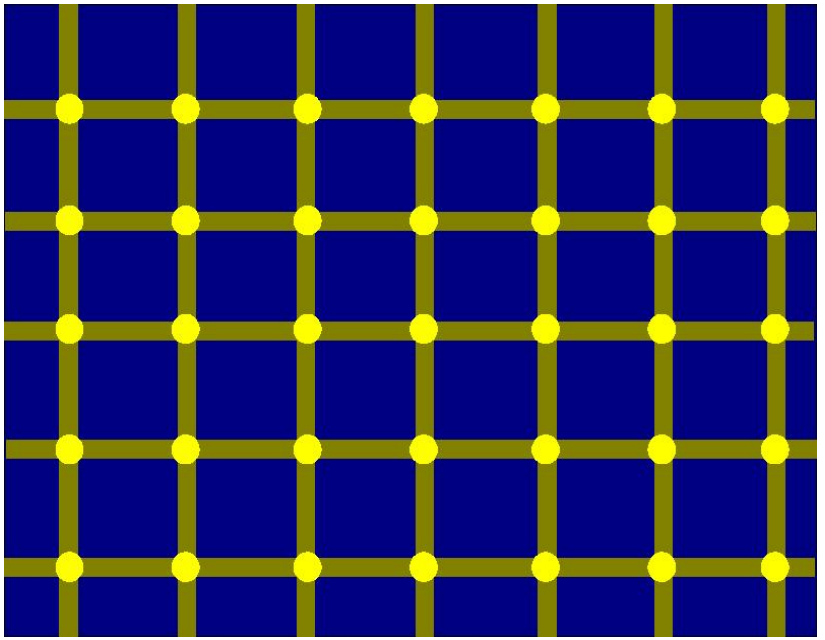
The underlying code for a Web page that was presented by firefox in this case came from expressions written in the hypertext markup language (html) shown at right...

It recurses to bits and physics

- Symbols you see are a visualization of the bits interpreted using the American Standard Code for Information Interchange (ASCII) ...
  - The visualization portrays specific bit sequences as specific characters in a particular font ...
    - The bit sequences reflect the traces identified by ...
      - Traces are generated by the finite state machines that operate the ...
        - The finite state machines store the traces as magnetically charged areas ...
          - The current theories of physics...
  - The “wrap-around” of the output mechanism shows lines continuing ...
- All of which you need to understand...

Outline

- Background of the speaker and subject
- The Big Picture
- Some things you may see – or not...
- Now you see it - now you don't
- What's the difference?
- I can see clearly now
- Questions / Comments?



Making a difference

- Party claimed that identical systems, processes, and methods were used to produce two proffered items
  - They looked pretty much the same
    ... to the tools that party used
  - But when examined with different tools
    ... differences became apparent
  - Non-printing characters were systematically different between the two items
- The visualization caused a process error to go unnoticed and all of the evidence was spoliated

What you don't understand

- Most “experts” presented in court do not understand these issues to their full depth
  - What you don't understand you are making guesses about
  - The result is misrepresentation of visualizations
  - Most misrepresentation goes unnoticed
  - But some of it does not...
- What you don't understand may mislead the court
  - Most of what we understand about any particular case is a direct result of the visualization of traces and analysis results

A common example

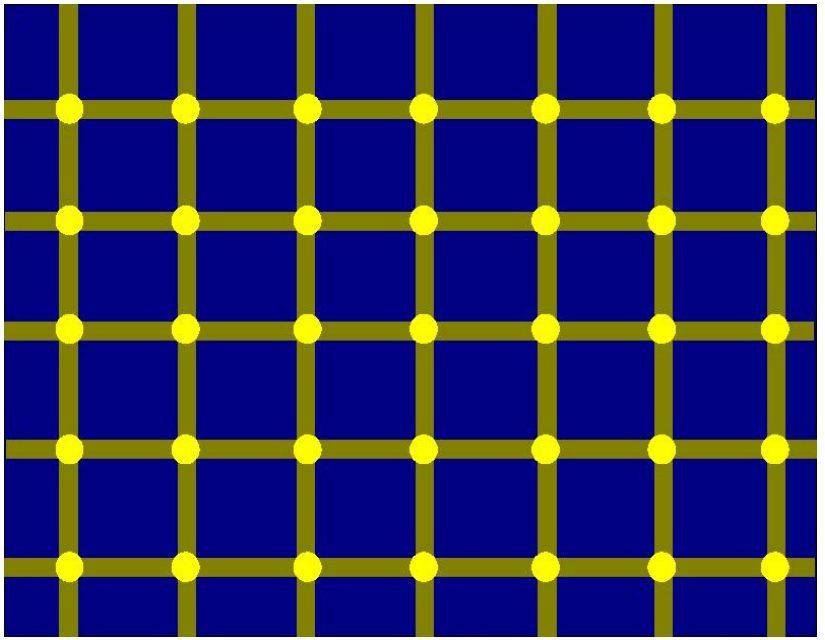
- Precision vs. Accuracy
  - 5 out of 7 items had this characteristic
    - Is that 70%? 71%? 71.42857%?
  - 50 out of 70 items had this characteristic
    - What presentation?
  - 500 out of 700
    - What presentation?
  - 50000000 out of 7000000
    - What presentation?
- How do your tools present precision and accuracy? How do you compensate?

What tools are you using?

- Many (most?) so-called forensic tools
  - Present date and time information after interpretation and without providing the basis
  - Don't present non-printing characters well (at all)
  - Have a small number of presentation methods
  - Make unstated assumptions in presentations
  - Don't provide the detailed linkage between what they present and the original traces they are based on
  - The list goes on and on...
- Go look at each of the tools in this light and ask the hard questions – what is the basis?

Outline

- Background of the speaker and subject
- The Big Picture
- Some things you may see – or not...
- Now you see it - now you don't
- What's the difference?
- I can see clearly now
- Questions / Comments?



What questions?

- What are some of the obvious hard questions?
  - What is the basis for this presentation of those traces or analysis results?
    - What are the assumptions? Why are they valid? How are they tested? How do I test them?
  - What traces and processes are being depicted, and how can I be sure of that?
  - How do I know the presentation is an accurate depiction of the traces and processes?
  - What other visualizations might be used and how might they shed light on issues in the case?
  - Is the precision proper for the accuracy?

How can you see clearly?

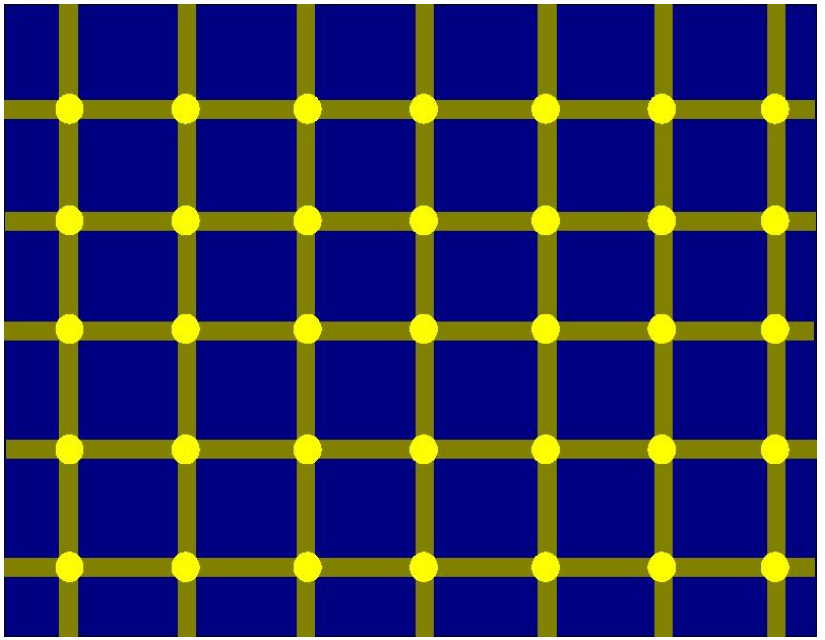
- Know your limitations
  - People who don't understand what they don't understand make claims and draw conclusions that are not justified and are often wrong
- Don't go beyond them
  - Knowing what you know and don't know helps to keep you from seeing what is not really there and helps you see what is really there for what it really is
- Be prepared
  - How will you explain what you don't know?
    - “I don't know”

One hour of CPE

- This class provides 1 hour of CPE credits
- To put it in context:
  - A M.S. degree from CalSci focussed on digital forensics involves about 160 contact hours and a total of about 500 total hours of effort over a minimum of 1 year.
  - A Ph.D. degree from CalSci focussed on digital forensics involves about 500 contact hours and a total of almost 2,000 hours of total effort over a minimum of 3 years.
- Expanding your mind expands your capacity to deal with your limitations.
  - How will you deal with yours?

Outline

- Background of the speaker and subject
- The Big Picture
- Some things you may see – or not...
- Now you see it - now you don't
- What's the difference?
- I can see clearly now
- Questions / Comments?



Thank You

http://calsci.org/ - calsci at calsci.org

http://all.net/ - fc at all.net