TOWARD A MORE SCIENTIFIC DETERMINATION: MINIMIZING EXPECTATION BIAS IN FIRE INVESTIGATIONS

Presented at the 3rd International Symposium on Fire Investigation Science and Technology, University of Cincinnati, May 20, 2008

John J. Lentini, CFEI, D-ABC
Scientific Fire Analysis, Big Pine Key, FL

ABSTRACT

This paper discusses the concepts of expectation bias and confirmation bias in science in general and in fire investigation in particular. Self examination, location by the discovery of errors in many fields of inquiry has led to a focus on the sources and influence of bias. This is one of the leading topics in forensic science today, and will likely be the focus of a National Academy of Sciences report on the status and needs of forensic sciences in the United States.

Bias has been recognized for centuries, but there is often not agreement about how to deal with it, particularly when individuals who possess biases are unaware that they are biased. Certainly, it is possible to gradually effect a cultural change by having frequent discussions to raise investigators consciousness about the existence and effects of bias. Certainly it is possible to attempt to be more "scientific" in one's approach by rigorously testing hypotheses and by entertaining the possibility that more than one hypothesis can explain the data.

Even these well-meaning steps, however, are unlikely to accomplish the goal of minimizing bias. Structural changes in the way investigations are conducted, limiting the dissemination of information on a need to know basis will not only improve the appearance of investigations, these changes will result in better and more defensible determinations.

INTRODUCTION

In the long-term, science is generally self-correcting. That is to say that the truth eventually emerges, or at least evolves, as our understanding of the universe improves. This built-in self-correction alone may suffice in many cases; however, for dealing with applications of science to shorter-term inquiries, the inherent self-correcting nature of science cannot be relied on exclusively. It needs to be supplemented and augmented by additional correction mechanisms and safeguards.

One concern is investigator bias, which may go unrecognized by an otherwise qualified and well-meaning scientific investigator. The bias may be the result of long-held beliefs or the result of recently acquired information that is peripheral to the scientific inquiry. The fact that the bias may not be recognized by its possessor makes it particularly insidious and difficult to deal with. Perhaps bias cannot be prevented from affecting human thought and reasoning, but steps can be taken to try to eliminate its undesirable effects.

THE TECHNICAL COMMITTEE ON FIRE INVESTIGATION RESPONDS

In recognition of the potentially harmful effects of investigator bias, the 2008 edition of NFPA 921 contains a new paragraph on the subject in Chapter 4 on Basic Methodology. Given the history of this chapter, the Technical Committee is ordinarily very reluctant to make any changes, but based on a proposal by committee member Douglas up with Carpenter, it was decided that the following language would provide some useful guidance.

> **4.3.8 Expectation Bias**. Expectation bias is a well-established phenomenon which occurs in scientific analysis when investigator(s) reach a premature conclusion too early in the study and without having examined or considered all of the relevant data. Instead of collecting and examining all of the data in a logical and unbiased manner to reach a scientifically reliable conclusion, the investigator(s) use the premature determination to dictate their investigative processes, analyses,

> and, ultimately, their conclusions, in a way that is not scientifically valid. The introduction of expectation bias into the investigation results in the use of only that data that supports this previously formed conclusion and often results in the misinterpretation and/or the discarding of data that does not support the original opinion. Investigators are strongly cautioned to avoid expectation bias through proper use of the scientific method.

Expectation bias and its close cousin, confirmation bias, are indeed well known. Julius Caesar, in his *Commentaries on the Gallic War* (55 BC) stated, "Men generally believe quite freely that which they want to be true."[1] Seventeen centuries later, Francis Bacon discussed confirmation bias:

> The human understanding, when any proposition has once been laid down, forces everything else to add fresh support and confirmation; and although instances may exist to the contrary, yet [the understanding] either does not observe or despises them.

Bacon also posited that

> It is the peculiar and perpetual error of the human understanding to be more moved and excited by affirmatives than negatives, whereas it ought duly to be impartial; nay, in establishing any true axiom, the negative instance is the most powerful.[2]

THIS IS NOT JUST ABOUT FIRES

There are many ways that expectation bias creeps into scientific analyses. Bias is not limited to fire investigations, forensic science, or law enforcement. Bias leads to error in many fields of science, but in forensic science, including fire investigation, the consequences of error can be most unpleasant.

---

[1] Caesar wrote in Latin. The quote is really "Homines fere credunt libentur id quod volunt."

[2] Francis Bacon, *Novum Organum*, Book I, 109, point 46 (1620), reprinted in *30 Great Books of the Western World* (Robert M. Hutchins ed., 1952).

In the medical art of differential diagnosis, Jerome Groopman has studied the sources of error, and he reports in his 2007 book, *How Doctors Think*, that the error rate may be as high as 20%. Dr. Groopman discusses the causes of these errors, and many of them are the result expectation bias. He states that in many cases, the doctor has made up his or her mind about a diagnosis in under 30 seconds.

Being “strongly cautioned” to avoid expectation bias is good guidance, but it is easier said than done. Disregarding context effects and colleagues’ expectations is difficult to do by a simple act of will. For the fire investigator, who frequently serves as the lead law enforcement official, as well as the individual making the most important “scientific” determinations in a case, eliminating bias can be particularly difficult. One way to accomplish this important goal may be to restructure the organization of the investigation, so that the expert who makes the scientific interpretation of the physical evidence is, to the extent possible, shielded from what is known as “domain irrelevant” information.

FINGERPRINT FIASCO

The corrosive impact of domain irrelevant information has been demonstrated repeatedly in forensic sciences other than fire investigation. The Brandon Mayfield case and studies conducted after it provide the perfect example.

Brandon Mayfield is an attorney from Portland, Oregon, who is a convert to Islam, and who once defended (in a family law case), an individual suspected of having ties to terrorists. After the March 11, 2004 Madrid bombings, a partial latent fingerprint found on a critical piece of evidence from one of the bombs was identified as Mr. Mayfield's. A senior FBI fingerprint analyst made the identification. He did so under extreme pressure. His job was to determine whether a latent print matched a known print. There was absolutely no reason for this fingerprint analyst to be told anything about the evidence other than being requested to make the comparison.

Instead, he was told that the unknown print was from the Madrid bombings and that the Director was very anxious for him to complete his work. Once he made the identification, the evidence was passed on to two other investigators with the request for a "peer review" of the evidence. Again, these "peer reviewers" had no need to know anything about the evidence. But they were told what case they were looking at and they were told that the senior fingerprint analyst had made the identification. Not surprisingly, they "verified" that the partial latent print was Mr. Mayfield's.

At the time the initial identification was made, the FBI's examiners did not know who Mr. Mayfield was. But they soon learned about his Muslim connections. It is possible to believe that if he had been a Maytag repairman from Des Moines, the Latent Fingerprint Unit might have re-evaluated their identification.

Spanish authorities requested and were provided copies of Mr. Mayfield's ten print card, and their analysis resulted in a negative finding. This was reported to the FBI in on April 13. Rather than reexamining their findings, the Latent Print Unit at the FBI went into a "defensive posture."

Mr. Mayfield was arrested as a "material witness", and his home and office were searched on May 6.

On May 17, the Court appointed an experienced fingerprint analyst in the private sector, who, in a May 19 report, concurred with the FBI's identification. On the same day, Spanish authorities reported that they had positively identified the real owner of the fingerprint, an Algerian national named Ouhnane Daoud.

At the request of the government, the Court released Mayfield to home detention on May 20. After reviewing Daoud's prints, the FBI Laboratory withdrew its identification of Mayfield on May 24, and the government dismissed the material witness proceeding.[3]

Mr. Mayfield eventually received an apology and $2 million for his trouble.

---

[3] US Department of Justice, Office of the Inspector General, A Review of the FBI's Handling of the Brandon Mayfield Case, Unclassified Executive Summary, January 2006.

The FBI convened a panel of experts for two days in June to investigate the fiasco. The bottom line of their findings: "To disagree was not an expected response" within the FBI's bureaucratic culture.
The story received much attention worldwide, and a subsequent study by the FBI Laboratory's Director of Quality Assurance revealed that context effects and confirmation bias (a subset of expectation bias) played a huge role in what was one of the most embarrassing episodes in FBI history. The laboratory has since changed the way they do business in fingerprints, and is in the process of implementing similar changes in all areas of forensic science.

The Mayfield episode becomes even more interesting. In order to test the effects of expectation bias, a blind study was set up in the United Kingdom. In this study, five senior fingerprint analysts were asked to re-examine one of their old latent fingerprint identifications. They were not, however, told that they were looking at one of their old identifications. Instead, they were told that they were looking at the Mayfield evidence.

In each instance, these senior fingerprint analysts had made an identification of a single latent print when compared to known suspect prints. When they re-analyzed the disguised prints, "knowing" about the Mayfield affair, they changed their determinations. Four of the five fingerprint analysts stated that the prints did not match, and the fifth stated that he could not reach a conclusion. The domain irrelevant information that was provided to these analysts, people at the top of their field, caused them each to make a very embarrassing mistake.

The UK test was an extreme way to test the effects of expectation bias, but it did prove a point. There have been numerous other examples of well-meaning individuals attempting to carefully evaluate evidence, who evaluated it the wrong way because they were told more than they needed to know.

EVEN DNA ANALYSTS CAN BE BIASED

DNA analysis is considered the "gold standard," the benchmark to which all other forensic sciences are compared. Yet, DNA analysts are subjected to the same pressures that every other forensic scientist, and indeed all human investigators, are subjected to. The following are three examples of case notes found in the files of DNA analysts, whose job it was to compare evidence from the scene or a victim with known samples from a suspect, nothing more, nothing less.

> Example #1. *"Jackson (vic) was a material witness in a case against Hodges (sus)…The suspect is a convicted felon and faced at least five years…Jackson was killed the Sunday (one week) before the grand jury was to consider the charges. It was believed Jackson was selling drugs for him…"*
> Example #2. *"Suspect-known crip gang member--keeps 'skating' on charges-never serves time. This robbery he gets hit in head with bar stool--left blood trail. [Deputy DA] wants to connect this guy to scene w/DNA …"*
> *"Death penalty case! Need to eliminate Item #57 [name of individual] as a possible suspect"*
>
> Example #3. *I asked how they got their suspect. He is a convicted rapist and the MO matches the former rape…The suspect was recently released from prison and works in the same building as the victim…She was afraid of him. Also his demeanor was suspicious when they brought him in for questioning…He also fits the general description of the man witnesses saw leaving the area on the night they think she died…So, I said, you basically have nothing to connect him directly with the murder (unless we find his DNA). He said yes."*

With this kind of notes in the file, even the most straightforward and unambiguous "match" becomes suspect. If there are any ambiguities in the analyses, such as those caused by multiple donor samples, these analysts could easily be shown to lack the objectivity required to correctly address the ambiguities in their analyses.

Context effects, caused by an overabundance of information, extend to all areas of forensic science. A fire debris analyst, for example, has no need to know that the fire investigators have determined that a particular fire was incendiary. All that chemist needs to know is that he or she has been asked to determine if there are ignitable liquid residues in the sample container. It should not matter one way or the other what the field investigator determined, and in fact, the analyst can be made to appear biased if it becomes known that he or she was told, "This is an arson fire."

The FBI's new procedure for fingerprint "confirmation" requires that the original fingerprint analyst prepare large blow-ups for comparing the latent print with the known suspect print. The QA manager requires that the identifying components be clearly marked on these blow-ups. Essentially, from the outset, the analyst is required to prepare his court exhibits.

Before the exhibit is submitted to another analyst for confirmation, however, all of the identifying notations are removed. The peer reviewer is provided no case-specific information, is not told who the original analyst was, nor is he or she told whether the original analyst made an identification or an exclusion. The peer review, therefore, meets most if not all of the requirements of a **true** scientific peer review.

"PEER REVIEW" IS IN THE MIND OF THE PARTICIPANTS

In most of the so-called peer reviews that we see in fire investigation, those scientific requirements are not met. The peer reviewer knows the case, knows all of the domain irrelevant information (the house was for sale, the marriage was in trouble, the suspect was drunk, etc.); he knows who said what about the fire and he knows that he is expected to confirm the original analysis.

In fire investigation, as investigations are currently structured, what we actually have might best be characterized as a "technical review." It is **not** a peer review. In the scientific community, peer reviews are primarily conducted in two areas: submission of manuscripts to scientific journals, and review of grant proposals. The submitter of the grant or article does not get to choose who reviews his work. The reviewer of the grant or article is asked to review in an unbiased fashion, and the results of the review will have zero impact on the reviewer's career, job prospects, or ability to get along with his coworkers. That is simply not the case in a fire investigation. An investigator who is asked to peer review and report for a colleague has many incentives to confirm the analysis and almost no incentive to disagree. It is easier to go along. One need look no further than the case of *Commonwealth of Pennsylvania v. Paul Camiolo* to see what can happen when context effects and confirmation bias influence an investigation.

Calling what happens in a typical fire investigation a "peer review" is not only inaccurate; it is intellectually dishonest. It misleads readers (or the

Court) into believing that something akin to a real peer review actually took place.

TOWARD A MORE OBJECTIVE DETERMINATION

It is one thing to understand the problems associated with expectation bias, but designing a solution to minimize bias in fire investigations is more difficult. While it is not possible to eliminate bias by an act of will, it is indeed necessary to possess the will to minimize bias. This can be accomplished by effecting a cultural change in an organization. Investigators should engage in frequent discussions about the proper application of the scientific method. These discussions can become internalized and be a source of institutional pride

Other techniques for reducing bias include making strenuous efforts to disprove hypotheses, and giving full consideration to alternate hypotheses. These "cultural" steps focus on attitude and approach, and laudable though they may be, they may not be enough. What may be required is a **structural change** in the way we approach investigations.
One structural solution is to separate responsibilities for examination of the physical evidence from case management. This necessarily requires at least two investigators: one who has the "big picture" to frame the questions about the case, and one to limit himself to the examination of the fire scene and physical evidence. The lead agent, who has all of the information, only provides the scene examiner with the information necessary to evaluate the physical evidence on the ground.

The case managers primary role is to limit information provided to analysts. The case manager has access to all of the information. He assigns analysts to collect the data, he formulates hypotheses, and he decides the best method for testing those hypotheses in consultation with the analysts.

The analysts only have access to the "domain relevant" information. They collect data and analyze it, but ideally, they do not know the "expected" results of their analysis. If an analyst needs more information, the case manager can provide it in a process of "sequential unmasking." A group of DNA scientists has proposed just such a scheme for providing analysts with information on a "need to know" basis.

An excellent example of this type of objective investigation applied to a fire took place in November 2006, following the Halloween night fire at the Mizpah Hotel in Reno, Nevada, which killed 12 people. ATF Special Agent Dan Heenan was in charge of the site inspection. He knew that there were eyewitnesses to the start of the fire, but adamantly refused to consider or even to listen to the eyewitness accounts until after his site inspection had been completed. Rather than being data for him to consider, the eyewitness statements became part of his hypothesis test. Once he independently determined the origin of the fire, which was carefully documented in his extensive report, only then did he compare what the eyewitnesses stated with the facts on the ground. It was even later that he learned the identity of the suspect and her prior history, none of which had anything to do with how the fire spread.

The result was a determination that was scientifically unassailable. The suspect pleaded guilty to 12 counts of murder and arson in exchange for avoiding the death penalty.

Eliminating all bias in a fire investigation is probably not possible. Fire investigators are human. An investigator coming upon a fire scene will be hard-pressed to ignore the fact that there is a “for sale” sign in the front yard. The emotionally charged context of a fire that kills children puts additional pressure on the investigator. **But the corrosive effects of bias can be significantly reduced.** This can be accomplished not only by the act of will to conduct the best scientific investigation possible and consider only relevant information, but also by the act of will to structure an investigation in such a way that irrelevant facts do not influence the scientific interpretation of the physical evidence.