J Forensic Sci, Jul. 1973, Vol. 18, No. 3

R. A. Harte,[1] *B.S., M.S.*

Evidence from Instruments—Their Validity, Reliability, and Role in the Adversary System

We have heard much today about the role of the expert witness and his place in the American adversary system of justice.

Since the expert witness tends generally to be someone with a technical background by education, training, and experience, and since the expert relies more and more on technological aids in his field of expertise, we see, being introduced into the courts, more and more instrument derived information.

I would like to confine my remarks today, not so much to the expert witness, but to his tools and the quality of evidence they produce for him.

We have agreed that the purpose of the adversary system is to bring out all facets of the case so that the court and the jury can see the ultimate truth in the matter and judge accordingly. Defenders of the adversary system state that two opposing advocates, doing all in their power to defend or prosecute the defendant, will reveal as much of the truth as is possible.

The evidence they gather is derived from people and things. The distortion of facts as presented by people is well known. Today, let us look at some of the distortions and prejudices produced by instruments.

In the first place, every piece of equipment—simple or complex—is the product of people. The man who dreamed up the concept, the mechanical or electrical engineer who reduced it to practice through his design, the machinest who puts it together, the technician who calibrates it, all introduce their own human imprint to the finished product. What the machine or instrument produces is what its creator intended it to produce.

This is similar to the problem confronting the computer programmer. The value of the information which the computer prints out is directly related to the validity of the information and relationships fed to the computer. As the programmers say, "Garbage In—Garbage Out."

To complicate the picture further, the instrument never really produces exactly what its creator hoped it would produce. If a device has more than one component, it is a system. Most modern devices are complex systems made up of many subsystems and components. This includes anything from sidearms to computers, from simple cameras to mass spectrometers.

In the design of any system, there are usually trade-offs that must be made in order to reach practical design goals.

Presented at the Twenty-Fourth Annual Meeting of the American Academy of Forensic Sciences, Atlanta, Ga., 2 Mar. 1972. Received for publication 27 Oct. 1972; accepted for publication 29 Jan. 1973.
[1] Omicron Systems Corporation, Palo Alto, Calif.

In order to optimize one of the performance criteria, you must sacrifice elsewhere. Aside from technical specifications, one must consider such factors as cost, weight, size, power drain, compatability with other equipment, shelf life, operating costs, training time required, complexity, resistance to temperature changes, humidity, type of handling, down time, repair costs, and so on. The poor systems engineer has his work cut out for him, and when he turns to the scientists, they say, "Do it your way as long as it meets all of the following technical specifications." They then proceed to give him a list of performance specs, many of which are mutually exclusive.

It's a little like the story of the poor grasshopper who spent the summer dancing and singing while the ants stored away food for winter. When the first snows came, the shivering, hungry grasshopper asked for food and was rejected by the prudent ants. So he went to the wise old owl who for a fee gave him advice. "Turn yourself into a field mouse, for during the winter field mice sneak into Farmer Jone's barn, are warm and have all of the leftover food from the cows and horses." The grasshopper considered that for a while and then, puzzled, asked, "That's fine, but how do I turn myself into a fieldmouse?" The wise old owl replied, "I've given you the basic concept, now you provide the engineering."

Information and Noise

The purpose of a device is to provide its operator with information. Ever since Shannon introduced the concepts of information theory, many learned articles have been written to expand the concept and apply it to various scientific fields. In general, the basic tenets can be simply stated: Information which is transmitted from a source to a receiver tends to be degraded. This degradation is due to the nature of the transmitter, the code the information is put into, the disturbances brought about in the transmission media, the unscrambling or decoding of the message, and the nature of the receiver.

A good example is the child's game of "Telephone" in which the first child in a long-line is given a whispered message by the teacher. Child 1 repeats it into the ear of Child 2, and thence to Child 3 and so on. The last child repeats the message to the teacher, who writes it down under the original. The results are often hilarious. The nature of the message can be so distorted as to bear no resemblance at the receiver end of the line to the originally transmitted one.

One of the causes of distortion is the ever present, unwanted background interference called noise. It can be interference of any nature, but was first used in sound transmission, so bears the name "noise." It is the factor which limits the sensitivity of the system; like the loud hub-bub at the crowded cocktail party so that you have to strain to hear what your host is saying; or the size of the grains of silver chloride in the emulsion of the films, so that when the crime lab darkroom blows up a photo as large as possible to see some very small detail, the detail is lost in the graininess.

It is the distortion of shape and color due to lens aberrations in a light microscope at the highest practical magnification; or it is the density of haze that an airplane pilot has to stare through before he sees a landmark at a distant airport; and it is the ripple in the atmosphere due to thermal currents which cause a small, distant target to dance in and out of the field of view in a high magnification binocular. You begin to see what noise is, and while much of it can be removed by clever design technique or careful shielding and screening from ambient sound, light, heat, and radiation of all types, there is a fundamental noise inherent in every piece of equipment, every transmission situation, which just must be lived with.

So much for noise. There are other ways to degrade or reduce information which may be fundamental to the technique. Take for example the photograph. It is often introduced into evidence because it not only conveys much information, but may have frozen that information at the time when more information was available. The verbal description of a bruise on a body as presented by a medical examiner, for instance, may not convey the impact of a photograph; and since that bruise has long since faded or perhaps even been buried on the victim, the photo holds significant information value for the jury. Some would argue that the camera, being totally objective and unable to lie or be forgetful or confused like an eye witness, is the better informer. But look at the black and white photo and see it for what it is—just a poor copy of reality.

In the first place, it lacks color. It has rendered the scene in black, white, and various shades of gray. How much information is lost in this manner? The photo of the bruise in black and white gives a good idea of the size and shape of the injury, but what to the trained physician's eye of the degree of lividity?

The next fact to consider is the lack of stereopsis or depth. How much does the lack of three-dimensional quality rob the scene on the photograph of information? How much more meaningful would the scene appear in depth?

Not only is it a flat, two-dimensional rending of a real scene, but even the best of cameras with the most expensive lenses will do some distorting due to aberrations, especially at the edges of the scene.

Finally, the photograph has inherent, in its restricted field of view, an enormous advantage and disadvantage. It can focus in on the particular object of interest and automatically crop out all the surrounding extraneous clutter, or visual noise, if you will. But by the same token, it can lose important information at some distance from the primary object. A photo of a man clutching his chest and grimacing in pain may immediately trigger in the viewer's mind the idea of a coronary attack. The photo with a larger field of view, showing simultaneously a man with a drawn revolver several feet behind the victim, introduces a substantially important new piece of information.

From what I have said about photographs then, one might conclude that he should view any so-called "objective evidence" in a new and more skeptical light.

Validity, Reliability, and Accuracy

When dealing with any measurement, either through an instrumental technique or by judgement with the human senses, a question usually arises which is stated simply this way: How true is the measurement? This is answered by evaluation of the validity, the reliability, and the precision or accuracy of the measurement.

When I stepped off the plane at Atlanta Airport, I looked at my watch which still read California time. It said 6:15 p.m. and I therefore knew that it was 9:15 p.m. local Atlanta time. I knew this because my watch was reliable even though it was reading an incorrect or invalid value. It was reliable because I knew from its tested performance over a long period of time that it does not gain or lose time. I could, therefore, calibrate the true local time accurately by adding three hours, representing the three intermediate time zones between Eastern Standard and Pacific Standard time.

Back home in my top dresser drawer, I have an old watch which is broken and with half the works gone. The hands on the watch face read 9:15 and always read 9:15. At the time I was glancing at my watch in Atlanta, the watch back home was reading the true Atlanta time. It was, therefore, a valid reading. But it is entirely unreliable since its reading some time later will be incorrect and the error will not be known. All that can be said for this watch is that two times a day it gives valid readings. Otherwise, it is worthless as a timepiece.

I will confine myself to timepieces again in discussing the precision or accuracy of a reading.

In grade school, we had a large clock on the wall over the teacher's desk, and every 60 seconds the minute hand jumped to the next minute position. It then remained there for another 60 seconds and jumped again.

It gave reliable and valid readings of time to within 60 seconds precision since a glance at the face, showing 9:15, could equally represent 15 minutes and 1 second after 9:00 or 15 minutes and 59 seconds after 9:00.

Precision, of course, has to do with the resolution of the measurement.

Again, in grade school, we used to run the 60 yard dash in the school yard, and the physical education teacher clocked us with a stopwatch that measured to the nearest tenth of a second. To me, that was a fantastic degree of accuracy. Imagine splitting so small a time unit as a second into 10 equal parts.

In high school I learned in science classes that fast cameras could be shuttered in milliseconds or thousandths of a second. Incredible!

At that time, during World War II, radar was being rapidly developed. Electromagnetic waves were being transmitted and propagated at the speed of light, bounced off objects and returned to a receiver to locate and range moving targets. To see an aircraft of 50 foot length at several miles distance required a resolution involving microsecond accuracy of the return signal.

Electronic shutters, of course, made this easily practical. Imagine accurately measuring to the millionth of a second.

In college chemistry I learned that molecular reactions were taking place in billionths of a second. The unit used was the nanosecond, and it was a new word and concept for me. This rate of molecular change was determined then on theoretical grounds and today is routinely measured by equipment like photomultiplyer tubes.

I must remind you that light travels at the incomprehensible speed of 186,000 miles, or seven times around the equator of the earth, every second. In 1 nanosecond, it has gone only 12 inches.

Today, if you attend scientific symposia of physicists and engineers, you begin to hear talk of picoseconds measurements—1/1000 of a nanosecond—the time it would take light to travel across the diameter of a human hair. As far as I am concerned, that is the limit of my imagination.

In any case, you begin to see that the precision of a reading has to do with the number of divisions you can divide your scale into and the ability to read the subdivision.

A man whose watch is calibrated in picoseconds will be able to give you the right time with a degree of precision you will never need.

To make the subject of truthfulness of a reading more meaningful to this audience, let us look at a device I am more familiar with—a breath alcohol monitor.

To specify its validity, one must compare its performance against the measurement it is supposed to be reading, that is, the blood alcohol concentration. In the State of California, the law reads that a breath alcohol testing device must consistently read within ±.01 percent of the blood alcohol concentration if it were drawn at the same time. This is a reasonable requirement of validity. If one had to match the blood concentration within ±.001 percent, a measure of precision which is unattainable, then it would be an unworkable standard.

Of course, a problem arises when one considers the accuracy and validity of the blood concentration reading itself as determined by a chemical test on a venous blood sample. Chemical analysis itself can produce errors in excess of 10 percent from true value.

This, of course, always raises that thorny question, "How accurate is the measurement of the standard by which other measurements are being judged?"

As part of the specification of validity, it must be shown that for all practical purposes the instrument is specific. That is, it will not measure anything but ethyl alcohol. The existance of any possible interfering substances tends to reduce the validity of the measurement.

In addition to specification of validity, the breath alcohol monitor must have its accuracy or precision measured. This is done by running a series of consecutive tests on the same subject or standard, and finding a series of repeat measurements that do not vary from each other by, for example, $\pm .01$ percent blood alcohol.

Finally, there is the problem of false alarm. No errors of any type are permitted to give a false positive reading when no alcohol is present in the breath. This should also be specified quantitatively as, for example, no zero blank can produce a reading greater than $\pm .01$ percent blood alcohol.

If any device or instrument to be employed is put to these tests, that is a careful evaluation of the validity, specificity, precision, and lack of false readings, one can greatly reduce the submission of inexpert evidence from his equipment entered into court records.

Not only should the forensic scientist be made to measure up to certain standards before he be considered an expert witness, but his information derived from instrumentation should be examined so that the instruments themselves meet certain standards.

Omicron Systems Corporation
1052 East Meadow Circle
Palo Alto, Calif. 94303