APPENDIX D

MODERN SHAMANISM AND OTHER FOLDEROL— THE SEARCH FOR CERTAINTY

JOHN J. FLAGLER*

You will quickly recognize that shamanism refers to the religious belief of ancient tribes that good and evil spirits pervade the world and can be summoned by inspired priests, or shamans, acting as mediums. The shamans who are the intended target of this discourse are the gurus of the electronic lie-catching business.

The icons of modern shamanism include the polygraph, the voice stress analyzer, the facial ticometer, the pupillometer, and sundry other "fidgetometers." One would have hoped that any serious doubts among arbitrators over the invalidity of electronic lie-catching apparatus were put to rest after Arbitrator Ted Jones' masterful paper on the subject presented at the 31st Annual Meeting of the Academy in 1978.[1]

I am sorry to report, however, that the electronic lie-catching business is not only alive and well, despite Ted's splendid fulminations, but indeed prospers. At the time Jones spoke to us on the question in 1978, not a scrap of respectable research support for the polygraph could be found. Indeed, the research reports offered by the polygraph industry up to that time were about on a par with those releases from the Tobacco Institute which keep telling us that cigarette smoking is quite nifty for our health.

---

*Member, National Academy of Arbitrators; and Professor of Industrial Relations, University of Minnesota, Minneapolis, Minnesota. This paper was presented at the NAA Continuing Education Conference at St. Paul, Minnesota, November 8–10, 1985.

[1]*See "Truth" When the Polygraph Operator Sits as Arbitrator (or Judge): The Deception of "Detection" in the "Diagnosis of Truth and Deception,"* in Truth, Lie Detectors, and Other Problems in Labor Arbitration, Proceedings of the 31st Annual Meeting, National Academy of Arbitrators, eds. James L. Stern and Barbara D. Dennis (Washington: BNA Books, 1979), 75–152.

I learned from my University of Minnesota colleague Dave Lykken, however, that at least three scientifically respectable studies have been published which provide some facial support for the validity of the polygraph[2]—respectable studies in that they hew to the recognized protocols of scientific inquiry. Dave Lykken who is perhaps the leading scientific critic of the polygraph extant expresses concern that these studies might conceivably reverse the national trend which has seen appellate courts in about half the states rule polygraph evidence inadmissible, even under stipulation.[3]

Indeed the latest setback for the polygraph business was the reporting out by the full House Education and Labor Committee of the Polygraph Protection Act of 1985 bill which would place significant restrictions on the use of the polygraph in private sector employment.[4] Despite this strong evidence of a public policy trend unfavorable to the use of the polygraph, my limited research on reported cases suggests that labor arbitrators may be moving in the opposite direction.

In 15 cases I found reported for the 20 years before 1975, 75 percent of those arbitrators refused to admit polygraph evidence. In the past 10 years, I was surprised to learn arbitrators admitted the lie detector evidence in six out of nine cases. What is perhaps even more disturbing, the majority of arbitrators who admitted polygraph evidence in both periods, assigned varying degrees of weight to this form of evidence.[5]

---

[2]Barland & Raskin, *Validity and Reliability of Polygraph Examinations of Criminal Suspects*, U.S. Department of Justice, Report No. 76-1, Contract 75-NI-99-001 (Washington: Government Printing Office, 1976); Horvath, *The Effect of Selected Variables on Interpretation of Polygraph Records*, 62 J. Applied Psychology 127 (1977); Kleinmuntz & Szucko, *On the Fallibility of Lie Detection*, 17 Law & Soc'y Rev. 85 (1982).

[3]Alaska, Colorado, Hawaii, Illinois, Kentucky, Louisiana, Maine, Maryland, Michigan, Minnesota, Mississippi, Missouri, Montana, Nebraska, New Hampshire, New York, North Carolina, Oklahoma, Pennsylvania, South Dakota, Tennessee, Texas, West Virginia, and Wisconsin.

[4]The Employee Polygraph Protection Act (H.R. 1524). Subsequent to the presentation of this paper the bill was passed in the House (March 12, 1986) and was sent to the Senate. The bill prevents private employers from using submission to lie detector tests as a condition of employment. Exemptions to the ban on polygraph testing include federal, state, and local government employees; defense contractors with access to confidential information; employees investigating the theft of narcotics; armored car guards; public transportation police; security alarm personnel; security guards at nuclear and electric power plants; some drug company employees; and some day care center and nursing home employees. A companion bill (S. 1815) was introduced in the Senate. The Senate Labor and Human Resources Committee has scheduled a hearing on S. 1815 for April 23, 1986.

[5]Subsequent to this presentation, a more thorough research of reported cases was published in the March 1986 issue of the Arbitration Journal. See Janisch-Ramsey, *Polygraphs: The Search for Truth in Arbitration Proceedings*, which reports that 67% of arbitrators denied such evidence in the period 1958–1966, while 67% admitted it in the period 1976–1984.

For those arbitrators who are disposed to admit and indeed credit electronic lie-catching evidence—perhaps favorably persuaded by claimed research support for the polygraph—I recommend to your attention the thoughtful critique of these studies by Dr. David Lykken.[6] The essence of his criticism can be reduced to two main propositions:

1. These recent studies actually reinforce the evidence that the lie detector is strongly biased against the innocent. Averaging the "false positive" results of these three studies shows that about half the time the polygraph examiner identified innocent subjects as "deceptive." (Subjects who were later proved innocent by extrinsic means.)

2. The "ground truth" baseline relied on to establish the validity of these studies is hopelessly compromised.

Concerning the first of these criticisms, Lykken points out that the odds that the polygraph will mark as a liar a person unjustly accused stand at no more than random chance—coin-flip odds. A few centuries of respect for the principle of "innocent until proven guilty" among all branches of civilized jurisprudence would seem to persuade most arbitrators to be a tad reluctant to credit evidence from a contraption that places the innocent at such risk.

Undaunted by evidence of its demonstrable bias against the innocent, however, some prominent fans continue to extol the virtues of the polygraph. In a remarkable role reversal, defense attorney F. Lee Bailey has recently opted for a novel burden of proof principle. On his daily television series, "The Lie Detector,"[7] Bailey zestfully proclaims that now "Every American has the right to prove his innocence"—so much for requiring the accuser to prove guilt. Bailey joins company with the nation's chief law enforcement official, Attorney General Ed Meese, who recently suggested that only the guilty need fear the lie detector.

As offensive as these propositions are to cherished principles of justice, the better case against the polygraph stands on Lykken's challenge to research claims for its validity. Validation studies examine whether the polygraph actually measures what it purports to measure—the telling of lies. The extrinsic validation baseline, known as the "ground truth," requires that means

---

[6]*See* A Tremor in the Blood: Uses and Abuses of the Lie Detector (New York: McGraw-Hill, 1981) and *Detecting Deception in 1984*, 27 Am. Behavioral Scientist 481 (1984).

[7]Now mercifully canceled.

independent of the polygraph itself must be developed to confirm which subjects told the truth and which of the subjects lied when tested.

The fatal flaw in research claiming to validate the accuracy of polygraph results springs from the way ground truth is developed in real-life situations. "Hits" are scored when a subject who tested out as deceptive on the polygraph later confesses or is found guilty by a jury. A hit must also be scored when a person deemed truthful by the polygraph is later exonerated by another's confession or is acquitted in a criminal proceeding.

The polygraph "misses" conversely, when a subject tests out as deceptive but is later cleared of suspicion by some extrinsic means. A miss must also be scored whenever the polygraph deems a suspect truthful who later confesses or is convicted by a jury.

The two types of lie detector misses, therefore, can be called "believing-a-lie" mistakes (false negatives) and "disbelieving-the-truth" mistakes (false positives). Ekman describes disbelieving-the-truth mistakes as the Othello error[8]—recall that Othello misread Desdemona's anxiety over his unfounded suspicions as the guilty anguish of an unfaithful wife. Regrettably for the innocent Desdemona, and for the quick-tempered Othello, he discovered his false positive error too late.

The believing-a-lie mistake can also be given a literary reference—The "Bad Seed" error. In W.E.M. Campbell's[9] book of that name, Mrs. Penmark steadfastly rejects to a particularly bitter end the mounting evidence that her darling little Rhoda has killed a playmate and the family handyman. Whenever Rhoda detects that her mother may be catching on to the terrible truth she asks "Mommy, what will you give me for a basket of kisses?" and Mrs. Penmark feels assured of her daughter's innocence as she answers "I'll give you a basket of hugs."

Returning now to the problem of ground truth pollution, the first taint is introduced by reliance on confessions as validating hits. The reason for this lies in the fact that the polygraph is far more effective as a confession inducer than a lie detector. Keep in mind that the testing procedure relies on convincing the subject of the polygraph's infallibility.

---

[8]Telling Lies (New York: Norton, 1985).

[9]Written under the pseudonym William March. The Bad Seed (New York: Rinehart, 1954).

Polygraph examiners are taught to gull subjects into this belief. One simplistic ploy calls for the use of a stacked deck of playing cards. The subject selects a marked card and returns it to the deck. He is then instructed to answer "no" as each of the cards is shown him by the polygrapher. Later, the polygrapher picks out the card he selected and informs the subject that the machine spotted the lie the instant he had answered "no" when shown the card during the test. While more sophisticated ploys may be used, none are anything but elaborate variations of this ancient parlor trick.

Once convinced of this polygraph's infallibility, the suspect becomes a virtual patsy for a plea bargain. If found "deceptive" by the lie detector, the innocent person is thrown into confusion. Knowing he is innocent, he nonetheless fears that no jury will believe him in the face of "infallible" lie detector results showing him to be guilty. Out of such turmoil, it must be expected that some significant percent of Othello errors (false positives) will subsequently be improperly scored as hits when innocent subjects "confess" in return for a reduced charge.

The ground truth is, of course, further tainted when an "innocent" subject is convicted by a jury influenced by false-positive polygraph results in those jurisdictions which still admit such evidence in trial court.

On the other hand, ground truth is also subject to taint by Bad Seed errors (false negatives). Bear in mind that while the polygraph lacks validity, it reliably measures one phenomenon—the suspect's level of anxiety. Obviously, the psychopathic liar feels no anxiety—he believes his own lie. Neither are most skilled, habitual liars whose long years of telling falsehoods callus over any misgivings they may once have felt. Indeed simple techniques for fooling the lie detector abound. Merely constricting the anal sphincter throws off the sensor readings as do more sophisticated biofeedback and other relaxation techniques.

Thus, some unknown number of the guilty pass their lie detector tests and later are confirmed as hits when these same favorable polygraph results contribute to their acquittal in jurisdictions where polygraph results are still admissible. Furthermore, there is no way of knowing in an absolute sense whether all so-called true positives, identified as such in their polygraph tests, should be confirmed as hits merely because they are later convicted—after all, juries have been known to err.

In sum, even under the best case scenario presented by respectable researchers in support of the polygraph's validity, the same findings show that lie detectors are strongly biased against the innocent. The most reasonable explanation for this result would seem to be that among the innocent who register high levels of anxiety when asked impertinent questions are numbered many of the most sensitive and scrupulous people in society. Anxiety and anger at being falsely accused generate exactly the same polygraphic readings as do fear of detection responses among many (but not all) of the guilty.

Established rules of scientific inquiry hold the proclaimer of a proposition responsible for proof of its validity. In regard to the polygraph, the voice-stress analyzer, and all other known forms of electronic lie-catching hardware, however, the far better scientific case has been made by those researchers who set out to prove the negative—that no universal physiological response unique to lying has yet been identified.

Despite the demonstrable lack of scientific support for the existence of a lie response, the belief persists in the lie-catching trade that superior technology will ultimately generate the level and quality of proof needed to silence the nonbelievers. This optimistic belief founders on the rocks of an elemental truth—if no such phenomenon as a universal lie response exists, it matters not how sophisticated the apparatus designed to detect it. Like fishing for bass in a bathtub, it makes no difference if one uses a canepole or the finest boron rod, the best tackle in the world cannot catch fish that are not there.

If no electronic divining rod has yet been devised to find that which is not there to be found, certainly no psychic divining rod can be any more successful. Despite this truism, the standard reporting services are replete with arbitration awards, court decisions, and tribunal rulings which appear to rely on the psychic powers of the particular fact-trier.

I call to your attention the numerous NLRB credibility resolutions which reportedly turned on the "demeanor" of the competing witnesses. Neither would you search long to find arbitration awards where some erstwhile colleague found the ever-elusive "ring of truth" in the "straightforward manner" on the witness stand of the grievant or the accusing supervisor.

Fascinated by the lore of the mysterious ring of truth, I have sought in vain these past 30 years to hear its clarion call. I have even sought out those seers who profess to hear it with consider-

able frequency in order to tune in on the proper wave length. Alas, the stock answer of those gifted mystics runs as follows: "No, I can't define the 'ring of truth,' but I sure as hell know it when I hear it."

One must be cautious in this line of criticism, however, lest the argument prove to be purely a matter of semantics. I have no quarrel with findings based on those observable aspects of a witness' behavior which tend to comport with or to conflict with the intrinsic circumstances of a case. For example, a grievant charged with assaulting a supervisor would seriously compromise his defense that he was really quite a placid fellow if he were to blow his cool in an ill-tempered outburst on cross-examination.

My criticism is directed at the arbitrator who would make some oblique reference in this case to the grievant's general demeanor, rather than stating forthrightly that the grievant's angry show of bad temper at the hearing gave credence to the charge against him. Similarly, my quarrel is with the arbitrator who notes in the award that "the testimony carries an undeniable ring of truth" rather than stating decisively that "the testimony squares with the known facts."

These are not mere literary criticisms. The consequential point is that the only valid means for detecting lies (or verifying truth) worthy of serious study is the deductive logic of the skilled and persistent fact-trier and the only reliable source of valid findings are the case facts.

I am particularly gratified in this regard to note that the Conference Manual contains an excellent article on credibility resolution by Dean Mason Ladd.[10] Fortune treated me kindly back in the late 1950s when I interned with Clarence Updegraff[11] at the University of Iowa and had Mason Ladd as a neighbor. Indeed, most of the ideas for this presentation I learned many years ago under the generous tutelage of these able men. I share these thoughts with you as they were given to me—as gifts passed on from caring friends.

Early in my practice as a labor arbitrator I would seek out these mentors for advice on seemingly intractable problems of credibility. Their message never varied—scrub away at the record; look to the case facts. Over the years of hearkening to this advice,

---

[10]*Some Observations on Credibility: The Impeachment of Witnesses*, 52 Cornell L.Q. 239 (1967).

[11]Updegraff & McCoy, Arbitration of Labor Disputes, 2d ed. (Washington: BNA Books, 1961).

I have organized the basic principles of credibility testing into an analytical framework that you may find useful. (Figure 1)

For ease of illustration, the framework can be applied to the following arbitration case—a classic example of conflicting testimony with no eyewitnesses to the disputed events.

The Hundred-Pound Heavyweight

Shortly before midnight on a warm summer night, Officer Burleigh Kopp was dispatched to investigate complaints of a noisy party of juveniles gathered in Wilderness Park. As he made his way along a trail toward their bonfire, Kopp could hear loud laughter and shouting in the distance. Suddenly, a teenaged girl stepped out of the woods and tugged at his sleeve, saying "Cool it man, nothing radical going on here." She then shouted out a warning to the group of juveniles around the bonfire ahead, "The cops are here." What happened next is almost entirely in dispute.

The Accuser's Version

Tina Tiny, 17-year-old runaway with a record of truancy, charges that Kopp swung his flashlight full force across his body and struck her in the mouth. She suffered a five-stitch laceration under the right side of her lower lip, chipped teeth, and bruised gums. The friend drove her to the police station where she filed an assault charge against Kopp.

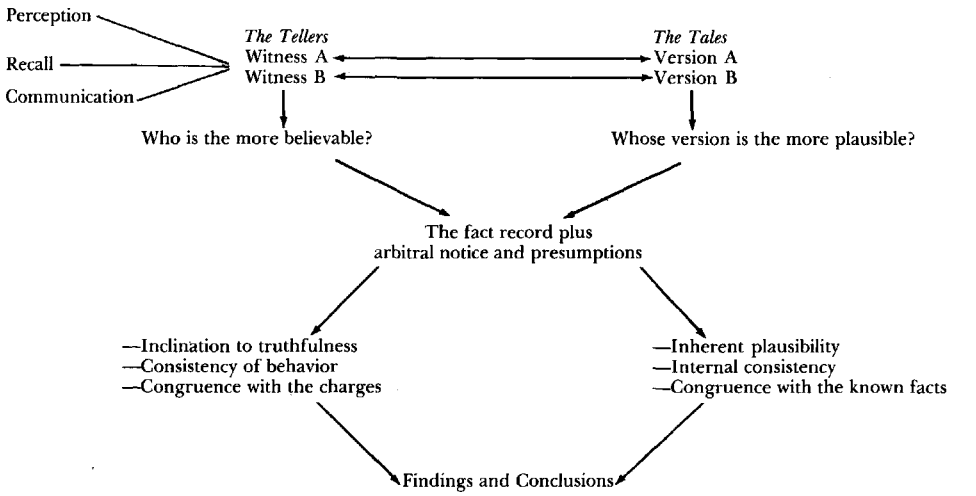
The Grievant's Version

An eight-year officer with an unblemished record, Kopp claims that he inadvertantly contacted the juvenile with the flashlight while shrugging off her grip on his sleeve. She cursed him and ran off through the woods as he proceeded to gain control of the illegal booze-and-pot party gathered at the bonfire site.

The Issue

The Police Chief ordered an investigation of the incident. On the basis of the injury report, he concluded that Kopp deliberately exercised undue force and discharged the officer. The City agrees with the Union that the issue reduces to the question of

FIGURE 1
A Framework for Resolution of Conflicting Testimony



whether the injuries were inflicted deliberately or were accidental and inadvertent. If deliberate, the discharge stands, if accidental, the discharge must be vacated.

As you can see from Figure 1 the essential elements of any credibility resolution reduce to the tellers and the tales. The credibility issue facing the trier of fact requires that both the tellers and their tales be separately considered as discrete parts of the problem. The testable questions thereby become:

1. Of the competing witnesses, who is the more believable?
2. Of the conflicting versions, which is the more plausible?

At the outset of a credibility analysis the seasoned arbitrator recognizes that neither witness, regardless of their conflicting testimony, may be "lying." People vary widely in three distinctive abilities which affect the reliability of their testimony—perception, recall, and communication. Advocates qualify their testimony so that the trier of fact can assess a witness' ability to perceive and recall facts accurately. Advocates counsel their witnesses in order that they can communicate their versions of the facts to the arbitrator effectively.

Perception

The critical foundation information must address the questions of whether the witness was in a position to see that which he reports he saw and to hear that which he claims he heard. If technical facts require expertise to be perceived correctly, the witness needs to be qualified as to training and experience. The opposing advocate may try to discredit the witness' perceptual capacities by pointing out that he wears thick glasses or has difficulty hearing questions and wears a hearing aid, or that he lacks expertise in the subject matter.

The challenge may be to the witness' mental acuity—that he was confused by fatigue or drinking, or simply that the witness is not a very bright person. Physical considerations may play a role in determining accuracy of perception—the noise level at the site of the events, the witness' sightline, dim lighting in the area, smoky or foggy conditions.

Recall

Some witnesses show a far greater ability to retrieve the perceived facts from their memory banks than others. The passage of time tends to blur the outlines of the facts and to distort the

witness' reconstruction of events. Often the "facts" drift loose from their original moorings as witnesses discuss events among themselves and challenge each other's recollection of what actually transpired. Witnesses may lose confidence in their original perceptions and conform their memories to some group consensus on the "facts." Skilled advocates seek to impress the fact-trier with their witness' ability to recall facts accurately by asking such questions as "How can you be sure this happened at 12:05?" Answer: "Because it was shortly after the lunch whistle blew," or "What makes you sure it was the Grievant you saw?" Answer: "Because he's the only guy on the crew who wears a beard."

Communication

A witness' ability to express his testimony effectively can be hampered by poor language habits or enhanced by a high level of verbal skills. Probably many cases have been won or lost on some arbitrator's mistaken notion that good grammar equates to veracity. Many a truth has been obscured by a witness' limited ability to articulate his message in a clear understandable manner.

The wisdom that credibility resolution may have nothing to do with truth or lying but rather with perception, recall, and communication is the central theme of the Japanese classic *Rashomon.*[12] Akutagawa spins his enigmatic tale of a dead husband through the confessions of a robber who claims he murdered the husband after tying him to a tree and raping his wife. The wife confesses she murdered her husband because he wished for death after the shame of his wife's violation. Finally, the murdered husband testifies through a medium that he committed suicide because his wife willingly chose the robber over himself. Each conflicting person's story squares with the extrinsic facts. Each is internally consistent in his or her several parts and all motivations are congruent with the respective versions.

This Zen tale is a favorite topic of debate among Japanese logicians and literati. Its conceptual point, according to one school, is that none of the three have "lied" but rather each has perceived the truth according to their respective needs and reported it to suit the expectations of their listeners—a vintage problem in credibility resolution.

---

[12]A film by Akira Kurosawa and Shinobu Hoshimoto based on *Vabu no Naka (In a Grove)* by Akutagawa, Ryunosuke (1921). Testimony and confession is taken from and given by various witnesses on a murder to provide conflicting accounts of what happened.

The Tests of Believability

In the case of the Hundred-Pound Heavyweight both advocates did an able job of developing a thorough and competent hearing record permitting a systematic analysis of the testimony and evidence. All relevant testable factors, therefore, could be effectively addressed.

Inclination to Truthfulness

Stated in the obverse, the propensity to lie, this factor refers to the witness' reputation as a truthful or deceitful person. A past disciplinary incident involving falsification of a time card, misrepresentation of employment history, or a perjury conviction would all bear on this consideration. Most telling would be the successful impeachment of the witness' testimony on a factual matter material to the outcome of the case.

In the case example, skilled cross-examination established that the accuser had a long record of truancy and juvenile delinquency, had been expelled from school for lying about her absences to teachers, counselors, and foster parents. Further, she denied being chemically dependent until confronted with the fact that her present address was that of a halfway house for chemically dependent adolescents.

By contrast, character witnesses portrayed Officer Kopp as a highly regarded professional police officer who served as president of the Christian Peace Officers Association and held two commendations for bravery. A stable married father of three children, he was a scout leader in his parish and was completing a degree in criminal justice studies at a community college.

The record facts, therefore, supported the grievant's reputation as a truthful person while showing the accuser as disposed to lying. There is yet another aspect to the testing of the inclination of witnesses towards truthfulness/lying—that of their motivation in the context of the case.

It came out at the hearing that the accuser had filed a five-million-dollar lawsuit against the grievant and the city. An arbitration award exonerating the grievant presumably would compromise the accuser's civil action scheduled for hearing later that same year. While some arbitrators have held that a grievant whose job is on the line may be more motivated to lie than his accuser, in this case the accuser also had a powerful interest in the outcome of the case.

Consistency

In regard to the believability of the witness, consistency refers to the degree of conformance or compatibility between a witness' testimony and his observable behavior at the hearing and/or the extrinsic facts on the record. By observable behavior, I am not referring to surface mannerisms such as nervous gestures, blushing, or sweating but rather to overt actions such as shows of ill-temper, evasiveness, or ability to recall events.

It would be inconsistent, for instance, for a witness to demonstrate exquisite recall of minor details surrounding events in dispute and then conveniently fail to remember really crucial aspects of the case. Such "selective amnesia," familiar to any seasoned arbitrator, clearly reduces the witness' believability.

In the case example, the accuser testified that she held no animosity to police officers in general, but later compromised her believability by incongruously referring to police as "pigs" in an angry slip during cross-examination. Before even considering the plausibility of the conflicting versions, therefore, it should be noted that the grievant clearly emerges as the more believable of the competing witnesses.

The Tests of Plausibility

Inherent Plausibility

This credibility factor assesses the truthfulness and accuracy of the testimony as a probability function against the world of human experience. Having swum around in the sea of human experience for some considerable span of years, seasoned arbitrators develop a reasoned capacity to array testimony along a spectrum of likelihood. A key ingredient of sound judgment is this ability to distinguish between the improbable, the possible, and the highly likely.

Rarely, can the fact-trier—in an absolute sense—capture the certain truth about past events in dispute. We have no time warp device through which we can revisit the scene and have the true events play out before our eyes. We must deal, therefore, in probabilities—favoring at the margin of credibility the likely version over the possible, while mentally discounting the improbable.

In the case example at hand, the plausibility of the accuser's version virtually collapsed when she testified that she never left

her feet when struck "full force" in the mouth with the flashlight. The record revealed that Officer Kopp was a 28-year-old athlete who stood 6 feet tall and weighed a muscular 190 pounds. A former intercollegiate wrestling champion, he still regularly bench-pressed 250 pounds.

The 17-year-old accuser, by contrast, stood slightly over 5 feet and weighed scarcely 100 pounds. The arbitrator in this case was a former university boxing coach. His experience told him that it was highly improbable—if not well-nigh impossible—for a person of the accuser's size to take a full-force blow to the face from a heavy service flashlight wielded by a man of the grievant's physical prowess and remain standing.

In his decisional thinking, the arbitrator reckoned that he had never coached a heavyweight boxer rugged enough to remain standing if struck by such a blow—hence the title of the case, "The Hundred-Pound Heavyweight." In sum, the arbitrator drew on his sea of experience to reject out of hand the accuser's version as totally implausible.

The Internal Consistency of the Testimony

When different parts of the testimony fail to hang together, the witness' version of disputed events loses plausibility just as surely as the witness whose story inconsistently changes on cross-examination loses believability. Plausible accounts tend to add up in the sense that the sum represents an integrated whole of its several parts.

In the case example, the accuser testified that she was bleeding profusely from the cut on her lip and sought out friends to take her to a doctor. Inconsistent with her professed eagerness to seek medical attention for what she claimed was a serious wound, however, she admitted taking the time to trash the grievant's squad car to the tune of over $3,000 worth of damage while he was occupied with the crowd at the bonfire. This inconsistency seriously reduced the plausibility of her claim to having been seriously injured.

Congruence With the Extrinsic Evidence

Probably the most reliable test of plausibility can be found in comparing the witness' testimony with the proven facts on record. Testimony that does not square with the known facts is definitively implausible.

The accuser's account of disputed events in the case example proved incongruent with the known facts in several regards. The swooping right-handed cross body blow she ascribed to the grievant while she was facing him would have caused injury to the left side of her mouth—the medical evidence described an injury to the right lower lip, however.

Indeed, only the grievant's version of carrying the flashlight in his left hand—and inadvertently contacting the juvenile while shrugging off her grip on his sleeve—squares with the medical evidence of injury to the right side of the face.

Adding up the scorecard in the case of the Hundred-Pound Heavyweight on all of the credibility factors examined illustrates a best case scenario for the arbitrator thirsting for confident findings and conclusions. The decision to vacate the discharge stood firm on the conclusion that the accuser proved to be a thoroughly unbelievable witness who related an utterly inplausible account of disputed events. Concomitantly the record showed the grievant to be a highly believable witness who presented a completely plausible version of what had transpired.

The worst case scenario perhaps confronts the arbitrator with two unbelievable witnesses who compete in relating conflicting but equally implausible testimony. Other variations on the theme present vexing credibility problems—including the more believable witness who relates the less plausible tale and the obverse, the less believable witness who gives the more plausible testimony.

The more complex the variation, the greater the need for critical evidentiary analysis—the more diligently must the arbitrator scrub away at the reluctant ore of the case facts to try out the precious metal of probable truth. On balance, when the more believable teller relates the less plausible tale the better part of reason should favor the teller, on the universal experience that implausible things sometimes happen to totally reliable folks.

Speaking of balance brings the discussion to the problem of equipoise where two competing and thoroughly believable witnesses relate conflicting but equally plausible testimony. When the evidentiary scales thus balance at equipoise, most arbitrators issue a "Scotch verdict," that is, the charge is dismissed as unproven. Indeed, some leading arbitrators hold that this is the

only situation where the "burden of proof" concept has any value in the arbitral setting.[13]

Improving the Odds

Before closing this discussion, a word should be said about the role of the arbitrator in improving the record. I remain an unabashed fan of the site visitation and the role-playing reenactment of physical events. We all recognize that words are not facts but rather abstractions used to describe facts—therefore words can be snares and delusions.

How does the fact-trier come to know what a witness means by "its noisy at the drop forge" without visiting the drop forge in operation? or, "it stinks in the tankage room" without experiencing that most malodorous section of the packinghouse? Even more baffling is how the arbitrator sorts out the probabilities when an accusing witness testified "He hit me on the shoulder" and the grievant testifies "I merely tapped him to get his attention." Assuming no reportable injury, it would seem important for the witness to physically demonstrate for the arbitrator what each means by "struck" and "tap."

Arbitrators who insist on site visitations are often amazed at what can be learned about eyewitness' sightlines, distances from disputed events, noise levels, and other significant details—details which often spell the difference between guesswork and a truly informed decision. Blueprints, maps, and photographs presented as evidence can be as illusionary as spoken testimony. These forms of evidence are merely two-dimensional representations of three-dimensional realities and as such may conceal more than they reveal of true fact.

A fact may have particular characteristics material to the outcome of a case, which cannot be understood without the use of the appropriate perceptual senses. I once settled a case where the union claimed the employer failed to provide prison guards their contractually promised "palatable and wholesome lunch at the prison site." I cannot fathom how an arbitrator could make an informed decision on such an issue without sampling the food provided at the guard's cafeteria. The union prevailed—the

---

[13]Decisional Thinking of Arbitrators and Judges, Proceedings of the 33d Annual Meeting, National Academy of Arbitrators, eds. James L. Stern and Barbara D. Dennis (Washington: BNA Books, 1981); *Decisional Thinking. West Coast Panel Report, Id.*, Chap. 5; Aaron, *Some Procedural Problems in Arbitration*, 10 Vand. L. Rev. 733, 741–742 (1957).

food proved well-nigh inedible. It was the same food served the prisoners—probably constituting a form of cruel and unusual punishment.

Unlike the courts, arbitration encourages the neutral to seek out material facts when the advocates fail to provide what may seem necessary for an informed decision.[14] Some arbitrators, however, hesitate to pursue information on their own initiative for fear of "making the case" for one of the parties. Discretion nourished by experience should counsel the arbitrator on the limits of the neutral's prerogatives in this regard.

The larger error, I would argue, is committed by the arbitrator who, out of an excess of caution, permits the hearing to close with substantial questions of fact remaining unanswered. The alternative to the risk of intruding with appropriate inquiries may find the arbitrator standing out in the parking lot after the hearing holding a wet finger to the breeze hoping for "Kentucky windage" to bring him an inspired answer to the dispute.

Unlike judges surrounded by all the trappings bespeaking the majesty of the law, the arbitrator can be found in the bowels of the earth checking a coal mining operation, up on the high iron of a construction job in the dead of winter, or on the killing floor of a packinghouse in the August heat. This is as it should be. Our special competence resides in the common law of the shop and must at times go where that law abides so as to understand its special character in its various settings.

If we forget that uniqueness which alone gives labor arbitration its eminence, we will most certainly become mere referees of word battles rather than adjudicators of fact in the context of shop law. Labor arbitration will then have been reduced to a mere vestibule at the front end of a continuing litigation process—rather than the final and binding forum for resolving industrial disputes.

By contrast, labor arbitration will prosper to the extent that we combine the best parts of the wisdom of law with that web of rules constructed from the self-governing system of industrial jurisprudence unique to this society. In particular regard to credibility resolution, the process will be advanced only to the

---

[14]Code of Professional Responsibility, Part 5, §A.1.b.: "An arbitrator may . . . request that the parties submit additional evidence. . . ." Arbitrator Richard Mittenthal developed this idea effectively in his paper, *Credibility—A Will-o'-the-Wisp*, in Truth, Lie Detectors, and Other Problems in Labor Arbitration, *supra* note 1, 61–74. Mittenthal's excellent paper contributed greatly to the Author's ideas on credibility resolution.

extent that arbitrators are willing to persevere at the demanding task of diligently sorting out the flyspecks from the pepper by applying the proven principles of deductive logic to the hard facts.

Given the state of research on all known forms of electronic lie-catching claptrap, I would suggest that it is premature to summon the carpenter to dismantle the jury box. Neither should arbitrators fear being rendered technologically redundant by these devices so long as we practice our craft faithfully.

In closing, may I share with you a passage from an elegant piece written by G.K. Chesterton in 1913. I was put on to this passage by my colleague Dave Lykken, who cites Chesterton to show that these parlous ideas about mechanized lie detection have been around for a very long time indeed.

> "I've been reading," said Flambeau, "of this new psychometric method . . .; they put a pulsometer on a man's wrist and judge by how his heart goes at the pronunciation of certain words. What do you think of it?"
>
> "I think it is very interesting," replied Father Brown; "It reminds me of that interesting idea in the Dark Ages that blood would flow from a corpse if the murderer touched it."
>
> "Do you really mean," demanded his friend, "that you think the two methods equally valuable?"
>
> "I think them equally valueless," replied Brown.[15]

[15]The Mistake of the Machine (1913).