

**Why Do You Always Go To Court?
Why Don't You Take the Polygraph?**

By O. D. Hamza

April 18, 2017

Abstract

Research concerning polygraph testing suggests it is around 90 percent accurate. By contrast, there are no measures indicating the accuracy of trial by judge or jury, or the JOJ system, which has existed since the Middle Ages. So, why does polygraph testing not replace the JOJ system in Canada? This article examines some of the major reasons for this in Canadian criminal law, critiques them, and suggests an alternative. This article is divided into four sections. The first section considers problems with the JOJ system. The second section considers the scientific status of polygraph evidence. The third section considers the legal status of polygraph evidence in Canadian criminal law, notably why it is inadmissible in court and whether the reasons for its exclusion are adequate. The fourth sections suggests corroborated polygraph testing (CPT) as an alternative to the JOJ system. The article concludes that CPT may be a viable alternative to the JOJ system that is modern, more cost-efficient, humane, simple, systematic, and above all more reliable.

Introduction

At a third year law school course on forensic science taught by three Crown Attorneys, two guest speakers, Sgt. Pascale Labine and Det. Andrea Giampaolo, gave a presentation on polygraph testing where they cited David Raskin's research which shows that polygraph testing is over 90 percent reliable. Labine also mentioned that trial by judge or jury (which this article will call the JOJ system, or simply JOJ) is only 50 percent reliable by contrast. (The author, however, has not found any research to vindicate this statistic.¹) According to Labine, the basic reason why polygraph testing does not replace the less reliable JOJ system to assess the credibility of witnesses is that the Supreme Court has ruled that it is the role of the JOJ to assess the credibility of witnesses. If what Labine says is true, then it may be argued that the current JOJ system is primitive, prejudicial, cruel, expensive, unscientific, artificial, and belligerent given the presence of a cheaper and more reliable alternative.

The first section of this article will discuss these problems with the JOJ system point by point. For instance, in the JOJ system lawyers ask witnesses questions and receive answers from them that may differ and which can affect the outcome of the trial. Judges and jurors can likewise come to different conclusions for different reasons, which are not predictable or even known in the case of juries, who do not have to give reasons. The reasons a judge gives may not be their true reasons (which could devolve to some bias), but whatever will pacify the parties (for instance, by depleting their financial resources) or satisfy the court of appeal. This means no two trials are the same, even of the same case.

Plus, a trial is artificial. It is an 'act,' or, as professor Scott Bushnell² and judge David Paciocco³ put it, a “morality play,” where argument and evidence play the role of a “shock

-
- 1 Bruce Spencer says it is not possible to know how accurate is the JOJ system except by comparing jurors' verdicts to judges'. See “Estimating the Accuracy of Jury Verdicts” (2007) *Journal of Empirical Legal Studies*. Other studies, like Kwangbai Park's “Estimating Juror Accuracy, Juror Ability, and the Relationship Between Them,” (August 2011) 35:4 *Law and Human Behavior* 288, similarly infer the reliability of juries from their ability to follow a judge's instructions to arrive at a legally correct judgment. The author has not, however, found any blind studies that test JOJs' ability to fact-find in a mock trial where the facts are known to the researcher, but not the JOJ, which would be an appropriate test of their fact-finding ability. (However, it will be difficult to draw general conclusions from the results since every case is different, although more research in this area will increase confidence in the results and generalizations.)
 - 2 Scott Ian Bushnell & John Desmond Morton, *Cases and Materials on Evidence* (London, ON: UWO Faculty of Law, 1971).
 - 3 David Paciocco (Welcome speech to participants in the Courthouse Mentorship Program delivered at the University of Ottawa Faculty of Law, September 2014) [unpublished].

absorber” to help the trier of fact (that is, the JOJ) make a decision. As historian of law Patrick Devlin writes: “[T]he jury system is not something that was planned on paper and has to be made to work in practice. It developed that way simply because that was the way in which it was found to work and for no other reason.”⁴ Concerning judges and juries, judge Oliver Wendell Holmes wrote: “I confess that in my experience I have not found juries specially inspired for the discovery of truth. [. . .] I have not found them freer from prejudice than an ordinary judge would be.”⁵

Part of the 'act' of a trial is for lawyers to prepare witnesses to testify in court. This is reminiscent of the musical scene “We Both Reached for the Gun” in the movie *Chicago*,⁶ where the defense attorney plays his client like a puppet to answer journalists' questions concerning her accusation of murder. Lawyers also ask witnesses loaded questions to get them to say things—or to seem to say things—they want them to say. This is reminiscent of the tap dance scene in *Chicago*⁷ where the attorney suggests the prosecution planted the evidence without proving the claim, leaving the jury with no other logical alternative but to doubt the whole system. Why, then, do we entertain such a primitive system of trial, which is neither adequate for proving facts nor laws? Oliver Wendell Holmes' answer is habit and lack of imagination: “continuity simply limits the possibilities of our imagination, and settles the terms in which we shall be compelled to think.”⁸

Professor Anthony Daimsis, an expert in trial advocacy, boasted that he could get a witness to say anything he wanted, which is not unrealistic in the author's experience. This is confirmed by the fact a trial presupposes that the JOJ could come to any conclusion. In the course of a trial, a witness can say something that will mislead the jury or give them the wrong impression. Or, a witness could make a foolish statement he or she cannot retract that can ruin his or her reputation. (Indeed, the very fact of being involved in a litigation can tarnish a person's reputation.) The author hence concluded he would never want to be a witness in any situation, because of the sheer responsibility and hazard it involves. As judge Learned Hand concluded: “I must say that as a litigant I should dread a law suit beyond almost anything short of sickness and death.”⁹

In short, the current JOJ system presupposes that a trial is sensitive to the actors—that is, the JOJ, the lawyers, and the witnesses—and the particular evolution of the 'act.' In principle, this ought not to be the case, since the facts and the law should determine the outcome, which should always be the same for the same facts and laws. If polygraph testing can render the outcome of a trial more accurate and uniform—that is, more scientific—it is definitely an alternative to consider. Hence, the second section of this article will consider the scientific status of polygraph testing, notably based on the research of David Raskin and others.

The third section of this article will examine the legal status of polygraph testing in Canadian criminal law, notably why the Supreme Court ruled this evidence to be inadmissible in major cases such as *R v Béland*¹⁰ (1976), *R v Phillion*¹¹ (1978), and *R v Mohan*¹² (1994). Criminal law cases were chosen because they are federal law that applies throughout Canada, as well as because a higher standard of clarity and rigor applies to this area of law, which requires proof

4 Patrick Devlin, *Trial by Jury* (London: Methuen, 1966).

5 Oliver Wendell Holmes, “Law in Science and Science in Law” (1899) 12(7) HLR 443 at 460.

6 *Chicago*, film, directed by Rob Marshall (USA: Miramax, 2002). See “Chicago - We Both Reached For the Gun” (19 March 2012), online: YouTube <<https://www.youtube.com/watch?v=C9dFKRZ8EbU>>.

7 *Ibid.*

8 Holmes, *supra* note 5 at 444.

9 Learned Hand quoted in Jerome N Frank, “Some Reflections on Judge Learned Hand” (1957) 24 U Chi L Rev at 666 at 675.

10 *R v Béland*, [1987] 2 SCR 398 [*Béland*].

11 *Phillion v R*, [1978] 1 SCR 18 [*Phillion*].

12 *R v Mohan*, [1994] SCJ No 36, 1994 CarswellOnt 66 (SCC) [*Mohan*].

beyond a reasonable doubt (unlike private law, where the standard of proof is a balance of probabilities).

The third section will also critique the law concerning polygraph evidence. For instance, while the Supreme Court ruled that polygraph evidence is inadmissible for various legal reasons, such as its violating the rules of evidence and impinging on the function of the jury, it provides no empirical evidence that polygraph testing delivers unreliable results. If polygraph testing is a more reliable method of fact-finding than witness examination and judgment by a JOJ, then it stands to reason that it should replace them. This section will argue that the legal principles the Supreme Court invokes to exclude polygraph evidence are not ends in themselves, since the search for truth is an overriding value of the justice system (as s. 24(2) of the *Charter*, for instance, implies). While the Supreme Court raises the legitimate point that polygraph evidence should be excluded because its appearance of infallibility (the “mystique of science”¹³) may prejudice the trier of fact, this may be countered with two arguments. First, the mystique of science presupposes the evidence is not scientific or that the jury is too ignorant to understand the evidence. Second, even if the evidence is scientific, but not very strong, albeit stronger than any unscientific alternative, it is worth considering the argument of professor Frederick Schauer that “Bad science is worse than good science, but not necessarily worse than the non-science that lurks in the heads of judges and jurors.”¹⁴

The fourth section of this article proposes corroborated polygraph testing (CPT) as an alternative to the JOJ system. This section explains the theory and methodology of CPT, which includes the Corroboration Rule and logical empiricism, particularly the philosophy of indirect verification. The section further discusses advantages of CPT over the JOJ system. For instance, CPT is not sensitive to variations in the conditions of a trial, such as the actors or players or the evolution of the a trial; it also asks relevant and comparison questions, without going into tangential storytelling, argumentation, or cross-examination; it is also subject to re-testing (*i.e.*, a person can challenge the results by taking another test); and it has to be corroborated by other independent material evidence (IME) in order to count as evidence or establish a fact.

This article concludes that CPT may be a viable alternative to the JOJ system that is modern, relatively inexpensive, humane, simple, systematic, and above all more reliable.

I. Problems with the JOJ system

This section will examine problems with the JOJ system, namely that it is primitive, cruel, expensive, complicated, unscientific, artificial, and belligerent. As mentioned in the introduction, the JOJ system is better explained as a “morality play,” which serves a twofold purpose. First, it acts as “shock absorber”—to borrow the words of professor Bushnell—for the decision of the JOJ.¹⁵ Second, it gives the community the impression that some justice is being done, as chief justice Hewart put it in 1924: “it is [. . .] of fundamental importance that justice should not only be done, but should manifestly and undoubtedly be seen to be done.”¹⁶

Historian of law Patrick Devlin sums up the purpose of the JOJ system thus: “the jury system is not something that was planned on paper and has to be made to work in practice. It developed that way simply because that was the way in which it was found to work and for no other reason.”¹⁷ This goes as much to the credit of the JOJ system as its discredit. As Oliver Wendell Holmes writes: “I believe that the claim of our especial code to respect is simply that it

13 *Béland, supra* note 10.

14 Frederick Schauer, “Can Bad Science Be good Evidence? Lie Detection, Neuroscience and the Mistaken Conflation of Legal and Scientific Norms” (2010) 95 *Cornell Law Review* at 36.

15 Bushnell & Morton, *supra* note 2.

16 *R v Sussex Justices, Ex parte McCarthy* ([1924] 1 KB 256, [1923] All ER Rep 233.

17 Devlin, *supra* note 4 at 4-5.

exists, that it is the one to which we have become accustomed, and not that it represents an eternal principle.”¹⁸ However, Holmes believes an aspiration toward an eternal principle, namely science, ought to be maintained.¹⁹ While polygraph testing may be less than an ideal solution, it constitutes a better alternative than the current system, as the following subsections will elucidate.

A. Primitive

The JOJ system is primitive, as it has been around since the Middle Ages. This could be taken to mean that the system works, even if it is not accurate. For instance, as American district attorney Vincent Bugliosi put it: “I have yet to see the man who can convince twelve reasonable men and women [. . .] that black is white and white is black.”²⁰ However, Bugliosi's remark seems to ignore reams of research that proves the contrary, *viz.*, that it is quite common and indeed the historical norm for any twelve reasonable men and women to be deluded. Philosopher Garret Merriam defends the notion that people are wrong most of the time in a lecture entitled *You Have No Idea How Wrong You Are*:

The question what is human nature is a question philosophers have been asking as long as there have been philosophers. So, I want to propose an answer to that question today. I want to say that human beings are the species that got it wrong. Now, not exclusively of course. [. . .] But no other species makes mistakes as frequently as we do about as broad an array of topics as we do or with as consequential outcome as we do. If you look at human history, everywhere you look you see mistake after mistake after mistake. It doesn't really matter which branch of history you look at: you can look at philosophy, science, medicine, mathematics, psychology, economics, religion, politics; and everywhere you turn people are making mistakes up, down, left, right, and center. Historically speaking, people have been wrong far more than they have been right. [. . .] This shouldn't surprise us, however, because there are a lot more ways to be wrong than there are ways to be right.²¹

The fact the JOJ system is ancient does more to show its fatal flaws in different historical settings than it does to show it is the 'old faithful' its proponents claim it to be. For instance, it is worth considering that this system is more or less the same that axed thousands of heads in England during the 18th c. under the 'Bloody Code,' which made just about every criminal offense a capital one. The infamous Salem Witch trials also used the same process. History does, however, suggest that the JOJ system is a transitional phase in the evolution of justice, as judge Oliver Wendell Holmes observes:

Who could fail to be interested in the transition through the priest's test of truth,' the miracle of the ordeal, and the soldier's, the battle of the duel, to the democratic verdict of the jury! Perhaps I might add, in view of the great increase of jury-waived cases, a later

18 Holmes, *supra* note 5 at 460.

19 *Ibid* at 462-463.

20 See closing arguments and verdict in Vincent Bugliosi, *The Trial of Lee Harvey Oswald*, two-part television film (USA: ABC-TV, September 1977). See “On Trial: Lee Harvey Oswald” (part 23) (February 2013), online: YouTube <http://www.youtube.com/watch?v=JDqBjDoOy_A>.

21 Garret Merriam, “You Have No Idea How Wrong You Are” (Lecture delivered at the Nammour Philosophy Symposium On Being Wrong, April 26-27, 2011) [unpublished], online: Center for Practical and Professional Ethics <http://www.csus.edu/cppe/nammour_2011_beingwrong.html>. See also Sisyphus Redeemed, “You Have No Idea How Wrong You Are” (25 May 2011) online: <<https://www.youtube.com/watch?v=E8V8rtdXnLA>>.

transition yet - to the commercial and rational test of the judgment of a man trained to decide.²²

Interestingly, some scholars have contested Holmes' suggestion that the JOJ system is a distinct transitional phase from trial by combat. Professor Ian Bushnell, for instance, argues that trial by combat never left the West, but evolved into something more brutal: trial by argument.²³ American defense attorney Francis Bailey—famous for defending O.J. Simpson—agrees with this view when he observes that litigation is “the true substitute for gladiatorial combat.”²⁴ He also expresses acute skepticism in regard to the efficacy of this process at finding truth when he remarks that “Those who think the information brought out at a criminal trial is the truth, the whole truth, and nothing but the truth are fools. Prosecuting or defending a case is nothing more than getting to those people who will talk for your side, who will say what you want said.”²⁵

Professor Patrick Devlin argues that the English jury is not what it is because a lawgiver decreed it so; rather, “It developed that way simply because that was the way in which it was found to work and for no other reason.”²⁶ The notion that a lawgiver invented this system is inconceivable according to Devlin. Devlin agrees with Holmes' observation that we are accustomed to the JOJ system; as a result, we are bound to it by habit and lack of imagination or motivation to think or do otherwise.²⁷ Writes Devlin:

We are used to it and know that it works; if we were not, we should say that it embodies a ridiculous and impracticable idea. Consider what the idea is. Twelve (why twelve?) men and women are to be selected at random; they have never before had any experience of weighing evidence and perhaps not of applying their minds judicially to any problem; they are often, as the Common Law Commissioners of 1853 tactfully put it, “unaccustomed to severe intellectual exercise or to protracted thought.” The case may be an intricate one, lasting some weeks and counsel may have in front of them piles of documents, of which the jury are given a few to look at. They may listen to days of oral evidence without taking notes—at least, no one expects them to take notes and no facility is provided for it in the jury-box, not even elbow room. Yet they are said to be the sole judges of all the facts. At the end of the case they are expected within an hour or two to arrive at the same conclusion. Without their unanimous verdict no man can be punished for any of the greater offences. Theoretically it ought not to be possible to successfully enforce criminal law by such means.²⁸

In short, the JOJ system only 'works' in the sense that we are used to it and have little motivation to change it. This does not mean there is no *reason* to change it. The system could be absurd and dysfunctional; but at least we are used to it and may be too lazy or carefree to change it. That we are used to the system and too lazy or carefree to change it does not mean it actually works to our benefit, nor does it mean that it actually achieves the transcendental goal of truth or justice. The systems that 'work' best in this sense are the most ancient and primitive systems. For instance, in many parts of Ghana witches are tried—as they have been tried for thousands of years—by slaughtering a chicken; if its feet point up, she is guilty and imprisoned for life; if they

22 Holmes, *supra* note 5 at 445.

23 Bushnell & Morton, *supra* note 2.

24 Francis Bailey quoted in “Bailey, Francis Lee” in *West's Encyclopedia of American Law* (USA: The Gale Group, Inc., 2005), online: Encyclopedia.com <<http://www.encyclopedia.com/law/encyclopedias-almanacs-transcripts-and-maps/bailey-francis-lee>>.

25 Francis Bailey quoted in the *New York Times* (20 September 1970).

26 Devlin, *supra* note 2 at 4-5.

27 Holmes, *supra* note 5 at 460.

28 Devlin, *supra* note 2 at 4-5.

point down, she is innocent and imprisoned all the same in case some members of the village still believe she is guilty.²⁹ There is no more reason to think that the JOJ system has persisted on account of its virtues than there is to think that the Ghanan system has. That the system 'works,' in short, does not rationally justify it.

B. Cruel

The JOJ system is cruel to everyone involved in it, except the legal actors (the judge and lawyers). Jurors are interrupted from work and sequestered from their families and society for weeks or months, sometimes years. Witnesses are asked probing questions in public for days or weeks. The JOJ system is especially cruel to witnesses with mental disorders, of which low IQ may be counted as an example. Such witnesses comprise the majority of people involved with the justice system, and are particularly vulnerable to the 'mind games' of clever and prepared lawyers, who are likely to manipulate them. Witnesses with mental disorders are also most likely to say or do 'stupid' things in court that can unjustly undermine their credibility in the eyes of the JOJ or raise suspicions against them.

C. Expensive

The JOJ system is expensive due to the high cost of lawyers and judges, difficulty to predict the process and outcome (which can become very protracted), as well as the general operation of the court system. For instance, the *Canadian Lawyer's* 2015 Legal Fees Survey estimates the national average cost for a two day trial at \$31,330,³⁰ which is shy of the median total individual income in Canada in 2014 of \$32,790.³¹

A more efficient and accurate system to process evidence could greatly reduce costs and unpredictability of the law, process, and outcome for lawyers and clients. This would in turn enable more rapid case resolution and increase the number of cases lawyers and judges can process, which would reduce backlog.

D. Unscientific

The principal problem with the JOJ system is that it is not scientific. Those who believe this system is accurate tend to believe, like Bugliosi who was quoted earlier, that it is impossible to fool twelve reasonable men and women.³² However, most of history and a wealth of psychological research shows that twelve reasonable men and women can believe just about anything.³³ Indeed, a trial presupposes that the JOJ could believe one party or the other. If the outcome were predetermined, then there would be no point in a trial.

It is largely because of the indeterminate nature of a trial—and its cost, length, and hazards—that the author shares judge Learned Hand's conclusion that “I must say that as a

29 *The Witches of Gambaga*, documentary film, directed by Yaba Badoe (Fadoa Films, 2010).

30 Michael McKiernan, “The Going Rate: The 2015 Canadian Lawyer Legal Fees Survey Shows Litigation Fees Have Returned to Pre-Recession Rates” (June 2015) *The Money Issue*, online: <www.canadianlawyermag.com>.

31 Statistics Canada, CANSIM, table 111-0008, “Individuals by total income level, by province and territory (Canada)” (14 July 2016), online: <<http://www.statcan.gc.ca/tables-tableaux/sum-som/l01/cst01/famil105a-eng.htm>>.

32 Bugliosi, *supra* note 20.

33 See, e.g., Kerry Kawakami *et al.*, eds, *Journal of Personality and Social Psychology*, online: <<http://www.apa.org/pubs/journals/psp/>>. See also Valerie P Hans *et al.*, eds, *Journal of Empirical Legal Studies*, online: <<http://onlinelibrary.wiley.com/journal/10.1111/%28ISSN%291740-1461>>.

litigant I should dread a law suit beyond almost anything short of sickness and death.”³⁴ A further reason for this is that

the trial judge's belief about the facts, based on his choice of trustworthy witnesses, results from subjective, undisclosed, un-get-at-able, private factors in the trial judge, i.e., his prejudices, for or against witnesses who wear glasses, or who are women, or blond women, or fat women, or Irishmen, Englishmen, Jews, Catholics, negroes, or red-headed men, or men with facial tics or who speak with a foreign accent or with a Southern drawl, etc., etc. It would be bad enough were those prejudices merely knowable by others. But the situation is worse than that: most of these prejudices are unconscious. That is, they are unknown to the trial judge himself.³⁵

Professor Jerome Frank comments that this “problem will never be solved so long as it is ignored, shoved under the rug, as it has been and still is, by judges, lawyers, and law professors. Moreover, if the problem is insoluble, we but delude ourselves by disregarding its existence.”³⁶ Frank proposes to psychoanalyse judges to evince some of their biases, although he recognizes this is a limited solution. However, it can help the process of judgment become more democratic by comparing judges' biases to those of the community. Learned Hand describes the problem of judicial bias in more detail as follows:

[A]lthough we speak of a trial judge “finding” the facts of a case, in truth he can't “find” them, ready-made, waiting to be discovered. The relevant events happened in the past, before any law suit began. The judge, therefore, never sees them happening. The facts, for purposes of judicial decision, are not what really happened, only what he believes—or guesses—to have happened. He forms this belief, or guess, by listening to witnesses who purport to have observed those past events when they occurred. Often the witnesses tell conflicting stories. Some of them lie; some are honestly mistaken in their original observations or memory. The judge “finds” the fact by choosing which witnesses he will believe. On this choice—the exercise of his fact-discretion, as I call it—depends the fate of the litigants. For in most suits, the entire or principal part of the dispute between the parties relates to the facts, not to the applicable legal rules, so that most law suits could be labeled “fact suits.” Usually, for reasons I'll state later, the trial judge's fact-finding is accepted by the upper court on an appeal, if there is an appeal. Because of the incontestable mistakes in the trial judge's findings (or guesses), innocent men go to jail, and other men suffer defeat in civil law suits they ought to win, thus losing their property, their fortunes, their jobs or their reputations. On that account, the work of the trial judge in determining the facts—that is, in choosing some rather than other evidence as reliable—represents the most important part of the judicial process. And note that the decisions, in most law suits not yet commenced, can seldom be predicted because the trial court's findings of fact cannot be prophesied. No reading of precedents, no study of legal treatises, can overcome such obstacles to prediction.

This job of trial-court fact-finding is far more difficult, if well done, than that of making and applying legal rules. Yet it shows up, scarcely at all, in law school study of the courts, a study confined chiefly to examination and criticism of the legal rules as expounded in judicial opinions. Accordingly, in law school, students learn almost nothing about those trial judges who are truly great in performing the task of finding the facts.

The most significant factor in trial court fact-finding is this: no one has ever been

34 Hand, *supra* note 9 at 675.

35 *Ibid* at 677-678.

36 *Ibid* at 678.

able to contrive any rules to guide the trial judge in making his choice of the believable witnesses. That choice, that exercise of fact-discretion, therefore escapes the discipline of rules. It is un-ruly. When the witnesses testify orally in the presence of the trial judge, as ordinarily they do, he is guided, in very considerable measure, by his observation of the witnesses' demeanor—their facial expressions, falterings, or glibness. As I once said: “the demeanor of an orally-testifying witness is 'always assumed to be in evidence.' It is 'wordless language.' The liar's story may seem uncontradicted to one who merely reads it, yet it may be 'contradicted' in the trial court by his manner, . . . his grimaces, his gestures, and the like—all matters which 'cold print does not preserve' and which constitute 'lost evidence' so far as an upper court is concerned.”³⁷

As Hand put it: “The whole nexus of sight and sound . . . is lost in a written record. . . . The words that a witness utters . . . are again and again of no probative weight at all because of address, his bearing and his apparent lack of intelligence.”³⁸ Moreover, the trial judge's finding of fact is “unassailable” by the higher courts, since they cannot observe the demeanor of the witnesses.³⁹ In the words of Hand: “The facts are cooked when a case reaches an appeal court.”⁴⁰ This is true even today in Canada, as Brian Dickson CJ confirmed in *Lensen v Lensen*:⁴¹ “It is a well-established principle that findings of fact made at trial based on the credibility of witnesses are not to be reversed on appeal unless it can be established that the trial judge made some “palpable and overriding error which affected his assessment of the facts.”⁴² Hand concluded:

Our courts [. . .] set great store by demeanor evidence. But you see how tricky is this reliance on demeanor evidence: one trial judge will interpret it one way, another judge, another way. There exists no common or objective measure usable in choosing the reliable witnesses.⁴³

Overall, Hand raises three points concerning how the JOJ system works which are very important—and true—to this day, *viz.*, (1) judges and juries find facts on a discretionary basis; (2) lawyers, judges, and juries are not educated in fact-finding; and (3) there are no rules of fact-finding. Concerning juries, Holmes observed:

I confess that in my experience I have not found juries specially inspired for the discovery of truth. I have not noticed that they could see further into things or form a saner judgment than a sensible and well trained judge. I have not found them freer from prejudice than an ordinary judge would be. Indeed one reason why I believe in our practice of leaving questions of negligence to them is what is precisely one of their gravest defects from the point of view of their theoretical function: that they will introduce into their verdict a certain amount—a very large amount, so far as I have observed—of popular prejudice, and thus keep the administration of the law in accord with the wishes and feelings of the community.⁴⁴

It is further worth considering that trials depend on how the JOJ assesses the credibility of

37 *Ibid* at 675-677.

38 *Ibid* at 677.

39 *Ibid*.

40 *Ibid*.

41 *Lensen v Lensen*, [1987] 2 SCR 672. This principle is confirmed in *Housen v Nikolaisen*, [2002] 2 SCR 235, 2002 SCC 33 at paras 12, 13, 20, and 24, *i.a.*

42 *Ibid*.

43 Hand, *supra* note 9 at 677.

44 Holmes, *supra* note 5 at 460.

witnesses: that is, who appears to be more trustworthy. In Canadian law, there is a presumption of credibility, which is assessed after evidence is presented⁴⁵ based on what is said, how, and by whom.⁴⁶ What is said can only be assessed based on its consistency and verisimilitude (the 'ring of truth'). (Hand explains this is because the JOJ cannot observe past events.) Character and demeanor are residual considerations. Bastarache and Abella JJ make the point in *R v Gagnon*⁴⁷ that

Assessing credibility is not a science. It is very difficult for a trial judge to articulate with precision the complex intermingling of impressions that emerge after watching and listening to witnesses and attempting to reconcile the various versions of events.⁴⁸

The Federal Court further cautions that “[t]here can be no consistency on findings of credibility”⁴⁹ between different cases.

While the assessment of credibility is not a science, it certainly pretends to be so when a JOJ declares a person to be guilty or innocent and thus deprives him of his property or liberty. McLachlin CJ confirms that credibility assessment is not a science when she writes in *R v REM*:⁵⁰

While it is useful for a judge to attempt to articulate the reasons for believing a witness and disbelieving another in general or on a particular point, the fact remains that the exercise may not be purely intellectual and may involve factors that are difficult to verbalize. Furthermore, embellishing why a particular witness’s evidence is rejected may involve the judge saying unflattering things about the witness; judges may wish to spare the accused who takes the stand to deny the crime, for example, the indignity of not only rejecting his evidence and convicting him, but adding negative comments about his demeanor. In short, assessing credibility is a difficult and delicate matter that does not always lend itself to precise and complete verbalization.⁵¹

Estey J lists as guides to the assessment of a witness' credibility “The general integrity and intelligence of the witness, his powers to observe, his capacity to remember and his accuracy in statement,” as well as “whether he is honestly endeavouring to tell the truth, whether he is sincere and frank or whether he is biased, reticent and evasive.”⁵²

In *Faryna v Chorny*⁵³ O'Halloran JA notes, however, that “A witness by his manner may create a very unfavourable impression of his truthfulness upon the trial judge and yet the surrounding circumstances in the case may point decisively to the conclusion that he is actually telling the truth.”⁵⁴ It follows from this observation that credibility assessment is at best of no assistance to justice (as its relevance presupposes a lack of evidence besides a witness'

45 See points 1 and 2 in *White v The King*, [1947] SCR 268 at 274 [*White*], quoting Beck J.

46 Ian Binnie, “The Changing Role of the Expert Witness” (Paper delivered at the International Association of Procedural Law in Canada, Toronto, 2009) at 8. Emphasis by Binnie J.

47 [2006] SCC 17 (CanLII).

48 *Ibid* at para 20. See also “Assessing Credibility is not a Science” (10 February 2015), online: <<http://labourlawblog.typepad.com/managementupdates/2015/02/assessing-credibility-is-not-a-science.html>> for a brief and informative survey of Canadian law on witness credibility assessment.

49 See Refugee Protection Division, *Assessment of Credibility in Claims for Refugee Protection Legal Services of the Immigration and Refugee Board* (31 January 2004), s 1.2, online: <<http://www.irb-cisr.gc.ca/Eng/BoaCom/references/LegJur/Pages/Credib.aspx#note12>>.

50 [2008] SCC 51 (CanLII).

51 *Ibid* at para 49.

52 See *White*, *supra* note 45.

53 [1951] 4 WWR (NS) 171 (BCCA).

54 See *Raymond v Bosanquet Tp.* [1919] CanLII 11 (SCC) at 460.

demeanor) and at worst misleading. This is apparent in the case of *Chiu v Chiu*,⁵⁵ a lawsuit for damages from a car accident, where Romilly J cites the mother's breaking into tears—a classic—to support his finding that she is credible:

The mother of the plaintiff gave her evidence in a very forthright manner. When asked to describe what Jason was like since the accident she broke into tears. When she returned to the stand after many other expert witnesses had testified she was again asked the same question. At that stage she asked that Jason be excused from the court room and again broke down into an uncontrollable sobbing when she attempted to describe the plight of her son, the plaintiff, since the accident. In my view, this could not have been rehearsed.⁵⁶

In *R v NS*,⁵⁷ a recent case of national significance, McLachlin CJ wrote the following for the majority concerning the witness' wearing an Islamic full-face veil or niqab at the stand:

There is a deeply rooted presumption in our legal system that seeing a witness's face is important to a fair trial, by enabling effective cross-examination and credibility assessment. The record before us has not shown this presumption to be unfounded or erroneous. However, whether being unable to see the witness's face threatens trial fairness in any particular case will depend on the evidence that the witness is to provide. Where evidence is uncontested, credibility assessment and cross-examination are not in issue. Therefore, being unable to see the witness's face will not impinge on trial fairness. If wearing the niqab poses no serious risk to trial fairness, a witness who wishes to wear it for sincere religious reasons may do so.⁵⁸ [Underlines added.]

The underlined portions—that seeing a witness' face is important and that 'the record' has not refuted this presumption—raise the question to what record the Supreme Court is referring: the legal record where it is the normal method of credibility assessment or the scientific record where it is a condemned pseudoscience called physiognomy?

Another unscientific aspect of the JOJ system is that every trial is unique. Different trials of the same case can produce different outcomes, as was the case in the O.J. Simpson civil and criminal trials. Footprints of a rare brand of Bruno Magli shoes were found at the scene of the crime. O.J. denied ever wearing such shoes, although he was photographed wearing them. This photograph, which was produced at his civil trial, was not produced at his criminal trial. Nor was the theory that police framed O.J. presented at the civil trial or detective Mark Fuhrman (who was in charge of the investigations and accused of racism) summoned thereat. Justice, however, requires that the verdict should always be the same and based on a monotonic inference, which means that evidence sufficient to warrant a guilty verdict must be such that more, less, or different evidence cannot change it.

The JOJ system also fosters all kinds of logical fallacies, as it is not constrained by the rules of logic or science (which do not always overlap with the rules of evidence). A particularly common fallacy in litigation goes by various names, including kitchen sinking, crap-shooting, the shotgun effect, or kettle logic. The following statement by Bugliosi exemplifies this fallacy:

My state of mind would have been the realization that if just one of the twelve jurors would have bought the police conspiracy argument, it's a hung jury, and the prosecution would have had to try the case all over again. As to that potential one or more jurors, I

55 *Chiu (Guardian ad Litem of) v Chiu*, [1999] 5633 (BCSC) at para 6.

56 *Ibid.*

57 *R v NS*, 2012 SCC 72, [2012] 3 SCR 726.

58 *Ibid.*

could never know which of my anti-conspiracy arguments might resonate with them. Eight of them may not have, but the ninth might have.⁵⁹

The famous American trial lawyer Irwin Younger made the following joke about this way of thinking or arguing:

[A] farmer has a patch of cabbages. His neighbor has a goat. The goat gets through the fence, gets into the cabbage patch, and eats the cabbages. The farmer brings a lawsuit against his neighbor for damages, alleging that the defendant's goat ate his cabbages.
 “You didn't have any cabbages!
 “If you did, they weren't eaten!
 “If your cabbages were eaten, they weren't eaten by a goat!
 “If they were eaten by a goat, it wasn't my goat!
 “If they were eaten by my goat, he was insane at the time!”⁶⁰

Why, one may ask, is it so important that the trial process should be scientific? The simple reason is that a person should not be deprived of his or her property or liberty on anything but scientific evidence. It is in these matters that science matters most. According to the received view among scientists and philosophers of science, a scientific theory is definite⁶¹ and empirically verifiable and falsifiable.⁶² The JOJ system is neither of these things. This is apparent in its reliance on subjective assessment of the credibility of witnesses and the fact that re-trying the same case can lead to different outcomes. This is why the Fifth Amendment of the American Constitution uses the macabre designation of “double jeopardy of life and limb” for a second trial. As the *NS* case highlights, the JOJ system basically devolves to an exercise of physiognomy (a textbook example of pseudoscience) by people who are not even physiognomists. (At least physiognomists tried to be scientific.)

The sheer randomness of the JOJ system has prompted a group of physicists to attempt to better predict the votes of US Supreme Court (USSC) judges than legal experts by using statistical mechanics (which is used to predict stochastic systems). The model of Roger Guimera and Marta Sales-Pardo, for instance, predicted the USSC votes 83 percent of the time, compared to the roughly 50 percent rate of legal experts.⁶³ Edward Lee *et al*, on the other hand, used an Ising model (comparing judges' positive or negative votes to magnetic dipole atomic spin) to predict USSC judges' votes with similarly accurate results. Lee *et al* compare USSC judges' votes to a randomly fluctuating “energy landscape”⁶⁴ which is predictable only because it

59 See Vincent Bugliosi, *Absolutely 100% Guilty*, film, (USA: Platinum Productions, 1999). The film is a video extension of his book, *Outrage* (USA: W W Norton & Company, 1996) on the O.J. Simpson murder case. See “Vincent Bugliosi vs. O.J. Simpson (“Absolutely 100% Guilty”) (1999) (Part 1)” (12 February 2017), online: YouTube <<https://www.youtube.com/watch?v=6vonDXxuZXY>>.

60 Quoted in Steve Schroeder, *The Lure: The True Story of How the Department of Justice Brought Down Two of the World's Most Dangerous Cyber Criminals* (USA: Course Technology, 2012) at 179.

61 The video of this lecture is available online. See nonstampcollector, “Feynman on Scientific Method” (18 February 2011) online: YouTube <<https://www.youtube.com/watch?v=EYPapE-3FRw>>.

62 Richard Feynman, one of the most eminent physicists of the 20th c., defined science in these terms in his famous 1965 Messenger Lectures at Cornell University. See “Feynman on Scientific Method” (18 February 2011), online: YouTube <<https://www.youtube.com/watch?v=EYPapE-3FRw>>. Karl Popper, one of the most eminent philosophers of science of the 20th c. defined science in similar terms in his 1953 lecture “Science: Conjectures and Refutations.” See “Philosophy of Science: a Personal Report” in *British Philosophy in Mid-Century* (UK: 1957), s 1.

63 Roger Guimerà & Marta Sales-Pardo, “Justice Blocks and Predictability of U.S. Supreme Court Votes” (2011) PLoS ONE 6(11): e27188, online <<https://doi.org/10.1371/journal.pone.0027188>>.

64 Edward D Lee et al, “Statistical Mechanics of the US Supreme Court” (June 2013) ARXIV, eprint

exhibits a certain degree of uniformity on the whole.

To conclude, trial and judgment under the JOJ system are fundamentally flawed. Some of the reasons for this may be summarized as follows. First, a trial can produce different outcomes under slightly different circumstances, even though the facts are supposed to be the same. Second, the JOJ's opinion concerning the credibility of witnesses is useless, not only because it is irrelevant, but also because it is unreliable. There is no reason to think that any number of judges or *12 Angry Men*⁶⁵ have magic powers to assess the credibility of witnesses, which is what judgment devolves to in the absence of scientific evidence. (Woe to him who has not Henry Fonda on his jury!) A person should not go to jail over reasons juries do not have to give, or reasons a judge may give that cannot be published in a science journal. Absent scientific proof, judgment is cleromancy. It is empty arguments, allegations, and speculation that wastes time and resources, where a coin toss, or even slaughtering a chicken in the Ghanan fashion, would be more economic and less judgmental.

E. Artificial

The JOJ system is artificial, meaning the entire process is fake or rehearsed. Evidence in a resolution negotiation (which comprises 95 percent of cases) is an empty threat. By contrast, evidence in a trial (which comprises the remaining 5 percent of cases) is synthetic, as it is mostly a product of the trial as it develops. Mooting competitions are a clear example of this. Participants in a moot competition are given an artificial case to argue in front of an artificial panel of judges, lawyers, or jurors. (The author, for instance, has participated in moots where lawyers or members of the public have acted as judges or jurors.) Thus, the moot teams are given the same 'initial conditions' to argue. In the course of a moot competition, one will witness a variety of evolutions of the case with different players or participants (namely, the different moot teams) and actors (namely, the different witnesses and judges or jurors). Yet, in principle the outcome of the trial should be the same in all cases and should not be influenced by the performance of the players or actors. Thus, while moot trials are moot trials, they are not moot experiments, but actual ones that prove the JOJ system does not work, but favors whomever the judges or jurors happen to favor for whatever reasons. (Contradictory evaluations are the norm in moot competitions.)

Mooting competitions show that a trial can evolve in many different ways and that it is not possible to predict how (otherwise, there would be no need for a trial in the first place). This implies that evidence is not just what is brought into the trial, but also a product thereof. That is to say, evidence is not merely what is *presented* at a trial, but what is *created* therein. There is a paucity of research on how trials create evidence, which is a ripe area for research to answer such important questions as how cross-examination can be used to get a witness to say—or appear to say—just about anything, or how abbreviation, embellishment, and negative or positive spin, can be used to tell completely different stories or versions of the same story at direct examination. Such things are known to practitioners, but are not part of an official legal canon or curriculum. Research on these subjects might not only help practitioners who may use these tactics (which are not all dirty), but can also expose their methodological flaws or 'dirty' aspects, which could lead to further research and provide additional impetus to improve the system.

For instance, it is an elementary point of science that talk proves nothing. If methods such as direct and cross-examination are supposed to be necessary, this presupposes that there is no material evidence to work with and that the only way to judge the case is to judge the witnesses as reliable or not based on their presence and testimony in court. That facts can be determined this way is a baseless assumption on which the JOJ system is built. This simply means the

arXiv:1306.5004.

65 *12 Angry Men*, film, directed by Sidney Lumet (USA: MGM Studios, 1957).

'players' in the justice system (the lawyers and their witnesses) must 'put up an act' to make up for the lack of facts.

Other fake aspects of the JOJ system may not be as easily apparent to the public as the aforementioned. For instance, most members of the public are probably not aware—as the author was not aware before studying law—that lawyers prepare their witnesses before they go to court. Telling or hinting to a witness what or what not to say (what some practitioners call 'wood shedding') or confirming what they are going to say in court can mislead the jury to think the witness is testifying spontaneously. (From the vantage of evidence law, this means the jury will conflate prepared statements for *res gestae*, which do not carry the same evidential weight.) In short, preparation for trial is preparation for an act, which is not a fact. Rather, the act creates the facts and not the other way around.

F. Belligerent

The JOJ system is an adversarial system, which is really just a euphemism for its being a *belligerent* process. Argument alone cannot prove anything, which only material facts can do (and when they do, argument is not required). As the proverb states: "Talk is cheap," since it does not prove anything. What is more, arguments do not generally tend to solve problems, but exacerbate them. The light bulb of popular humor will not be screwed no matter how many lawyers argue about it, although the client will be.

Arguments culminate in fights. As an argument progresses, the parties hate each other more and become increasingly committed to their claims and their corollaries, which increases errors and lies (since the parties cannot often foresee the corollaries or go back on their word without hurting themselves). The argument eventually reaches a 'point of no return' where it becomes a fight and the truth no longer matters, whence the proverb: "In war, truth is the first casualty." At this point, the conflict no longer has anything to do with the facts, but becomes a matter of sheer vindictiveness where getting at the other party is all that matters. The 'facts' or the 'system' then become mere instruments for victory or vengeance on both sides.

In principle, the purpose of the justice system should be to put an end to arguments before they develop into fights, or to put an end to fights, by checking the facts. At any rate, the truth should not become a casualty; otherwise, justice dies with it. This is especially problematic in criminal law, where much time is spent preparing and delivering arguments, even though argument not only does not help, but diverts time and resources that should be spent on checking facts or the law. This results in enormous backlogs and illegal pre-trial detentions to prevent people awaiting trial from escaping or re-offending in the meantime. The author, for instance, spent over 300 hours preparing for a 3 hour argument over 8 photographs and a couple of police statements. While the photographs may assist in the determination of the facts of the case, since they can be hard to fabricate, argument is futile, since it does not ultimately prove anything and can go both ways, as astrophysicist Neil deGrasse Tyson remarked in an interview with Bill Maher:

Tyson: When you look at law, well, law is what happens in the courtroom. It doesn't go to what's right, it goes to who argues best. And there's this urge—the entire profession is founded on who the best arguers are.

Maher: Right, a courtroom is not about the truth. The theory, if I get what you're saying, is that each side argues their version and then the truth somehow emerges.

Tyson: That's the premise. However, the practice—which is bred in debating teams, for example—where you know the subject, but you don't know what side you're going to be

put on to argue. So, the act of arguing and not agreeing seems to be fundamental to that profession, and Congress is half that profession.⁶⁶

The JOJ system does not foster academic argumentation—such as scientific or philosophical argumentation—either. Rather, a 'good argument' by the standards of the JOJ system is whatever it takes to convince the JOJ. The JOJ, for its part, tends to be quite confident in its own 'good sense' of which the French philosopher René Descartes quipped:

Good sense is, of all things among men, the most equally distributed; for every one thinks himself so abundantly provided with it, that those even who are the most difficult to satisfy in everything else, do not usually desire a larger measure of this quality than they already possess.⁶⁷

G. McLachlin's Defense of the JOJ System

A critique of the JOJ system would not be as compelling without at least briefly addressing some of the arguments of one of its most respectable advocates: philosopher and chief justice Beverley McLachlin. In her 2004 lecture entitled *Judging in a Democratic State*,⁶⁸ McLachlin asserts that “By and large, our court system succeeds in getting at the truth.” She explains that

Each party presents evidence through witnesses and documents. The evidence is tested through the engine of cross-examination. Expert witnesses testify on matters that are beyond ordinary knowledge. The judge or jury considers and weighs the evidence and determines where the truth lies. In a civil case the judge must be satisfied of the truth on the balance of probabilities. In a criminal case she must be satisfied of the accused person's guilt beyond a reasonable doubt before she can convict. Lawyers and judges who spend their lives working in the system are generally satisfied that in the vast majority of cases the judge or jury reaches the right conclusion on the facts.⁶⁹

Thus, according to McLachlin, the JOJ system is generally effective to determine the correct facts; but how does she arrive at this judgment without any data to compare? Her conclusion that the JOJ system is accurate is based solely on the fact that jurists do not generally oppose it (and she cites no data concerning this). This is like saying the monarchy is good because there are no riots or because the royal court (which basically runs the system) believes it is good. The opinion of lawyers and judges cannot be counted as data for at least two reasons: first, because judges cannot know what really happened in any given case, except for what they may infer from testimony they hear in court; second, because lawyers should be of the opinion that the system does not work at least 50 percent of the time since 50 percent of cases are lost; for in every case there is a winner and a loser and in every case both parties are convinced they should win.

McLachlin carefully cautions, however, that

66 See Neil deGrasse Tyson & Bill Maher, “Neil deGrasse Tyson and Bill Maher” (August 2011), online: YouTube <<http://www.youtube.com/watch?v=v-gKXlaEEeY>>

67 René Descartes, *Discourse on Method*, translated by Henry Lee (NY: Collier Books, 1962 [1637]).

68 Beverley McLachlin, “Judging in a Democratic State” (Remarks of the Right Honourable Beverley McLachlin, CJ, at the Sixth Templeton Lecture on Democracy at the University of Manitoba, 3 June 2004), online: <<http://www.scc-csc.ca/court-cour/judges-juges/spe-dis/bm-2004-06-03-eng.aspx>>.

69 *Ibid.*

[O]ur system for finding the truth through the trial process is a good system, but for all that a human system. It may fail under the stress of fabrication, manipulation, bad faith, or simple error. Judges know this. They are constantly on the look out for it. They use the rules of evidence to achieve maximum reliability, and they examine their own minds and warn jurors against attitudes and preconceptions that may lead them to undervalue, overvalue or misinterpret conduct in testimony. Appeal courts, while they accept facts validly found, examine the record to ensure that the rules of evidence were properly applied in that the evidence taken as a whole supports the factual conclusion or verdict. Despite all these measures, judges and juries occasionally arrive at the wrong conclusion. When this happens we must do our best to find out why and to ensure that it does not happen again. But we should not condemn the entire system as unjust.⁷⁰

McLachlin therefore recognizes that judicial errors occur. However, her argument implies that they do not occur so often or so badly as to justify condemning the entire process. In response it is expedient to distinguish between actual error, which is an error of fact, and judicial error, which is a term of art. Indeed, the concept of judicial error is circular reasoning, since the judiciary defines judicial error by its own criteria and standards of review. In other words, 'judicial error' consists of judges correcting judges' poor judgment based on their own supposedly good judgment. Error is not evaluated by any external factors to the judiciary, except in patent miscarriages of justice, which are few and far between. Otherwise, there is no way to know how many actual errors the judiciary commits, and therefore no reason to put any confidence in the system, since there is no way to know it is accurate, except that judges judge their own judgment to be so. To use a simple analogy, a scientist cannot put any stock in a theory or instrument whose errors are unknown, let alone high or low (if it were low—even very low—the scientist could at least put *some* stock in it).

McLachlin explains that judges are often faced with conflicting evidence and must therefore resort

t o time-honoured techniques for deciding [. . .] where the truth lies: the internal consistency of the witness's evidence; demeanor; how the witness responded to questions in cross-examination. Empathy and an understanding of human nature are essential to this task.⁷¹ [Underline added.]

What are these “time honored techniques”? Oral argument? Direct and cross-examination? A closer examination of these “time-honored techniques” reveals them to be a mixture of pseudoscience (notably, physiognomy) and vulgar intuition. No million dollar Nobel prize was ever won over a person's testimony concerning the continuity of space and time, although million dollar awards are often decided this way in court. A person breaking into tears will not earn them a Nobel Prize, but it could put someone in jail for years and ruin their social relationships and reputation over an allegation of rape. In short, the JOJ system regularly metes out such expensive consequences on such cheap 'evidence.'

One of the main contentions of this article is that these so-called “time-honored techniques” should not be valued as evidence or in terms of consequences more than techniques that deserve a Nobel Prize. The probative and economic value of scientific and legal evidence should be proportional. (That is to say, evidence should not be more plausible in court, or have more stakes attached to it, than it would in a scientific context.) To illustrate this point, imagine if scientists' method to determine the color of the sky consisted of arguing in front of a JOJ. Why, then, is this method acceptable to determine who rots in jail?

⁷⁰ *Ibid.*

⁷¹ *Ibid.*

It is interesting to note that the aforementioned “time honored techniques” are often associated with a pathological contempt for expert witnesses, whose expertise tends to usurp the authority of the JOJ to judge the matter. Muldoon J, for instance, wrote in *Unilever PLC v Procter & Gamble Inc.*⁷²

Expert witnesses - called because, one supposes, of their eminence in the chemical science in which they proudly purport to be expert - are a large hindrance rather than much help because, of course, they are paid to contradict the eminent scientists on the opposite side ... A judge unschooled in the arcane subject is at difficulty to know which of the disparate, solemnly mouthed and hotly contended ‘scientific verities’ is, or are, plausible. Is the eminent scientific expert with the shifty eyes and poor demeanour the one whose ‘scientific verities’ are not credible? Cross examination is said to be the great engine for getting at the truth, but when the unschooled judge cannot perceive the truth, if he or she ever hears it, among all the chemical or other scientific baffle gab, is it not a solemn exercise in silliness?

This quotation betrays the frustration of a “judge unschooled” in chemistry who blames the expert witnesses and the fact they are paid to testify on either side for his confusion, when he might reasonably blame his own lack of education in science and the adversarial system for allowing experts to be bought and creating the incentive to do so. Muldoon J says he has no choice but to fall back on credibility assessment (a “time honored technique”) like shifty eyes and poor demeanor to assess the scientific evidence, which is an entirely inappropriate method reminiscent of comedian Ali G's improv with creationist Kent Hovind and futurist Graham Molitor.

Ali G: Where did men come from?

Molitor: Men evolved out of lower species in the way that all the species evolved out of lower species.

Hovind: This is ridiculous. This is ludicrous. If he wants to believe this happened he's welcome to believe his grandpa came down from a tree or something, but that is pure religion. We offer a quarter million dollars for anybody with any real empirical testable scientific evidence for evolution.

Ali G: So, you' saying we ain't come down from monkeys?

Hovind: No, monkeys are still having babies so why don't they have another human today?

Ali G: Okay, has you eva' ea'en a banana?

Hovind: Oh yeah. I eat all kinds of food, yeah.

Ali G: Yo. That's proof!

Hovind: That's not proof of evolution.

72 (1995), 61 CPR (3d) 499 at 488-89, quoted in Binnie, *supra* note 46 at 4.

Molitor: The people engaged in this kind of nonsense always set up the rules so you can never win the money.

Ali G: But the banana thing got him, 'innit? He started going red when I said that. He went red 'dinnit?

Hovind: Because I'm surprised at how illogical your question was.⁷³

However, there is a legitimate aspect to Muldoon J's frustration, which 19th c. British explorer and litterateur Richard Burton elegantly expresses in the following quatrain:

When doctors differ who decides
 a m i d t h e m i l l i a r d - h e a d e d t h r o n g ?
 Who save the madman dares to cry:
 “'Tis I am right, you all are wrong”?⁷⁴

Binnie J sees Muldoon J as raising the issue how do “judges *really* assess expert testimony”? Do they do this “based on what is said or how it is said or who says it”?⁷⁵ There is a real danger that experts, who best know their subject-matter, are in a position to misrepresent the 'truth of the matter' (*i.e.*, what they know or do not know) to others, who are usually in a position to take their word as authority. How, then, can a non-expert tell that an expert is misrepresenting the facts?

It is not that we were engaged in formulating lies; there was nothing as crude and naïve as that. But we were using facts, emphasizing facts, bearing down on facts, sliding off facts, quietly ignoring facts and, above all, interpreting facts in a way to do what Marshall said we had to do - ‘to get by those boys down there.’⁷⁶

This is a real problem, but the solution is not to resort to “time-honored techniques” like looking out for shifty eyes and poor demeanor to 'smoke out' the foxy experts, as such techniques are more likely to occlude fact-finding by resting the case on the testimony and demeanor of witnesses rather than provable facts. The question is: what is the alternative? This article will propose corroborated polygraph testing (CPT) as a viable alternative. The following section will begin by discussing the scientific status of polygraph testing.

II. The Scientific Status of Polygraph Testing

This section examines the strengths and weaknesses of polygraph testing from a scientific point of view.

73 See “Season 2, Episode 5: Science” of *Da Ali G Show*, TV series, directed by Dan Mazer, created by Sacha Baron Cohen (USA: HBO, Channel 4, 2003), online: <http://www.imdb.com/title/tt0508527/fullcredits?ref_=tt_ov_st_sm>. See “Ali G – Science” (28 August 2006), online: YouTube <<https://www.youtube.com/watch?v=eB5VXJXxnNU>>.

74 Richard Burton, *The Kasidah of Hâjî Abdû El-Yezdî* (Google EPUB, 1870), online: Internet Sacred Text Archive <<http://sacred-texts.com/isl/kas/index.htm>>.

75 Binnie, *supra* note 46 at 4.

76 See Alfred H Kelly's speech in "Hearings before a Subcommittee of the Committee on the Judiciary, United States Senate" (May 1973) 87th Congress, 2nd Session, where he describes how he helped to present the *Brown* case to the American Supreme Court.

A. Strengths

Polygraph testing, also known as the “lie detector,” is used for forensic purposes in Canada and the US; it is also used by the US federal government for testing job applicants and current employees. In 2003, the US Department of Energy assembled a Committee to Review the Scientific Evidence on the Polygraph (henceforth “the Committee”) comprised of members from various scientific backgrounds, none of whom was previously involved in polygraph research *per se*. The Committee's mandate was to conduct a scientific review of research concerning the reliability and validity—that is, the accuracy and precision—of polygraph testing. Over nineteen months, the Committee held a series of meetings; visited polygraph facilities at government agencies; examined reports and publications; listened to representatives on both sides of the debate on polygraph accuracy; and reviewed the evidence. The National Research Council (NRC) published the Committee's report in its seminal work *The Polygraph and Lie Detection*,⁷⁷ which shall be referred to as the NRC report.

The NRC report concludes that DNA profiling is the “model of cooperation between science and the law,”⁷⁸ because it produces “almost certain conclusions with a probability of error of one in billions or less (except for identical twins).”⁷⁹ By contrast,

The accuracy of polygraph testing does not come anywhere near what DNA analysis can achieve. Nevertheless, polygraph researchers have produced considerable data concerning polygraph validity [. . .]. However, most of this research is laboratory research, so that the generalizability of the research to field settings remains uncertain. The field studies that have been carried out also have serious limitations [. . .]. Moreover, there is virtually no standardization of protocols; the polygraph tests conducted in the field depend greatly on the presumed skill of individual examiners. Thus, even if laboratory-based estimates of criterion validity are accurate, the implications for any particular field polygraph test are uncertain. Without the further development of standardized polygraph testing techniques, the gulf between laboratory validity studies and inferences about field polygraph tests will remain wide.⁸⁰

The report also observes that

Much of the expert opinion that has been presented as “scientific” in courts is not based on what scientists recognize as solid scientific evidence, or even, in some cases, rudimentary scientific methods and principles. The polygraph is not unusual in this regard. In fact, topics such as bite mark and hair identification, fingerprinting, arson investigation, and tool mark analysis have a less extensive record of research on accuracy than does polygraph testing. Historically, the courts relied on experts in sundry fields in which the basis for the expert opinion is primarily assertion rather than scientific testing and in which the value of the expertise is measured by effectiveness in court rather than scientific demonstration of accuracy or validity.⁸¹

Thus, the Committee suggests that if polygraph testing compares well with other forensic

77 National Research Council, *The Polygraph and Lie Detection* (USA: National Academies Press, 2003) at 203 [NRC report].

78 *Ibid.*

79 *Ibid.*

80 *Ibid* at 204.

81 *Ibid* at 202-203.

sciences, then the courts may consider it because of its relative success.⁸² The Committee also concedes that

The available evidence indicates that in the context of specific-incident investigations and with inexperienced examinees untrained in countermeasures, polygraph tests as currently used have value in distinguishing truthful from deceptive individuals. However, they are far from perfect in that context, and important unanswered questions remain about polygraph accuracy in other important contexts. No alternative techniques are available that perform better.⁸³

Thus, the report mainly criticizes the validity of polygraph testing, its sensitivity to the examiner and examinee, and the comparability between the results of specific-incident tests in labs and field tests. However, it should be noted that polygraph testing, and research on these topics, has considerably developed since the NRC report was published in 2003. Indeed, a wealth of research has since emerged of which a few examples may be discussed. Professors Charles Honts and Racheal Reavy, for instance, conducted a mock crime experiment with 250 paid participants,⁸⁴ which compares the validity of probable- and directed-lie variants of the comparison question test (CQT), “the most widely applied technique for physiological detection of deception.”⁸⁵ Subjects were assigned at random to one of eight conditions in a guilt (guilty/innocent) × test type (probable-lie/directed-lie) × stimulation (repeat stimulation/no stimulation) factorial design. The data were scored by an experienced polygraph examiner who was unaware of subject assignment to conditions and with a computer algorithm known as the Objective Scoring System Version 2 (OSS2).

The results showed that guilt had a substantial effect in the OSS2 and human scores. Probable- and directed-lie variants of the CQT produced equivocal results in terms of detection accuracy. The directed- and probable-lie variants of the CQT had equivalent validity. Vasomotor responses were found to be very useful to detect deception, and the directed-lie variant of the CQT was found to be more advantageous than the probable-lie variant, because it is more standardized. Professors Heinz and Susanne Offe's research shows that 90 percent detection rates can be achieved using this method and that the significance of comparison questions was hardly affected by the different testing conditions (*viz.*, explanation of comparison questions and lack thereof).⁸⁶ These studies and others show that polygraph testing has gone a long way since—and perhaps largely thanks—to the critical NRC report, which professor David Raskin calls a “misguided and disingenuous attack on the polygraph.”⁸⁷

Raskin specifically analyzes the diagnostic reliability and validity of polygraph tests and compares it to other commonly used psychological and medical diagnostics. While psychophysicologists generally accept that there is no “specific lie response” or pattern of reactions that is peculiar to deception,⁸⁸ a variety of factors may cause subjects to react more

82 *Ibid.*

83 *Ibid.* at 178.

84 Charles Honts & Racheal Reavy, “The Comparison Question Polygraph Test: A Contrast of Methods and Scoring” (2015) 143 *Physiology & Behavior* 15.

85 David Raskin & John Kircher, eds, *Credibility Assessment: Scientific Research and Applications*, 1st ed (USA: Academic Press, 2014), c 3.

86 Heinz Offe & Susanne Offe, “The Comparison Quest Test: Does It Work and If So How?” (2007) 31 *Law Hum Behav* 291.

87 Raskin & Kircher, *supra* note 85.

88 J A Podlesny & David Raskin, “Physiological Measures and the Detection of Deception” (1977) 84 *Psychological Bulletin* 782. See also David Raskin, “Orienting and Defensive Reflexes in the Detection of Deception” in H D Kimmel et al, eds, *The Orienting Reflex in Humans* (NJ: Erlbaum, Hillsdale, 1979), 587–605.

strongly to questions about crimes of which they are suspected than to neutral questions. Strong reactions, such as nervousness, distrust of the examiner, anger, or disgust at the accusation, may cause autonomic reactions to relevant questions even when answered truthfully. Without a properly constructed test protocol, neither polygraph examiners nor psychophysicologists can accurately distinguish these reactions from those associated with deception.

Comparison questions (CQs) turn the subject's attention from relevant questions (RQs), thereby causing the subject to react more strongly to CQs or RQs. If the subject reacts more strongly to CQs, his or her reaction is interpreted as truthful. Conversely, if he or she reacts more strongly to RQs, his or her reaction is interpreted as deceptive. Inferences to truth or deception are drawn from the relative strength of the subject's reactions to CQs and RQs, which circumvents the lack of a specific lie reaction.

Whereas scientists debate the validity of CQ polygraph tests,⁸⁹ most psychophysicologists support their usefulness for credibility assessment.⁹⁰ In response to concerns expressed by the American Psychological Association concerning the scientific basis and specific applications of polygraph testing,⁹¹ Raskin examines the scientific literature on the subject. A couple of his findings bear mentioning. In 1983, the Office of Technology Assessment of the US Congress selected ten field studies that showed that the overall accuracy of the polygraph decisions was 90 percent on criterion-guilty suspects and 80 percent on criterion-innocent suspects.⁹² The Committee of Concerned Social Scientists similarly surveyed available field studies⁹³ whose results independently showed that the CQT had an average accuracy of 90.5 percent.

In 2012, Avital Ginton⁹⁴ conducted a field study using using 64 paired polygraph tests he obtained from the Israel Police where two subjects provided contradictory statements in each pair. Background information showed that only one of the two subjects could be telling the truth on the relevant questions, but which of them was unknown. Since the proportion of pairs with the same results is inversely related to the accuracy of the test, it was algebraically determined that the CQT was 94 percent accurate on guilty suspects and 84 percent accurate on innocent suspects, which reinforces the 90 percent overall accuracy findings of the aforementioned field studies.⁹⁵ Ginton's novel algebraic approach using paired testing eliminates the need for external verification, such as a confession or other evidence if it is known that one the subjects has to be lying.

Further to this evidence, in 2011 the American Polygraph Association Ad Hoc Committee on Polygraph Techniques⁹⁶ (the Ad Hoc Committee) conducted a four year study of the scientific

89 See C R Honts *et al*, "The Scientific Status of Research on Polygraph Techniques: The Case for Polygraph Tests" in D L Faigman *et al*, eds, *Modern Scientific Evidence: The Law and Science of Expert Testimony* (St Paul, MN: West Publishing, 2005). See also W G Iacono & D T Lykken, "The Scientific Status of Research on Polygraph Techniques: The Case Against Polygraph Tests" in D L Faigman *et al*, eds, *Modern Scientific Evidence: The Law and Science of Expert Testimony* (St Paul, MN: West Publishing, 2005).

90 S L Amato & C R Honts, "What do Psychophysicologists Think About Polygraph Tests? A Survey of the Membership of SPR," (1994) 31 *Psychophysiology* at S22. See also Gallup Organization, "Survey of the Members of the Society for Psychophysiological Research Concerning their opinions of polygraph test interpretations" (1984) 12 *Polygraph* 153.

91 David Raskin, "The Polygraph in 1986: Scientific, Professional, and Legal Issues Surrounding Applications and Acceptance of Polygraph Evidence" (1986) *Utah Law Review* 29 at 73.

92 Office of Technology Assessment, *Scientific Validity of Polygraph Testing: A Research Review and Evaluation* (Washington, DC: US Government Printing Office, 1983).

93 C R Honts & C F Peterson, Brief of the Committee of Concerned Social Scientists as Amicus Curiae to *United States v Scheffer*, 523 US 303 (1998).

94 Avital Ginton, "A Non-Standard Method for Estimating Accuracy of Lie Detection Techniques Demonstrated on a Self-Validating Set of Field Polygraph Examinations" (2013) 19:7 *Psychology, Crime & Law* 577.

95 Raskin & Kircher, *supra* note 85.

96 Ad Hoc Committee on Polygraph Techniques, "Meta-Analytic Survey of Criterion Accuracy of Validated Polygraph Techniques" (2011) 40 *Polygraph* 193.

basis for polygraph techniques in response to concerns various experts expressed about the methods and motives of the NRC report. The Ad Hoc Committee summarized results from 52 different experiments and surveys published in 37 different studies, including results from 289 scorers who provided a total of 12,665 scores for 2300 criterion deceptive examinations and 1983 criterion truthful exams. The results for CQT specific-incident diagnostic techniques produced an average criterion accuracy of 92 percent, which supports the aforementioned field studies' findings of an overall accuracy of 90 percent.

Finally, Philip Crewson⁹⁷ conducted a meta-analysis of data from 198 studies that compares four characteristics of polygraph tests with standard tests for medical and psychological diagnoses (Crewson discusses these standard tests in his article). Table 1 (below) presents the results of his study.

Table 1. Comparison of four characteristics of polygraph tests with standard tests for medical and psychological diagnoses.

Diagnosis	Polygraph	Medical	Psychological
Agreement ⁹⁸	77%	56%	79%
Sensitivity	92%	83%	72%
Specificity	83%	88%	67%
Accuracy	88%	86%	70%

These results show CQT is at least as reliable as other types of accepted expert testimony, notably medical, psychiatric, and psychological opinions. This is consistent with the NRC report's conclusion that polygraph evidence meets a comparable standard to other forensic sciences and that courts may consider it because of its relative success.⁹⁹ Indeed, the NRC report concedes that “specific-incident polygraph tests can discriminate lying from truth telling at rates well above chance, though well below perfection.”¹⁰⁰ It also concedes that “No alternative techniques are available that perform better.”¹⁰¹ A 90 percent overall accuracy rating may not be perfect, but it is certainly useful and the best alternative, especially compared to the unknown accuracy of JOJ verdicts. This is because the correct verdict cannot be known in a lot of cases. Moreover, the author is not aware of any research that directly tests the accuracy of JOJ verdicts. However, some research may seem to point to an answer.

Bruce Spencer,¹⁰² for instance, argues that the average accuracy of jury verdicts can be inferred by soliciting judges' verdicts, and then comparing them with juries' verdicts to see how often they agree. His study presumes that how often the judge and jury agree on the verdict of a case they both heard reflects how often the verdict is accurate. However, a flaw with this methodology is that the judge and jury may agree on a false verdict. Or, they may disagree under slightly different circumstances, such as the presence of a doubting Thomas or a hanging judge. It is also worth considering that the verdict of a jury is often strongly influenced by the judge's instructions as to what constitutes a legal inference and proof. Spencer, at any rate, finds that

97 Philip Crewson, “A Comparative Analysis of Polygraph With Other Screening and Diagnostic Tools” (2001) 32 Polygraph 57.

98 This is based on kappa coefficients.

99 NRC report, *supra* note 77 at 202-203.

100 *Ibid* at 4.

101 *Ibid* at 178.

102 Bruce D Spencer, “Estimating the Accuracy of Jury Verdicts” (2007) Journal of Empirical Legal Studies, DOI: 10.1111/j.1740-1461.2007.00090.x.

judges and juries agree in 87 percent of the cases of the National Center of State Courts.

A more appropriate way to determine the accuracy of JOJ verdicts (and hence the JOJ system) would be to test judges and juries in cases where they are blind to the facts, which are known to the researcher. Research on polygraph testing calls this testing by a 'mock crime.' The polygraph test has been so thoroughly tested with mock crimes, and polygraph examiners are so confident in the results and procedures, that Sgt. Labine and Det. Giampaolo performed such a test in class in front the students, including the author. They had one of three volunteer students commit a crime, namely the theft of a wallet, and then proceeded to polygraph test the three volunteers in front of the class. The test was a success, and the guilty student confessed to stealing the wallet. Although the result of the polygraph test was accurate, the test does not prove that the student stole the wallet, which is why this essay proposes that it should be corroborated with independent material evidence in order to constitute a scientific test. Conversely, the author has not found any studies that test the accuracy of JOJ verdicts with mock crimes.

However, the author has found such studies as Thomas Reidy *et al's* in 2013, which suggests that jurors are terrible at predicting the violence of offenders at the sentencing phase of aggravated murder trials. Reidy *et al* review the disciplinary records of 115 male inmates (65 sentenced to life and 50 to death) in Oregon from 1985 to 2008. The data shows that juries erred in their predictions 90 percent of the time. Of the 10 percent of them they got right, 90 percent of the violations caused no harm or only minor injuries.¹⁰³ Studies like this suggest that juries are poor judges of future events (which can be verified), and hence probably at least as poor judges of past events (which cannot be verified), since both judgments involve a symmetrical inference. (For instance, the inference that an accused probably committed a crime under certain conditions is the same that he would probably do it in the future under the same conditions.)

B. Weaknesses

The NRC report presents the main and commonest criticisms of polygraph evidence, judging by the response of its advocates, such as David Raskin, researchers he cites, and the American Polygraph Association. The following subsections will examine two of the main criticisms the NRC report raises, as well as some other criticisms, respectively. The NRC report's first main criticism of polygraph testing is the absence of a general signature response to indicate deception (meaning the polygraph test cannot be entirely valid). Its second criticism is that studies of the accuracy of polygraph testing are unrealistic inasmuch as they extrapolate from laboratory tests to field tests, which are not alike.

1. The Lack of a General Signature Response to Indicate Deception

The NRC report notes that the lack of a general signature response to indicate deception is a permanent flaw with polygraph testing, as it states:

Almost a century of research in scientific psychology and physiology provides little basis for the expectation that a polygraph test could have extremely high accuracy. Although psychological states often associated with deception (e.g., fear of being judged deceptive) do tend to affect the physiological responses that the polygraph measures, these same states can arise in the absence of deception. Moreover, many other psychological and physiological factors (e.g., anxiety about being tested) also affect those responses. Such phenomena make polygraph testing intrinsically susceptible to producing erroneous results. This inherent ambiguity of the physiological measures used in the polygraph

¹⁰³ T J Reidy *et al*, "Probability of Criminal Acts of Violence: A Test of Jury Predictive Accuracy" (2013) 31:2 Behav Sci Law 286, DOI: 10.1002/bsl.2064.

suggests that further investments in improving polygraph technique and interpretation will bring only modest improvements in accuracy.¹⁰⁴

This is indeed a serious issue; however, proponents of polygraph testing acknowledge it and counter with the argument that stronger reactions to relevant questions correlate to truthfulness. Moreover, background information or the use of an algorithm to narrow down the possibilities, as in Ginton's study, can corroborate the analysis.

It may be further answered to this criticism that the cardinal virtue of polygraph testing is not its ability to detect truth or lies, but to evidence hidden information which can be used to predict evidence or narrow down contradictions that might not otherwise be apparent. For instance, instead of interpreting polygraph data to show the subject is truthful or lying, the data could be simply interpreted to indicate a strong or weak reaction to certain statements, which can be used to verify the statement or compare it to other evidence, such as the reaction of another subject in a paired test. The results of this analysis could in turn be compared with background information, such as other evidence or knowledge one of the paired subjects must be lying, as Ginton does in his study. The power of this approach to narrow down possibilities can be quite formidable, as Ginton's study demonstrates, and should not be underestimated simply because the polygraph does not measure truth or lies.

Polygraph data alone cannot tell that a person is truthful or lying for at least two reasons. First, the subject may be deluded. Second is the absence of a general signature response indicating deception. However, if independent material evidence (IME) corroborates the data, then a legitimate inference could be made as to the truth or falsehood of the claim. To put it crassly, a person's mouth may lie while their systolic blood pressure may not. Whether or not the subject is *actually* lying (as opposed to deluded, for example) is irrelevant. What matters is the truth, and the way to know the truth is to verify it. Polygraph testing can thus collect a number of clues that may not otherwise be collected. It can therefore be useful to predict evidence and to narrow down the possibilities. By contrast, the only data that can be obtained from testimony is the witness' word and presentation in court, which ends up being 'his word against hers' without an objective measure of the witnesses' reactions. (As Hand observed, these reactions are lost evidence and left to the whim of the JOJ to determine their significance.)

That being said, it should be borne that polygraph testing may not always predict evidence (or uncover hidden information) or narrow down the possibilities. When it does, however, it can be very strong evidence that is not easily explained away as a conspiracy or coincidence. Polygraph evidence also has the advantage that it can be objectively reviewed and subjectively challenged. (For instance, if polygraph fails to point to corroborating IME, the subject can potentially repeat the test, which will either confirm or confound the data.)

Counter to the NRC report's suggestion that polygraph testing is unscientific or a weak science (due to the absence of a general signature response to indicate deception), it may be argued that polygraph testing exhibits the scientific attribute of prediction. To wit, polygraph data can predict evidence by uncovering hidden information. Two quote Bill Nye ('the science guy'):

[W]hat we want in science, science as practiced on the outside, is an ability to predict. We want to have a natural law that is so obvious and clear, and so well understood, that we can make predictions about what will happen. We can predict that we can put a spacecraft in orbit, and take a picture of Washington DC. We can predict, that if we provide this much room for an elephant, it will live healthfully for a certain amount of time.¹⁰⁵

104 NRC report, *supra* note 77 at 2.

105 Ken Ham & Bill Nye, "Ken Ham vs. Bill Nye Debate," transcript by Bill Browning (4 February 2014), online: <<http://www.youngearth.org/index.php/archives/rmcf-articles/item/21-transcript-of-ken-ham-vs-bill-nye-debate>>.

According to some scientists the ability to predict is a sufficient attribute of science. Richard Dawkins, for instance, argues that

Quantum theory yields experimental predictions which have been verified to an accuracy—to a number of decimal places—so accurate, that the great theoretical physicist Richard Feynman compared it to the accuracy of predicting the width of North America to the accuracy of the width of one human hair. That is why quantum theory has to be taken seriously. And it doesn't matter—well it does matter, but one can take one stride because of the brilliance of the experimental verification—it doesn't matter that quantum theory is so mysterious that, as Feynman himself once said: “If you think you understand quantum theory, you don't understand quantum theory.”¹⁰⁶

However, other scientists might argue that prediction alone is not sufficient to qualify as science, and that a consistent theory and methodology are also necessary. In this case, the scientific status of polygraph testing may be contested on the basis that the lack of a general signature response to indicate deception is a fatal flaw in its underlying theory. However, this argument only holds if polygraph testing is supposed to tell if the subject is truthful or lying without other evidence. If, on the other hand, polygraph testing is supposed to uncover hidden information that can be contradicted or corroborated, then this argument will not apply.

Conversely, if prediction is enough to count as science, then polygraph testing evidently meets the test. Consider, for instance, the case of a bank heist with two suspects. Polygraph testing can uncover important hidden information that could help to solve the case. For instance, it could help to determine where is the stolen money based on their responses to relevant questions, especially if background information shows that one of them knows where is the money. Paired polygraph testing can also determine the likelihood that a suspect is lying if background information indicates one of them is lying. For these reasons, polygraph data should not be ignored or deprecated as unscientific or weak science despite its useful applications and obvious successes.

2. Lack of Realism and Generalizability from Lab Tests to Field Tests

A second major criticism of polygraph testing which is found in the NRC report is that it lacks realism inasmuch as it extrapolates from lab tests to field tests, which are not alike. More specifically, the NRC report notes that the studies it examined tend to generalize from specific-incident tests (typically mock crimes) to field tests.¹⁰⁷ However, a weakness with this criticism is that the NRC report only examined 57 studies, all of which examine specific incidents. However, there are now abundant field studies, some of which were discussed in the previous subsection.

3. Other Criticisms

Other criticisms of studies on the accuracy of polygraph testing include that of Iacono and Lykken, who argue that the false-negative rate of polygraph tests tends to be underestimated and that confessions are not independent from the test results.¹⁰⁸ Contrary to this view, Raskin argues that methods such as Ginton's, which consists of cross-checking paired tests based on

106 Richard Dawkins, “The God Delusion” (Lecture concerning The God Delusion delivered at Book TV, C-Span 2, 23 October 2006), online: <<https://www.youtube.com/watch?v=-KhHtYEpvho>>.

107 NRC report, *supra* note 77 at 4.

108 Raskin & Kircher, *supra* note 85.

background information that one of the two suspects is lying, overcomes this problem.¹⁰⁹

Professor Jeffrey Roskey adds another criticism: that any method of polygraph testing cannot account for many extraneous variables and parameters that may confound the results and analysis. For instance, his study that examines the relationship between recidivism and sexual history polygraph tests shows that the lack of statistically significant different rates of passing these tests between juveniles and adults can be related to numerous issues, such as lack of independence or adjustment for differing rates of opportunity across ages, poor construct validity of deceit, failure to adjust for base rates of deceit in subsequent analyses, failure to include recidivism as an outcome, and use of recidivism as an outcome to judge post-conviction polygraph.¹¹⁰ This is a difficult criticism to counter if one supposes that the polygraph is a truth-measuring device or a 'lie detector.' However, if the polygraph is regarded as instrument which points to evidence by indicating where there might be hidden information, then the fact it does not precisely measure probity is not really relevant. That it often does, however, can be very useful.

III. The Legal Status of Polygraph Evidence in Canadian Criminal Law

This section is divided into three subsections: the first examines the legal status of polygraph evidence in Canadian criminal law; the second considers why polygraph evidence is not admissible in Canadian law; and the third considers when and why it should be admissible.

This section specifically focuses on the legal status of polygraph evidence in criminal law for three reasons. First, criminal law is the same throughout the country, since it is federal law. Moreover, criminal law is the subject of many important Supreme Court decisions, so criminal laws and their justifications tend to be quite clear and rigorous. Second, criminal law is a good model, since the standard of proof in criminal law is beyond a reasonable doubt, unlike civil law, where the standard of proof is on a balance of probabilities. It is therefore more cautious to make inferences from criminal law principles to civil law ones, since criminal law principles are more rigorous. (For instance, evidence that is admissible in a civil trial may not be admissible in a criminal trial, whereas evidence that is admissible in a criminal law trial will most likely be admissible in a civil trial.) Third, focusing specifically on criminal law serves to narrow down and simplify the analysis, which can be extrapolated to a civil law context.

A. Polygraph Testing in Canadian Criminal Law

This subsection examines five influential Canadian criminal law judgments concerning polygraph testing from 1978 to the present in chronological order, *viz.*, *R v Phillion*¹¹¹ (1978), *R v Béland*¹¹² (1987), *R v Mohan*¹¹³ (1994), *R v Oickle*¹¹⁴ (2000), and *R v Rafferty*¹¹⁵ (2016). Except for the last judgment by the Ontario Court of Appeal (ONCA), all of these judgments are by the Supreme Court of Canada (SCC).

R v Béland,¹¹⁶ 1976 SCC

109 See Iacono & Lykken, *supra* note 85.

110 Jeffrey W Rosky, "More Polygraph Futility: A Comment on Jensen, Shafer, Roby, and Roby" (2015) 31:10 Journal of Interpersonal Violence 1956.

111 *Phillion*, *supra* note 10.

112 *Béland*, *supra* note 11.

113 *Mohan*, *supra* note 12.

114 [2000] 2 SCR 3 [*Oickle*].

115 2016 ONCA 816 [*Rafferty*].

116 *Béland*, *supra* note 10.

Béland and Phillips were charged with conspiracy to rob an armored truck. However, the robbery did not take place because their co-conspirator, Grenier, ratted them out to the police. Grenier's testimony implicated Béland and Phillips, who testified in their own defense, asserting there was no conspiracy. Béland applied to take a polygraph test, but the trial judge denied the application because the results would be inadmissible. The issue went up to the SCC, which had to determine the admissibility of the results.

La Forest J, writing for the majority, said there are two compelling reasons to exclude polygraph evidence in court. The first is human fallibility in weighing the evidence, which is cloaked under the “mystique of science.”¹¹⁷ Second is the “inadvisability of expending time on collateral issues.”¹¹⁸ La Forest J elaborates her reasons thus:

[A]dmission of polygraph evidence will serve no purpose which is not already served. Such admission will disrupt proceedings, will open the trial process to the time-consuming and confusing consideration of collateral issues and will deflect the focus of the proceedings from the fundamental issue of guilt or innocence. It will also lead to numerous complications which will result in no greater degree of certainty in the process than that which already exists. The results recorded by the polygraph instrument, their nature and significance will reach the trier of fact through the mouth of the operator. Human fallibility will thus still be present, but now fortified with the mystique of science. [Underline added.]

La Forest J's observation (underlined in the quotation above) that polygraph testing “will [. . .] lead to numerous complications which will result in no greater degree of certainty in the process than that which already exists” is particularly interesting. She cites no proof of this, even though there is ample evidence that the contrary will be the case, *viz.*, that polygraph testing will save time in court and improve confidence in the results. Her conclusion that polygraph evidence should be “has no place in the judicial process where it is employed as a tool to determine or to test the credibility of witnesses,”¹¹⁹ however, is based on the rules of evidence rather than “fear of the inaccuracies of the polygraph.”¹²⁰

Against what she describes as the “superficial appeal”¹²¹ of the argument that polygraph testing is better than the “imperfect methods of the past,”¹²² La Forest J argues that “this cannot prevail in the face of the realities of court procedures.”¹²³ Table 2 (below) summarizes the principal rules of evidence that polygraph evidence is supposed to violate according to La Forest J and how. A critique of these reasons follows.

Table 2. Rules of evidence polygraph evidence violates and how.

Rule of Evidence	How Polygraph Evidence Violates It
1. Oath helping	Polygraph evidence serves only to bolster or undermine a witness' credibility, which is a form of oath helping. ¹²⁴ As such, the polygraph examiner's testimony cannot increase or reduce the weight of the examinee's statements. ¹²⁵

117 *Ibid.*

118 *Ibid.*

119 *Ibid* at para 18.

120 *Ibid* at para 19.

121 *Ibid* at para 18.

122 *Ibid.*

123 *Ibid.*

124 *Ibid* at para 9.

2. Prior consistent statements	Repeating a witness' statements cannot increase or reduce their weight. ¹²⁶
3. Usurpation of the role of the trier of fact	Only the JOJ can judge the credibility of a witness based on their presence and testimony in court. ¹²⁷
4. Character evidence	Polygraph evidence draws an illegal inference to probity from physical characteristics.
5. Opinion evidence	Polygraph evidence is an unnecessary expert opinion concerning credibility assessment, which the JOJ can do without expert assistance. ¹²⁸

Rules 1 and 2 in the Table above—namely, that polygraph evidence violates the rules against oath helping and prior consistent statements—comes down to the claim that polygraph evidence is essentially self-serving evidence. However, this is not always the case. It was only the case in *Béland* where the accused sought to submit polygraph evidence to prove his innocence. Whether or not polygraph evidence is self-serving depends on its use and is not intrinsic to the evidence itself.

As for rule 3—that polygraph evidence impinges on the role of the trier of fact (the JOJ)—it is not clear why the judiciary maintains the principle that only the JOJ can rule on the credibility of witnesses other than the fact it is constitutionally mandated. It is not even clear that the JOJ is best or at all suited to this task. However, it is indeed their constitutionally mandated role, which is not for any judge to change. From a legal perspective, therefore, that the constitution makes the JOJ the sole judges of the credibility of witnesses is a sufficient reason to exclude polygraph evidence if it impinges on their role.

So far, the reasons mentioned for excluding polygraph evidence have only to do with law and nothing to do with science. La Forest J's point with regard to rule 4, however, comes pretty close to being scientific; namely, that polygraph evidence runs afoul of the rule against character evidence. The rule against character evidence is a crucial rule of evidence that is connected to the so-called Prohibited Inference—a fallacy with a bloody history in the judiciary—which basically consists of the inference that the accused 'did it' because he or she are 'just the type of person who would do it' or exhibits the sort of character traits the JOJ expects to find in the offender.

La Forest J's reasons for rule 5—that polygraph evidence constitutes inadmissible opinion evidence because it is within the experience of the average JOJ—are less compelling. It is not obvious why credibility assessment is not within the purview of expert testimony. The advantage of expert testimony, even if the evidence is weak, is that it produces a report that can be checked. The conclusions of experts, if they are properly screened according to the criteria established in *R v J-LJ*,¹²⁹ will be based on theories and methods that can be tested and whose errors are known. Unlike the unwritten gut feelings of any JOJ. While an expert might write in his report something to the effect that 'In 10 percent of cases examined from such and such a population under such and such conditions, an eye twitch can indicate mendacity' a JOJ might simply think: 'Jail the bugger. I don't like how his eye twitches,' and nothing could stop them for thinking that. Nor could anyone know.

At any rate, the *Béland* decision is particularly important, since prior to it, in 1976, *R v Wong*¹³⁰ allowed the admission of polygraph evidence, provided it is not proffered as an opinion

¹²⁵ *Ibid* at para 12.

¹²⁶ *Ibid*.

¹²⁷ *Ibid*.

¹²⁸ *Ibid* at para 17.

¹²⁹ [2000] 2 SCR 600 [*J-LJ*].

¹³⁰ 1976 BCJ No 1299 at para 33.

that the individual is truthful or lying. La Forest J, on the other hand, concludes in *Béland* that the polygraph has no place in court because it is self-serving evidence that impinges on the role of the JOJ. In short, only the JOJ can assess credibility based on the presence and testimony of witnesses in court. Notwithstanding these provisions, lawyers and police continue to use polygraph testing for investigative purposes, such as screening clients and suspects.

***R v Phillion*,¹³¹ 1978 SCC**

Mr. Phillion was accused of the first degree murder of Mr. Roy. The defense called John Reid—a polygraph examiner—to testify against Mr. Phillion, although the latter had answered all the relevant questions in the negative. Mr. Phillion was convicted on this evidence.

In 2009, by which time Mr. Phillion had spent thirty two years in prison, it was determined that Mr. Phillion had been wrongfully convicted. Mr. Phillion sued the Crown and police for \$14 million. This case reinforced the principle in *Béland* that polygraph tests are inadmissible, as it is the sole function of the JOJ to assess the credibility of witnesses based on their presence and testimony in court.

Sgt. Labine remarked concerning this case that a trial essentially concerns the credibility of the accused. However, by law only two groups of people are allowed to form an opinion thereon, namely the JOJ. Therefore, polygraph evidence is inadmissible because it impinges on this role of the JOJ, which is the moral of the case.¹³²

***R v Mohan*,¹³³ 1994 SCC**

Dr. Mohan, an Ontario pediatrician, was charged with sexual assault of four teenaged patients. After direct examination of Dr. Mohan, the defense intended to call Dr. Hill—a psychiatrist—to testify concerning the psychosexual profile of the offender. In a *voir dire*, Dr. Hill testified that the perpetrator of the first three offenses was likely a pedophile, while the perpetrator of the fourth offense was likely a sexual psychopath. In Dr. Hill's opinion, Dr. Mohan's profile was neither consistent with that of a pedophile nor of a sexual psychopath.

While the *Mohan* case does not deal specifically with polygraph evidence, it does deal with expert evidence more generally and with a number of issues that are relevant to polygraph evidence in particular, three of which may be noted:

1. Whether the offense has unusual features indicating the offender is a member of a limited and distinct group.
2. Whether Dr. Hill's experience of interviewing only three doctors accused of sexual misconduct is a sufficient basis to generalize.
3. Dr. Hill's inability to diagnose a profile such as pedophilia until a patient committed an act consistent with that profile, such as an act of pedophilia.

The first issue relates to the problem with the theory behind polygraph testing, *viz.*, that there is no general signature response to indicate deception, which is analogical to the “limited and distinguishable group” criterion *Mohan* introduces. (Should the lack of a “limited and distinguishable group” criterion also apply to the polygraph?) The second issue has to do with sample size and induction, which relates to the problem of generalizing from specific-instance tests to field tests on the accuracy of polygraph testing. (*I.e.*, what is an adequate sample to generalize from?) The third issue relates to prediction. What weight should be given to Dr. Hill's

¹³¹ *Phillion*, *supra* note 11.

¹³² O D Hamza, “CML 3193: Forensic Science,” Class Notes (2016) [unpublished].

¹³³ *Mohan*, *supra* note 12.

diagnosis of a pedophile *after* he commits an act of pedophilia (that is, his diagnosis *ex post facto*) if he cannot identify a pedophile *before* he commits an act of pedophilia (that is, his diagnosis *ex ante*)?

The trial judge ruled that Dr. Hill's evidence is inadmissible. Dr. Mohan was convicted at trial; however, the Ontario Court of Appeal overturned his conviction, concluding that the trial judge erred in excluding Dr. Hill's evidence, since the trial judge's ruling was based on an evaluation of its reliability, which is a matter of weight rather than admissibility. (Wilson J, writing for the majority of the SCC, noted in *R v Lavallee*¹³⁴ that the developing trend at the time was that reliability was a matter of weight rather than admissibility in cases involving psychiatric evidence where the issues were beyond the scope of normal human experience.) This decision was again appealed to the SCC, which had to determine whether Dr. Hill's expert testimony was admissible or ran afoul of the rule against character evidence.

Sopinka CJ, writing for a unanimous court, allowed the appeal and ruled that Dr. Hill's evidence should be excluded because it violates the rule against character evidence. Sopinka CJ was also not satisfied that there was a clear standard to determine the profile of a pedophile or psychopath and did not, therefore, consider the evidence to meet the standard of threshold reliability in order to be admissible in court. He also found that the evidence was not sufficiently relevant to help the trier of fact.

Three points of this judgment are worth emphasizing in connection to polygraph evidence. First, it is noteworthy that relevance is both a legal criterion and a logical one. The logical criterion of relevance in law relates to probabilities, namely whether the evidence makes a matter at issue more or less likely to be true as a matter of commonsense.¹³⁵ Sopinka CJ also emphasizes that admissible evidence should be sufficiently reliable (*i.e.*, should satisfy a criterion of threshold reliability) in order to be admissible. The closer the evidence comes to the ultimate issue (that is, the question of guilt or innocence), the more rigorously this principle applies.¹³⁶ (Note that *R v Burns*¹³⁷ had just abolished the ultimate issue rule which excludes evidence concerning it.) In short,

The expert's group profiles were not seen as sufficiently reliable to be considered helpful. In the absence of these indicia of reliability, it cannot be said that the evidence would be necessary in the sense of usefully clarifying a matter otherwise inaccessible, or that any value it may have had would not be outweighed by its potential for misleading or diverting the jury. Given these findings and applying the principles referred to above, I must conclude that the trial judge was right in deciding as a matter of law that the evidence was inadmissible.¹³⁸

In considering whether the probative value of the evidence is likely to outweigh its prejudicial effect (in terms of moral or reasoning prejudice to the JOJ) Sopinka CJ specifically considers whether the JOJ is likely to uncritically accept the evidence “dressed in scientific language” and delivered by “a witness of impressive antecedents [. . .] as being virtually infallible and as having more weight than it deserves.”¹³⁹

In conclusion, the *Mohan* criteria do seem to militate against the admission of polygraph testing, as they do against Dr. Hill's opinion of Dr. Mohan's psychological profile based on the former's experience of only three interviews with doctors who were also accused of sexual

134 [1990] 1 SCR 852.

135 See *R v Truscott* (2006), 213 CCC (3d) 183 (Ont CA) at para 23.

136 *Mohan*, *supra* note 12 at para 28.

137 [1994] SCJ No 30.

138 *Mohan*, *supra* note 12 at para 46.

139 *Ibid.*

misconduct. Moreover, polygraph evidence, like Dr. Hill's opinion, tends to be “dressed in scientific language” that jurors may not understand or properly evaluate. Polygraph evidence also typically concerns the ultimate issue (that is, whether the accused is lying or truthful). Plus, there is no clear general profile of a liar or truthful person or of characteristics such a person might exhibit to compare with the accused. Finally, polygraph evidence impinges on the JOJ's exclusive prerogative to judge the credibility of witnesses. These are all very good legal reasons to exclude polygraph evidence, although they are poor scientific reasons, except for the lack of a “limited and distinguishable group”¹⁴⁰ to compare with the accused.

***R v Oickle*,¹⁴¹ 2000 SCC**

Oickle was a volunteer firefighter. Over the span of ten months, eight fires were set in his community, and he responded to each of them. To help narrow the list of suspects, police asked seven or eight people including Oickle to take a polygraph test. Five of the people passed the test and were removed from the list of suspects. Oickle was initially reluctant to participate. He was informed of his right to counsel, told he could leave anytime, and informed that polygraph evidence is inadmissible, although things he says could be admissible. After failing the polygraph test, Oickle confessed to setting all eight fires.

The issue at trial concerned the admissibility of Oickle's confession, which depends on whether it was vitiated by the circumstances. The Court of Appeal ruled that the confession is admissible, since it was voluntary. The SCC concluded that insofar as the operation of the polygraph test does not impact the voluntary character of a confession, errors such as misleading the subject concerning the length, accuracy, or inadmissibility of the polygraph test will not render the confession inadmissible.¹⁴² In short, if the statement is voluntary, given regard to the circumstances,¹⁴³ then it is admissible. The SCC mentions the following examples of what could vitiate the voluntary character of a confession:

1. suggesting it would be better for the respondent to confess;
2. mistreating or oppressing the respondent;
3. questioning the respondent in a harsh tone;
4. offering strong enough inducements to raise a reasonable doubt as to whether the confession was voluntary.¹⁴⁴

The *Oickle*¹⁴⁵ case shows that while polygraph evidence is inadmissible, it can still be used outside of court to generate statements that can be admissible in court. Put simply, polygraph testing cannot be used *as* evidence, but it can be used to *generate* evidence. For instance, McLintic, Rafferty's co-accused in the Victoria Stafford murder, gave a polygraph statement that was admitted into evidence pursuant to a *KGB*¹⁴⁶ application. The Ontario Court of Appeal did not fault this judgment.¹⁴⁷ What is peculiar about this is that the law treats the polygraph as a sort of dysfunctional instrument that can be used to get people to say things that may be received as evidence against them, although the data it produces and its analysis is not treated as evidence. In fact, it should be the opposite: whatever people say (unless it can be proven) should not be treated as evidence, whereas polygraph data and analysis may constitute or

¹⁴⁰ *Ibid.*

¹⁴¹ *Oickle*, *supra* note 114.

¹⁴² *Ibid* at para 88.

¹⁴³ *Ibid.*

¹⁴⁴ *Ibid* at para 104.

¹⁴⁵ *Ibid.*

¹⁴⁶ See *R v B(KG)*, [1993] 1 SCR 740 [*KGB*].

¹⁴⁷ See *Rafferty*, *supra* note 115 at paras 14-24.

point to evidence. *Oickle* thus exemplifies how the common law persists in a medieval paradigm where things people say matter, but measurement and analysis do not.

***R v Rafferty*,¹⁴⁸ 2016 ONCA**

In 2012, Mr. Rafferty was convicted of murder, violent sexual assault, and kidnapping of eight year-old Victoria Stafford. At trial, however, Ms. McLintic—Rafferty's co-accused—pleaded guilty to the murder of Stafford contrary to her agreed statement of facts. Rafferty appealed his conviction on this basis, but the Ontario Court of Appeal dismissed the appeal, citing both the absence of incriminating evidence against McLintic and an air of reality to the defense that he helped her hide the evidence despite not committing the crime.

McLintic gave a statement to a police polygraph examiner where she admitted to participating in the kidnapping, sexual assault, and murder, but said that Rafferty killed the child.¹⁴⁹ Then, she testified that she was the killer.¹⁵⁰ Part of her statement to police was admitted as a principled exception to the hearsay rule.¹⁵¹ Rafferty submits that the trial judge erred by admitting this statement, which meets the necessity requirement of the Principled Approach, but not the threshold reliability requirement.¹⁵² The trial judge admitted the statement based on the criteria in *Khelawon*¹⁵³ and *KGB*,¹⁵⁴ since it “was made under caution, rather than under oath, it was made following consultation with her counsel, it was videotaped and, importantly, McClintic was available for full cross-examination.”¹⁵⁵

Hence, the *Rafferty*¹⁵⁶ case shows that polygraph statements can be admissible if reliable and the person who made them is available to cross-examine. Like *Oickle*,¹⁵⁷ *Rafferty*¹⁵⁸ exemplifies how spurious statements made in a polygraph test are admissible, whereas the data and analysis are not.

B. Why Polygraph Evidence is Not Admissible in Canadian Criminal Law

The reasons for excluding polygraph evidence that were mentioned in the previous section can be grouped into three major categories as follows:

1. **The rules of evidence** preclude polygraph evidence because it is the role of the JOJ to assess the credibility of witnesses. (*R v Phillion* makes this point.¹⁵⁹) Polygraph evidence also violates rules of evidence such as the rule against oath helping.
2. **The constitution** establishes the right to a bench or jury trial.
3. **The adversarial process** requires parties to argue the evidence in court, so the JOJ can decide whom to believe; however, the “mystique of science” of polygraph evidence is supposed to hinder this process by compelling the JOJ to default to the opinions of experts.

¹⁴⁸ *Ibid.*

¹⁴⁹ *Ibid* at para 14.

¹⁵⁰ *Ibid* at para 15.

¹⁵¹ *Ibid* at para 16.

¹⁵² *Ibid* at para 17.

¹⁵³ See *R v Khelawon*, 2006 SCC 57, [2006] 2 SCR 787.

¹⁵⁴ See *KGB*, *supra* note 146.

¹⁵⁵ *Rafferty*, *supra* note 115 at para 19.

¹⁵⁶ *Ibid.*

¹⁵⁷ *Oickle*, *supra* note 114.

¹⁵⁸ *Rafferty*, *supra* note 115.

¹⁵⁹ *Phillion*, *supra* note 11.

While these are apt legal reasons to exclude polygraph evidence, the legal principles behind them need not exclude this evidence, as the SCC supposes, since these principles are not ends in themselves but means to justice, at the heart of which is the search for truth.

When the SCC dismisses polygraph evidence in *Phillion*,¹⁶⁰ *Béland*,¹⁶¹ and *Mohan*,¹⁶² it provides no evidence that it is unreliable. Indeed, La Forest J clearly states that her reasons for excluding polygraph are legal, not scientific. To wit, polygraph evidence offends the constitutional role of the JOJ, on which it impinges, and rules of evidence, such as the rule against oath helping. Hence, the point of this article is not that the SCC's legal reasons for excluding polygraph evidence are entirely inadequate (in fact, a lot of them are very good legal reasons). Rather, this article makes the point that the SCC's legal reasons contradict their fundamental purpose, which is the search for truth, when they are used to exclude polygraph evidence in some cases. This article also makes the point that considerations of the scientific status and accuracy of polygraph evidence are not irrelevant to justice but central to the determination whether it should be admitted or excluded. For instance, given that polygraph testing is more reliable than credibility assessment by a JOJ, it stands to reason that it should replace this procedure. Put simply, there is more reason to find the unconstitutionality of the JOJ system than there is to find the unconstitutionality of polygraph evidence in light of evidence that polygraph testing is more reliable than the JOJ system.

The following quotations by the SCC summarize its most salient reasons for excluding polygraph evidence, which will be followed by critiques thereof. For instance, in *R v Béland* the SCC states:

[A] basic tenet of our legal system [is] that judges and juries are capable of assessing the credibility and reliability of evidence.¹⁶³

This implies that evidence that tends to establish or undermine the credibility of a witness, such as polygraph evidence, is inadmissible because the JOJ is supposed to be able to assess the credibility of witnesses without any help.

Professors David Paciocco and Lee Stuesser further note that the SCC decision *Bleta v R*¹⁶⁴ establishes that “[T]he facts of the case are for the trier of fact to determine, not the expert. Unless those facts are uncontested.”¹⁶⁵ In other words, the JOJ is sovereign to determine the facts of the case based on the evidence presented in court. This is bizarre, because the JOJ is not sovereign to determine what evidence is presented in court (judges make this decision). The system supposes, on one hand, that the jury can determine the truth. However, with some kinds of evidence, such as polygraph evidence (or any evidence that is supposed to exude the “mystique of science” or cause the jury ‘moral or reasoning prejudice’¹⁶⁶), it supposes that they “can’t handle the truth,”¹⁶⁷ to borrow the famous line of *A Few Good Men*.

The SCC further states in *R v WH*:¹⁶⁸

It is for the jury to decide, notwithstanding difficulties with a witness’s evidence, how much, if any, of the testimony it accepts. Credibility assessment does not depend solely on objective considerations such as inconsistencies or motive for concoction;

160 *Ibid.*

161 *Béland*, *supra* note 10.

162 *Mohan*, *supra* note 12.

163 *Béland*, *supra* note 10.

164 [1964] SCR 561.

165 David Paciocco & Lee Stuesser, *The Law of Evidence*, 5th ed (Toronto: Irwin, 2008) at 212.

166 *Mohan*, *supra* note 12.

167 *A Few Good Men*, film, directed by Rob Reiner (USA: Columbia Pictures, 1992).

168 2013 SCC 22, [2013] 2 SCR 180.

accordingly, the jury is entitled to decide how much weight it gives to such factors. The reviewing court must be deferential to the collective good judgment and common sense of the jury. While appellate review for unreasonableness of guilty verdicts is a powerful safeguard against wrongful convictions, it is also one that must be exercised with great deference to the fact-finding role of the jury. Trial by jury must not become trial by appellate court on the written record.¹⁶⁹

What is interesting about this statement is that it shows deference to juries' "collective good judgment and common sense"¹⁷⁰ despite the fact they do not give reasons for their verdict, which means the appellate court is not provided with any reasons for the jury's decision that it can review. The SCC also says that juries can make a subjective credibility assessment, using their own 'good sense,' although it does not allow them to make an objective assessment by using polygraph data and analysis.

A further contradiction is notable in that the JOJ is supposed to be sovereign to determine the facts of the case. It follows from this assumption that more evidence can only increase the JOJ's sovereignty in this regard. Yet, polygraph evidence is excluded because it is supposed to undermine or 'usurp' the JOJ's sovereignty. It is as if there is an implicit Principle of Subjectivity that stipulates that if evidence becomes too technical (*i.e.*, if the JOJ cannot arbitrarily determine the fact of the matter), then it must be excluded.

According to Sgt. Labine the basic reason why polygraph evidence is inadmissible is because it impinges on the constitutional role of the JOJ as the sole judges of the credibility of a witness.¹⁷¹ Fear of the "mystique of science"¹⁷² is unfounded, not only because there is evidence that polygraph testing works in some situations, but also because the court admits less reliable forensic evidence, as the NRC report points out. Professor Frederick Schauer also makes the point that the admission of weak science is better than no science:

Bad science is worse than good science, but not necessarily worse than the non-science that lurks in the heads of judges and jurors. And flawed science may still be superior to the superstitions and urban legends that influence so much of public policymaking and legal decision-making.¹⁷³

The notion that questions of fact, credibility, and law must devolve to the discretion of a JOJ both normalizes and reflects a pathological attitude in society that 'there are no right or wrong answers' to these questions, when in fact there are legally and factually right and wrong answers. Legally, the right answer is the edict of a democratically elected legislator within a constitutional framework. From a factual point of view, the correct answer is the scientific one, meaning there are either known errors or *we simply do not know*, which is often the right answer.

C. When and Why Polygraph Evidence Should be Admissible in Canadian Criminal Law

Canadian courts should consider polygraph evidence when it can be useful to determine the facts of a case. Not all polygraph evidence should be admissible, however, since polygraph evidence will not always be useful for this purpose. Courts should screen polygraph evidence for threshold reliability using the *Daubert* factors elicited in *R v J-LJ*,¹⁷⁴ focusing mainly on the

¹⁶⁹ *Ibid.*

¹⁷⁰ *Ibid.*

¹⁷¹ O D Hamza, *supra* note 132.

¹⁷² B eland, *supra* note 10.

¹⁷³ Schauer, *supra* note 14 at 36.

¹⁷⁴ *Supra* note 129.

methodology, and only secondarily on the conclusions. Polygraph evidence may be excluded on the basis that it is solely proffered to establish the probity of witness, since whether a witness is honest or deceiving does not prove anything. It does not even prove that the witness is truthful or lying, but at best that they have the intention to do so. Rather, polygraph evidence should be used to uncover hidden information that can lead to or constitute evidence. In short, polygraph data and analysis, and the conclusions derived therefrom, can be very useful in some cases, and may even prove guilt or innocence beyond a reasonable doubt.

The difference between polygraph evidence and the JOJ's assessment of the credibility of a witness is that the JOJ will assess a witness' credibility based on their experience or intuitions, whereas polygraph evidence actually measures specific reactions and systematically analyzes them. The difference, in short, is between a qualitative and a quantitative assessment of the witness' reactions.

Not only should Canadian courts be wary of *what kind* of polygraph evidence is used when screening it for threshold reliability, but they should also be wary of *how* it is used or for what purpose it is submitted. For instance, there is a world of difference between using polygraph evidence to bolster one's credibility—as it is used in a number of Quebec family law cases¹⁷⁵—and using it to uncover hidden information. For instance, in a number of family law cases in Quebec, litigants take the polygraph test. If the results are in their favor, they submit them to the court as evidence they are truthful. If, on the other hand, the results are not in their favor, then they simply do not submit the evidence. This use of polygraph evidence results in a sort of judicial confirmation bias where polygraph evidence only serves to demonstrate that witnesses are honest, when, by the same reasoning, it should also be used to show that witnesses are dishonest. Moreover, this use of polygraph evidence offends the rule against self-serving evidence, which is a well established rule of evidence that should be invoked to exclude this use of polygraph evidence.

Another illegitimate use of polygraph evidence is to determine the ultimate issue of guilt or innocence. As aforementioned, polygraph evidence in and of itself cannot prove a person is lying or truthful, except through the analysis of a paired test or by comparison with other evidence. At best, polygraph evidence can show that a person shows some signs that tend to be associated with honesty or deception, or that they react strongly to some statements instead of others. Additional evidence and analysis are required to draw conclusions from the data. However, it is worth noting that the data itself often points directly to the culprit. For instance, tests involving mock crimes show that spikes in systolic blood pressure often do correlate to lying. However, it should be borne that even if polygraph evidence could prove that a person is being honest or deceptive, this is not the same as proving that the person is telling the truth or lying. This is something only actual evidence of the claim can do. In other words, polygraph evidence measures a person's physiological responses to certain stimuli such as relevant questions, which may correlate with their so-called *state of mind* and thereby point to certain hidden information. However, polygraph evidence does not establish the actual *state of reality*. The inference from mendacity to falsehood is a *non sequitur*.

With that, it should be noted that while this article impugns the JOJ system, it does not impugn the most well established rules of evidence, notably those which are not based on the presumption of the fallibility of the JOJ (of which there are many) but on hard-learned lessons from the history of judicial error. For instance, the rule against character evidence, which is based on the Prohibited Inference (“he is the type of person who would do it, therefore he did it”), is a 'good rule,' as it outlaws a spurious inference. Conversely, the rule against similar fact evidence, which is based on the fear that its admission would cause moral or reasoning prejudice to the JOJ, is a 'bad rule,' as it finds no fault in the evidence itself, but in the JOJ's ability to come

175 See, e.g., *J(M) c M(S)* (QCCS) 2016; *B c A* (QCCS) 2015; and *G(J) c B(F)* (QCCA) 2015.

to the right conclusion regarding it. (Professor Lisa Dufraimont defends the “jury-centered view” that “evidence rules are primarily directed at constraining jury decision making and preventing jury error”¹⁷⁶ against the “truth-seeking” view, which holds that the rules of evidence developed to support the search for truth. The best conclusion, however, is that both types of rules exist and that the jury-centered rules are generally bad, since they try to compensate for dysfunctions in the JOJ system, while the truth-seeking ones are good, since they support the search for truth.)

The position of this article with regard to the rules of evidence and the JOJ system is that while some of the rules of evidence are good (notably the truth-seeking ones), the rules of evidence as a whole are not sufficient safeguards against judicial error, contrary to McLachlin's contention that they are. In short, the rules of evidence are not sufficient to compensate for the inherent flaws of the JOJ system. A major flaw is that the JOJ system rejects polygraph evidence, which is more reliable than it, because it impinges on its basic function. However, while the polygraph is more accurate than a JOJ, it cannot replace the JOJ system on its own, since it cannot prove the facts of a case but only measure's a subject's responses to relevant information in order to uncover hidden information, which can be checked. Hence, this article proposes an alternative method of corroborated polygraph testing (CPT), which combines polygraph testing with the Corroboration Rule and its underlying theory or philosophy of logical empiricism to produce a modern, scientific process that is less expensive, testable and repeatable, humane, simple, systematic, and above all more reliable.

IV. Corroborated Polygraph Testing: A Viable Alternative

The previous sections discussed a number of flaws with the current JOJ system, where a JOJ is basically given to assess the credibility of witnesses based on their presence and testimony in court. This section, which is divided into four subsections, proposes corroborated polygraph testing (CPT) as a viable alternative to the JOJ system. The first subsection answers what is CPT. The second and third subsections explain the philosophy of logical empiricism and the principle of corroboration, the Corroboration Rule, and the connection between them. The fourth section discusses some of the major advantages of CPT over the JOJ system.

A. What is CPT?

CPT is a methodology which consists of corroborating or falsifying results or conclusions from a polygraph test with independent material evidence (IME). The ideal CPT report would be something like a 10 page science publication that outlines the theory, data, methods, discussion, and conclusion (being the verdict) in a publicly accessible journal of 'Justice' (like 'Nature') that anyone can check, as opposed to endless courtroom transcripts and the unpredictable (and often *ad hoc*) reasons of judges or non-existent reasons of juries.

The theory behind CPT is to combine polygraph testing with the Corroboration Rule, which is a rule based on the philosophy of logical empiricism—notably Karl Popper's principle of corroboration¹⁷⁷—that corresponds, to some extent, to the Corroboration Rule in Scots Law and Biblical law, as subsection C will explain. The Corroboration Rule, as it is understood in the framework of CPT, requires that for anything to count as evidence it has to be corroborated by IME. This means any piece of evidence must be *doubly* corroborated. For instance, polygraph results, or the conclusion they are supposed to entail, must be corroborated by (1) an independent polygraph test that yields the same results and conclusion upon analysis and (2) IME from an entirely separate source, such as evidence that proves the conclusion is correct, thereby

176 Lisa Dufraimont, “Evidence Law and the Jury: A Reassessment” (2008) 53 McGill LJ 199.

177 Popper elaborates the principle of corroboration in his seminal book: *The Logic of Scientific Discovery* (London: Hutchinson, 1959).

tremendously decreasing the odds it is spurious. Put simply, results and conclusions prove nothing unless IME corroborates or falsifies them. The following subsection will elaborate the philosophy of logical empiricism and its connection to the principle of corroboration and the common law.

B. Logical Empiricism and the Principle of Corroboration

While not stated explicitly, the philosophy of logical empiricism is the official philosophy of the common law. This is proven by two fundamental rules of evidence, namely the hearsay rule and the rule against opinion evidence. The hearsay rule is a central rule in common law, which may be crassly described as “he said she said” type of evidence or, more technically, what cannot be cross-examined in court. The rule against opinion evidence, on the other hand, basically requires a witness to testify only concerning what he or she heard and saw. Anything else constitutes an inadmissible opinion or hearsay. In short, these rules establish that the common law only wants to hear what witnesses saw and heard, and it wants to test this evidence in court by way of cross-examination. In other words, the common law is only interested in testable observation-statements. It does not want to hear about witness' feelings, beliefs, thoughts, or opinions, to which lawyers may object and which can be excluded from the record. It may be argued that expert opinion is an exception to the rule against opinion evidence. However, the term 'expert opinion' is misleading, since an expert opinion is supposed to summarize the observations of experts in a given domain.

A philosopher will immediately recognize this philosophy as logical empiricism, all the more since most of the major architects and theorists of the common law, such as Jeremy Bentham, John Henry Wigmore, and Oliver Wendell Holmes easily fit the description of logical empiricists. Science and the common law are thus basically based on the same theory or philosophy of logical empiricism.¹⁷⁸ However, they do not apply the same methods. Whereas cross-examination is used to test evidence in court, science actually tests evidence in the field or lab. Why, then, do science and the common law use these different methods of testing when they share pretty much the same logical empiricist values? One possible explanation is that lawyers lack scientific or technical training, and hence the ability or will to treat legal problems as scientific or technical ones. Another possible explanation is that legal problems are not scientific or technical ones and therefore must be solved by other means.

If the first explanation is correct, then the problem of reconciling legal methodology with scientific methodology is a pedagogical one. However, it may be, supposing the second explanation to be correct, that legal claims are not susceptible to what philosopher Chauncey Wright calls “direct verification” (as opposed to “indirect verification”¹⁷⁹). Direct verification consists of simply checking the claim—for instance, checking if there is a coat in the closet to confirm or refute the claim 'there is a coat in the closet.' Indirect verification, on the other hand, consists of inference to the best explanation. For instance, if Mandy did not see Jacob catch his flight at 5 AM, she can check his tickets to confirm that he did. However, this is not the same as seeing Jacob catch his flight, which is direct verification. Most legal claims can only be checked indirectly, since they often concern past events that are not repeatable or observable.

However, some of the most respectable sciences, such as astronomy, make inferences to the past based on theories that are testable in the present and future. Underlying predictions about the past (or retrodictions) is a principle of symmetry, which holds that retrodictions cannot be

178 In their study entitled “Convergent Evolution in Law and Science: The Structure of Decision-Making under Uncertainty” (2011) 10 L Prob & Risk 133, social psychologists Michael Saks and Samantha Neufeld discuss the convergent evolution of law and science toward similar evidential standards despite having developed in different social environments.

179 Chauncey Wright, *Philosophical Discussions* (New York: Burt Franklin, 1971 [1877]).

more reliable than predictions. The law may adopt a similar principle where the reliability of a witness may be inferred from their reliability in the past. This is consistent with the philosophy of logical empiricism, which entails that a witness is an instrument like any other. For instance, from a logical empiricist point of view, there is no difference between a human witness and a camera except that “the camera don't lie,”¹⁸⁰ to quote the lyrics of Daniel Powter's hit song *Bad Day*. That is to say, the only difference between humans and cameras, from an evidential standpoint, is that they are often less reliable than cameras.

Logical empiricism has a lot of other implications that should be salutary to the common law. For instance, it entails that the prosecution theory should be clear and falsifiable; otherwise, the defense cannot defend itself (it cannot refute the Crown's argument), which defeats the purpose of a trial or test. It also entails that the defense need only propose an alternative hypothesis that the prosecution cannot eliminate beyond a reasonable doubt. The advantage of this theoretical approach is that the true hypothesis can never be eliminated beyond a reasonable doubt, which gives the innocent defendant who tells the whole truth and nothing but the truth an absolute advantage. In short, science—which is based on the philosophy of logical empiricism—is a philosophy which is designed to eliminate all kinds of falsehood, although it can never eliminate the truth.

While it would be nice if trials followed all the rules of logical empiricism, in reality they do not. Cross-examination, which is supposed to test witnesses, can branch off in many directions, leading to an outcome of guilty or not guilty. For instance, a witness can say something stupid that might incriminate him in the eyes of the JOJ. Had the witness been quiet, they might have walked away with their credibility intact. The JOJ system is quite random this way, which is why lawyers can rarely predict the outcome of a trial, which is why they go to trial in the first place, and which is why there is always a loser in a trial.

A method like CPT is more consistent with the philosophy of logical empiricism that the common law and science share, because it treats witnesses not as prophets or oracles by taking them on their word, but literally as instruments, like cameras, whose reactions are measured and compared to IME that corroborates or refutes them. Measurement entails error, which means it is possible to put a sigma value on conclusions derived from the data. The sigma value could be less than one (that is, less than 68 percent accurate), which is not particularly remarkable in science, but it beats a guess, which is the current situation.

C. The Corroboration Rule

The Corroboration Rule is the most elementary rule of logical empiricism. In the words of comedian Ali G: “Check yourself before you wreck yourself.”¹⁸¹ Of course, the rule is more sophisticated, and this subsection will provide a brief overview of its history and implications.

At the beginning of the 20th c., British philosopher Alfred Ayer formulated the principle of [empirical] verification,¹⁸² as the most elementary rule of logical empiricism, and hence the criterion to distinguish between scientific and unscientific theories. (This principle was later replaced by, or evolved into, the principle of corroboration, which will later be discussed.) The principle of verification failed, since only the most rudimentary observation-statements can be verified. Even such seemingly simple claims as 'the sky is blue' cannot be verified, because the sky is gray or black at times and some people are colorblind. Karl Popper revolutionized logical empiricism by formulating a new principle of falsification,¹⁸³ which initially replaced the

180 See Daniel Powter's song “Bad Day” in his eponymous album *Daniel Powter* (USA: Warner Bros, 2005).

181 See Ali G, “Booyakasha” (10 July 2011), online: YouTube <<https://www.youtube.com/watch?v=b00lc92lExw>>.

182 Ayer elaborates this principle in his influential book *Language Truth, and Logic* (London: Gollancz, 1936).

183 Karl Popper elaborates this principle in his influential book *Conjectures and Refutations: The Growth of Scientific Knowledge* (London: Routledge, 1963).

principle of verification as a criterion to demarcate scientific theories from unscientific ones. According to this principle, a scientific theory is a definite theory which is empirically falsifiable. To test a theory, therefore, means to try to falsify it. (A theory that is not falsifiable is not scientific to begin.) For Popper, a theory is more scientific as it is more precise, makes more and riskier predictions, and is easier to refute. A major problem with the principle of falsification, however, is that it rules out strong inductive claims like 'swans are white,' even when only one in a thousand swans is black. (This is especially problematic since most scientific theories, including the best, are strong inductive claims.) So, Popper ultimately formulated the principle of corroboration at mid-century, according to which a claim can be strongly or weakly corroborated, even if falsified, depending on its scope, the number and rigor of tests to which it has been subjected, and the ratio of their successes and failures.

While this article champions the Corroboration Rule *à la* Popper, as it was just described, a more prosaic (or less technical or philosophical) version of this rule can be found in Scots Law,¹⁸⁴ which requires that each crucial fact should be supported by at least two different and independent sources of evidence in order to convict someone.¹⁸⁵ The Corroboration Rule can also be interpreted in the context of logical empiricism to imply that a person cannot be condemned on a word, but rather must be condemned on a fact. (This stems from the logical empiricist principle that testimony is not evidence, but merely points to evidence at best.)

The Corroboration Rule also has ancient precedents in the Bible and the Qur'an, which are quoted below.

Torah: “On the testimony of two or three witnesses a person is to be put to death, but no one is to be put to death on the testimony of only one witness.”¹⁸⁶

The gospel according to Matthew: “[E]very matter may be established by the testimony of two or three witnesses.”¹⁸⁷

Qur'an: “Call two of your men to witness.”¹⁸⁸

That Scotland is the only country that applies the Corroboration Rule, which embodies the most basic principle of logical empiricism and modern science, is surprising, as one would expect all modern societies to apply this rule. If nothing else, Canadian courts should at least adopt this rule, although this article more boldly proposes CPT as an alternative to the JOJ system. CPT simply consists of collecting hidden information from suspects and comparing it with IME. The odds a witness is truthful will be very low indeed if they persistently answer 'no' with a strong reaction to the question: 'Did you hide the cookie in the cupboard?' and the cookie is subsequently found there.

The JOJ system, by contrast, is at least ineffective (since a trial is unnecessary if the proof is decisive), at best arbitrary, and at worst unconscionable. When the SCC ruled in *R v NS*¹⁸⁹ that it is normally expedient to show one's face in court, even in a rape case, it basically confirmed that physiognomy is standard courtroom practice. (The case spun off into a matter of reasonable accommodation instead of evidence, which should have been the real issue.) Had NS and the complainant simply taken a paired polygraph test, which might have also pointed to decisive

184 Gareth Rose, "The Corroboration Rule, Unique to Scots Law" (2 October 2011), *Scotland on Sunday*, online: <scotlandonsunday.scotsman.com/scotland/Controversial-Scottish-law-escapes-review.6846161.jp>.

185 "Consultation Issued on Scots Law After Cadder Ruling" (2 October 2011) *BBC News Scotland*, online: <<http://www.bbc.co.uk/news/uk-scotland-13004941>>.

186 Deuteronomy 17:6.

187 Matthew 18:16.

188 Qur'an 2:282. See also verse 5:106.

189 2012 SCC 72, [2012] 3 SCR 726 [*NS*].

IME, the case could have been resolved without the need for any courtroom drama. This would be quite a relief, both to the system and parties, since rape allegations are the worst cases of 'his word vs. hers' and 'who is the JOJ going to believe,' which highlights the hazards of the current system.

D. Advantages of CPT Over the JOJ System

This section elicits five practical advantages of CPT over the JOJ system, that derive from the theoretical advantage of its being a scientific procedure that is subject to scientific scrutiny, *viz.*,

1. it enables lawyers to outsource evidentiary issues to independent experts, thus enabling them to focus their time and resources on legal issues rather than factual ones in which they are not experts;
2. it is not sensitive to variations in the modalities of a trial, such as the persons, feelings, or actions of the lawyers, witnesses, or JOJs, or the general context of the trial, such as the timing or venue;
3. it can be challenged by re-testing;
4. it asks only relevant and comparison questions, instead of variable open-ended or loaded questions which can go off on protracted, embarrassing, or humiliating tangents; and
5. the results and conclusions need to be corroborated by independent tests and IME to qualify as evidence.

1. Outsourcing Evidentiary Issues

CPT is an expert system, which means that independent polygraph examiners and investigators will be responsible to discover and disseminate the facts. The facts should be published, so that the public can check or challenge them. Lawyers can then use these facts to prepare a legal case, which is within their expertise, instead of arguing the facts, which is not their expertise.

2. Not Sensitive to Variations in the Modalities of a Trial

CPT is not sensitive to variations in the modalities of a trial, such as the persons, feelings, or actions of lawyers, witnesses, judges or jurors, or the general context of the trial, such as the timing or venue. For instance, it is not sensitive to whether the witness wears tie dye rags, spiral glasses, and a funnel hat, or a 'smart' outfit. It is not sensitive, either, to what 'stupid' things a witness might say or might have said, or whether the JOJ got up on the wrong side of bed that morning or was out for blood on that particular day.

A trial, by contrast, ultimately depends on what convinces the JOJ, which can depend on many factors. For instance, it can depend on whether they like or hate the client. It can also depend on how the 'argument' develops (if they actually follow the argument), although even that may have very little impact on the final judgment. For instance, the facts that transpire from a trial depend on the evidence, which depends on the questions the lawyers think to ask, the arguments and motions they think to make, and the objections they think to raise. What becomes of these arguments, motions, and objections in turn depends on whether the judge decides to allow or dismiss the motion or appeal, to sustain or overrule the objection, to admit or exclude the evidence, *etc.* All of these factors combine to make trial a real game of lottery or "jeopardy of life and limb," as the American Constitution aptly calls it. By contrast, CPT is supposed to deliver consistent and publicly verifiable results that could otherwise be challenged as

inconsistent or false and be dismissed.

3. Re-Testing

A simple way to challenge a polygraph test is to re-take it, which will either confirm or nullify the results, and therefore the analysis, unless IME proves it beyond a reasonable doubt. For instance, if an initial polygraph test points to damning evidence that is found, then even if the test is successfully challenged it will no longer be relevant because of the damning evidence. With that, it is worth bearing that polygraph data on its own is not evidence, unless analysis of shows that it proves something, as in the example of Ginton's¹⁹⁰ method of paired polygraph testing based on background information one of the subjects is lying or concealing information.

4. Relevant and Comparison Questions

Unlike direct and cross-examination, polygraph testing only asks relevant and comparison questions which have to do with the culpable act. It does not ask irrelevant questions that may waste a lot of time and resources, or which may be embarrassing to a person or harmful to their reputation, as often occurs during direct and cross-examination of witnesses. Unlike direct and cross-examination, polygraph testing is not a traumatizing process. It is very discrete and straightforward and easily repeated. CQT sometimes does not even use complete phrases to ask questions but merely repeats certain words to gauge the subject's reaction to them and narrow down an object or location that IME sometimes confirms.

5. Corroboration by Independent Material Evidence

CPT presumes that polygraph data does not prove that a person is lying or truthful, or even that they are being honest or deceitful. Rather it seeks to corroborate or refute with IME hidden information (or what may be called 'suppressed testimony') that may be gleaned from the data. This makes CPT a scientific process.

It is worth explaining the term IME, however, which may cause some confusion to readers who may conflate this article's use of the term 'material evidence' with its legal usage. By 'material evidence,' this article means what is called 'physical evidence' at law. (Lawyers use the term 'material' to refer to what pertains to a matter at issue.) Hence, corroborating evidence has to be material (that is, physical) in keeping with the logical empiricist principle that testimony is not evidence ("Talk is cheap"), but at best predicts or points to evidence. Similarly, a scientific theory is not evidence, but what predicts or points to evidence. Indeed, the notion of 'physical evidence' is redundant, since what is not physical is not evidence. This does not mean witnesses are to be ignored according to logical empiricism. Rather, they must either be tested for reliability, as a camera is tested for reliability, or their statements must be tested and corroborated, as a theory is tested and corroborated.

Conclusion: Why Don't You Take the Polygraph?

The title of this article is a pun on VIA Rail's viral commercial: "Why don't you take the train?"¹⁹¹ which asks why people do not change their habits. The title question: "Why don't you take the polygraph?" similarly asks why the justice system does not change its habits.

This article is premised on the notion that the JOJ system does not work and needs to be

¹⁹⁰ Ginton, *supra* note 94.

¹⁹¹ See, e.g., "VIA Rail - Computer & Ashley Perkins," online: YouTube <<https://www.youtube.com/watch?v=EzogiBfayA>>.

replaced—not just that it is broken and needs to be fixed. CPT is a possible alternative to consider, although more research needs to be done in order to design a more accurate and efficient system of justice. The current system of justice largely remains as it has for centuries not because it is effective at fact-finding (it is literally the same system that tried the Salem Witches), but because it is so firmly entrenched in societal custom. In the words of Oliver Wendell Holmes:

Judges commonly are elderly men, and are more likely to hate at sight any analysis to which they are not accustomed, and which disturbs repose of mind, than to fall in love with novelties. Every living sentence which shows a mind at work for itself is to be welcomed. It is not the first use but the tiresome repetition of inadequate catch words upon which I am observing,—phrases which originally were contributions, which, by their felicity, delay further analysis for fifty years. That comes from the same source as dislike of novelty,—intellectual indolence or weakness,—a slackening in the eternal pursuit of the more exact.

As Holmes puts it, the measure of objective improvement is “substitution of quantitative for qualitative judgments,”¹⁹² for “[I]t is of the essence of improvement that we should be as accurate as we can.”¹⁹³ Even if this goal cannot be fully realized, it is still worthy to pursue. Again, in the words of Holmes:

Very likely it may be that with all the help that statistics and every modern appliance can bring us there never will be a common wealth in which science is everywhere supreme. But it is an ideal, and without ideals what is life worth?¹⁹⁴

Like Holmes, the author concludes that

I have tried to show by examples something, of the interest of science as applied to the law, and to point out some possible improvement in our way of approaching practical questions in the same sphere.¹⁹⁵

The ideal of a justice system where science is “everywhere supreme”¹⁹⁶ is worth elaborating. A trial has to be scientific. A person should not be convicted on evidence that could not be published in a science journal. What cannot be published in *Nature* should not be published in *Justice*. In the words of American Supreme Court judge Stephen Breyer:

The search [for truth] is not a search for scientific precision. We cannot hope to investigate all the subtleties that characterize good scientific work. A judge is not a scientist, and a courtroom is not a scientific laboratory. But consider the remark made by the physicist Wolfgang Pauli. After a colleague asked whether a certain scientific paper was wrong, Pauli replied, “That paper isn’t even good enough to be wrong!” Our objective is to avoid legal decisions that reflect that paper’s so-called science. The law must seek decisions that fall within the boundaries of scientifically sound knowledge.¹⁹⁷

192 Holmes, *supra* note 5 at 455.

193 *Ibid* at 456.

194 *Ibid* at 462.

195 *Ibid* at 462-463.

196 *Ibid* at 462.

197 *Reference Manual on Scientific Evidence*, 3rd ed. (Washington DC: The National Academies Press, 2011) at 4.

However, some critics may argue that science sets too high a standard to find anyone guilty. Or, they may argue that legal facts are not amenable to scientific scrutiny; for instance, because they concern claims about the past or a person's intentions, which are not empirically verifiable. However, it may be answered that all of these issues come down to the problem of dealing with our own ignorance about the facts. The point worth emphasizing is that 'we don't know' does not mean 'there are no right or wrong answers' and therefore 'someone should decide,' which is the premise of the JOJ system. Rather, that 'we don't know' is always the right answer absent scientific proof to the contrary. This does not mean that only information above two sigma (95 percent accuracy) should be admissible in court. Rather, it means that only information whose accuracy is known should be admissible. Polygraph testing, for example, has known errors, and the data and conclusions can be published and checked. By contrast, the credibility assessments of JOJs do not have known errors, are not based on published data, and their conclusions are impossible to check. (The reasons of judges are unpredictable, unverifiable, and unfalsifiable. Juries, on the other hand, do not have to give reasons.)

Although polygraph testing may not yield perfect results or conclusions, it is still better than any unscientific alternative. For instance, even if polygraph testing is supposed to be 5 percent accurate, then knowing this and knowing it is the most accurate method available, makes it a legitimate basis for judgment, unlike a JOJ's verdict that is impossible to check and whose accuracy is unknown. This relates to professor Frederick Schauer's point that weak or 'bad' science is better than none.

Bad science is worse than good science, but not necessarily worse than the non-science that lurks in the heads of judges and jurors. And flawed science may still be superior to the superstitions and urban legends that influence so much of public policymaking and legal decision-making.¹⁹⁸

In short, preserving the JOJ system merely substitutes the “mystique of science”¹⁹⁹ with that of the JOJ; that is, it substitutes the apparent infallibility of science with the actual infallibility of the JOJ. The JOJ, for its part, is anything but infallible. It is rather, to borrow the words of German philosopher Friedrich Nietzsche, “human, all too human,”²⁰⁰ especially when its beliefs and prejudices are not made public and subjected to scientific scrutiny. Where people fail in their judgment (as they often do), they say: 'I am certain.' Where science fails in its judgment (as it often does), it says: 'I don't know.'

198 Schauer, “Can Bad Science Be good Evidence? Lie Detection, Neuroscience and the Mistaken Conflation of Legal and Scientific Norms” (2010) 95 Cornell Law Review at 36.

199 *Béland*, *supra* note 10.

200 Friedrich Nietzsche, *Human, All Too Human: A Book for Free Spirits*, translated by R J Hollingdale (UK: Cambridge University Press, 1996 [1878]).