

# *R v Bingley* and the Importance of Scientifically Guided Legal Analysis

Jason M. Chin & Helena Likwornik\*

*In R v Bingley, the Supreme Court considered a controversial subjective methodology used by police officers trained as drug recognition experts (DREs) pursuant to the Criminal Code. At issue was the admissibility of these experts' evidence. A 5–2 majority held that Parliament conclusively established the reliability of the DRE program's methodology and the DRE's qualifications to perform that methodology. In this article, the authors argue that the majority's holding is problematic for multiple reasons. First, the majority's statutory interpretation is insensitive to the science behind the drug recognition program. The Court's analysis puts this subjective methodology on the same footing as objective forms of evidence, like breathalyzer analysis, where human judgment and bias play almost no role. Second, the majority's decision comes on the heels of recent findings that several forensic scientific disciplines are not as reliable as they purport to be. Finally, the majority's reasoning seems largely driven by concerns about judicial economy, and in particular the worry that evaluating DREs would take too much court time. In response, the authors attempt to provide a more scientifically rigorous but less time-consuming way for trial judges to scrutinize DREs.*

---

\* Jason M. Chin, Lecturer, TC Beirne School of Law, University of Queensland. Helena Likwornik—the views expressed by Helena Likwornik represent only her personal views and are not intended to reflect the views of any institution(s) with which she is associated. We thank Jocelyn Bosse, Itiel Dror, Caitlin Goss and Hamish McDonald for their thoughtful and perspicacious feedback. This research was supported by a start-up grant from the University of Queensland awarded to the first author.

## Introduction

### I. The DREP as a Subjective Forensic Science

### II. Scientifically Informed Legal Analysis

### III. What Happened to Contested Science?

### IV. A (Scientific) Path Forward

The unmistakable overall trend of the jurisprudence, however, has been to tighten the admissibility requirements and to enhance the judge's gatekeeping role.

—*White Burgess Langille Inman v Abbott and Haliburton Co* (2015)<sup>1</sup>

## Introduction

The Supreme Court, reflecting on the past twenty years of expert evidence jurisprudence, made the above statement in 2015. It was an accurate assessment. Canadian appellate courts have been increasingly concerned with the dangers of expert evidence, especially when that evidence comes from scientists “cloaked under the mystique of science”<sup>2</sup> and police officers whose authority can provide a “superficial attractiveness”.<sup>3</sup> As a result, these courts have reminded trial judges of their duty to vet expert evidence before it reaches the trier of fact and have accordingly empowered them to do so.<sup>4</sup> The Supreme Court's recent decision in *R v Bingley* tested the extent of this trend towards expanding the trial judge's gatekeeper role and seems to have identified its limit.<sup>5</sup> We suggest that it is not a scientifically sound limit.

---

1. *White Burgess Langille Inman v Abbott and Haliburton Co*, 2015 SCC 23 at para 20, [2015] 2 SCR 182 [*White Burgess*].

2. *R v Moban*, [1994] 2 SCR 9 at 21, 114 DLR (4th) 419. See also *R v DD*, 2000 SCC 43 at para 41, [2000] 2 SCR 275.

3. *R v Sekhon*, 2014 SCC 15 at para 50, [2014] 1 SCR 272, citing *R v Sekhon*, 2012 BCCA 512 at para 27, 565 WAC 170. For a review of this jurisprudence, see Lisa Dufrainmont, “New Challenges for the Gatekeeper: The Evolving Law on Expert Evidence in Criminal Cases” (2012) 58:3&4 *Crim LQ* 531.

4. Note, however, that several commentators have questioned whether courts regularly make use of these powers. See generally Gary Edmond & Kent Roach, “A Contextual Approach to the Admissibility of the State's Forensic Science and Medical Evidence” (2011) 61:3 *UTLJ* 343; Gary Edmond & Emma Cunliffe, “Cinderella Story? The Social Production of a Forensic ‘Science’” (2016) 106:2 *J Crim L & Criminology* 219 at 264–70. Similarly, apparently expert opinion evidence is sometimes characterized as lay opinion in Australia, New Zealand, the US and Canada, thus skirting the expert gatekeeper. See Gary Edmond & Mehera San Roque, “Quasi-Justice: Ad Hoc Expertise and Identification Evidence” (2009) 33:1 *Crim LJ* 8; Edward J Imwinkelried, “Distinguishing Lay from Expert Opinion: the Need to Focus on the Epistemological Differences Between the Reasoning Process Used by Lay and Expert Witnesses” (2015) 68:1 *SMU L Rev* 73; Jason M Chin, Jan Tomiska & Chen Li, “Drawing the Line Between Lay and Expert Opinion Evidence” 63:1 *McGill LJ* [forthcoming in 2017].

5. 2017 SCC 12, [2017] 1 SCR 170.

The stakes in *Bingley* go well beyond the case itself. It arises in a time of upheaval within the forensic sciences that has seen peak scientific bodies, including the United States National Academy of Sciences (NAS), call into question the reliability of several long established forensic disciplines.<sup>6</sup> Similar issues abound in the basic sciences.<sup>7</sup> Simply put, this is not the time to punctuate a two decades long trend of empowering trial judges to question science. Trial courts need the flexibility to question a body of knowledge that is in flux.

The science at issue in *Bingley* was a standardized but subjective method employed by specially trained police officers, drug recognition experts (DREs), to determine whether a suspect is drug impaired (collectively, the Drug Recognition Expert Program or DREP). DREs perform a twelve-step assessment pursuant to sections 254 and 254.1 of the *Criminal Code*<sup>8</sup> and the *Evaluation of Impaired Operation (Drug and Alcohol) Regulations*.<sup>9</sup> This legislative scheme came into force in 2008. It provides a sister regime for laws in place to detect and enforce alcohol-impaired driving (e.g., breathalyzers to measure blood alcohol concentration). Under the DREP, a police officer with a reasonable suspicion that a motorist is under the influence of a drug may demand that motorist be examined by a DRE.<sup>10</sup> Pursuant to the Regulations, DREs must have taken courses with and been accredited by the International Association of Chiefs of Police (IACP), a non-governmental organization based in the United States.<sup>11</sup>

The majority and dissent disagreed as to whether the *Criminal Code* had established that a properly administered twelve-step DREP assessment was sufficiently reliable to serve as evidence of impairment in court, or alterna-

---

6. US, National Research Council, *Strengthening Forensic Science in the United States: A Path Forward* (Washington, DC: National Academies Press, 2009) [NAS Report].

7. See D Stephen Lindsay, “Replication in Psychological Science” (2015) 26:12 *Psychological Science* 1827; Marcus R Munafò et al, “A Manifesto for Reproducible Science”, online: (2017) 1 *Nature Human Behaviour* 0021 <[www.nature.com/nathumbehav](http://www.nature.com/nathumbehav)>; Denes Szucs & John PA Ioannidis, “Empirical Assessment of Published Effect Sizes and Power in the Recent Cognitive Neuroscience and Psychology Literature” (2017) 15:3 *PLOS Biology* 1; Florian Prinz, Thomas Schlägel & Khusru Asadullah, “Believe It or Not: How Much Can We Rely on Published Data on Potential Drug Targets?” (2011) 10:9 *Nature Reviews Drug Discovery* 712.

8. RSC 1985, c C-46; *Bingley*, *supra* note 5 at para 4. Note that this program has recently received considerable negative media attention in the US for alleged false positives. See e.g. Brendan Keefe & Michael King, “The Drug Whisperer: Drivers Arrested While Stone Cold Sober” 11 *Alive NBC News* (24 October 2017), online: <[www.11alive.com/news/investigations/the-drug-whisperer/437061710](http://www.11alive.com/news/investigations/the-drug-whisperer/437061710)>.

9. SOR/2008-196 [*Impaired Operation Regulations*]; *Bingley*, *supra* note 5 at para 4.

10. For a review of the DREP, see Amy J Porath-Waller, Douglas J Beirness & Erin E Beasley, “Toward a More Parsimonious Approach to Drug Recognition Expert Evaluations” (2009) 10:6 *Traffic Injury Prevention* 513.

11. *Ibid* at 514.

tively whether it was still open to scrutiny by the trial judge. The dissent agreed with the accused: the text of the DREP provisions contrasted with those surrounding breathalyzers and thus should not be treated the same in court.<sup>12</sup> The breathalyzer provisions render them “conclusive proof”<sup>13</sup> of intoxication, whereas there is no such language with regard to DREs’ evidence. The majority held otherwise: while a DRE’s conclusions may be scrutinized on limited grounds, the twelve-step process itself is irrefutably valid and reliable.<sup>14</sup>

We propose that the statutory interpretation in *Bingley* should have been guided by the science. It is not just that the words in the breathalyzer and DREP provisions are different; *so is the science*. Unlike breathalyzer results, the DREP is subjective and thus there may be vast differences between DREs in their ability and skill in applying the methodology. Further, errors in application are more frequent in subjective methodologies.

In the remainder of this Part, we will review the facts of *Bingley* and the Court’s reasons. Part I situates the DREP within the forensic sciences. Building on that foundation, Part II argues that Parliament did not intend for the DREP to elude the trial judge’s gatekeeper role because of its subjective methodology and because Parliament focused on its power to identify the presence of drugs as opposed to its sizable false positive rate. In Part III, we discuss *Bingley*’s ramifications for expert evidence law writ large. And, finally, in Part IV, we offer a solution to manage DRE evidence in court that is scientifically rigorous but also judicially economical.

In *Bingley*, the police were called after the accused was observed driving erratically. Pursuant to the DREP, he was evaluated by Constable Jellinek, a certified DRE.<sup>15</sup> DREs, including Constable Jellinek, employ a standardized subjective methodology to determine whether a target is impaired by drugs.<sup>16</sup> The methodology comprises twelve steps, including judgments of physical coordination, balance, pupil size, eye gaze, muscle tone and pulse.<sup>17</sup> If the evaluation affords “reasonable grounds” to believe the suspect’s ability to drive is impaired by a drug, the DRE may then demand a sample of blood, saliva or

---

12. *Bingley*, *supra* note 5 at para 54; *R v Bingley*, 2017 SCC 12, [2017] 1 SCR 170 (Factum of the Appellant at paras 44–59) [*Bingley* Factum].

13. *Criminal Code*, *supra* note 8, s 258(1)(c).

14. *Bingley*, *supra* note 5 at paras 27–28, 30, 32.

15. *Ibid* at para 2.

16. See Douglas J Beirness, Jacques LeCavalier & Deanna Singhal, “Evaluation of the Drug Evaluation and Classification Program: A Critical Review of the Evidence” (2007) 8:4 *Traffic Injury Prevention* 368 at 369; *Bingley*, *supra* note 5 at para 9.

17. *R v Bingley*, 2015 ONCA 439 at para 12, 126 OR (3d) 525 [*Bingley* ONCA], *aff’d Bingley*, *supra* note 5; Porath-Waller, Beirness & Beasley, *supra* note 10 at 514.

urine as confirmatory evidence.<sup>18</sup> The Constable concluded that Mr. Bingley was under the influence of a drug.<sup>19</sup> A urinalysis revealed the presence of THC, but was inconclusive as to when that exposure occurred.<sup>20</sup> At trial, Bingley successfully challenged the reliability of Constable Jellinek's subjective evaluation.<sup>21</sup>

The case eventually made its way to the Supreme Court of Canada<sup>22</sup> where McLachlin CJC, writing for the majority, held that the evidence of a DRE should be admitted without a *voir dire*.<sup>23</sup> The majority's analysis of the text, context and legislative intent<sup>24</sup> of the DREP provisions concluded that the qualifications of DREs had been "conclusively and irrebuttably established by Parliament"<sup>25</sup> and the associated scientific validity "comes from the statutory framework itself".<sup>26</sup> It was common ground at the Supreme Court that the bodily fluid analysis component of the DREP was admissible.

Justice Karakatsanis, writing for herself and Gascon J, disagreed; they would have held that DRE evidence is open to scrutiny by trial judges.<sup>27</sup> More specifically, they could find no wording in the DREP provisions, nor their legislative history, to indicate that Parliament intended that a DRE's evidence should evade the trial judge's gatekeeping function. As noted above, the DREP provisions were textually distinct from those covering breathalyzer results, the latter being admissible due to express language in the *Criminal Code* rendering them "conclusive proof [of] the concentration of alcohol in the accused's blood".<sup>28</sup> Rather, Parliament envisioned the DREP as an investigatory tool

---

18. *Criminal Code*, *supra* note 8, s 254(3.4).

19. *Bingley*, *supra* note 5 at para 2.

20. *Bingley* ONCA, *supra* note 17 at para 15.

21. *Ibid* at para 24.

22. The procedural history is lengthy but it can be summarized as follows: Bingley was acquitted in two trials and in both cases the Crown successfully obtained new trials by way of summary conviction appeal. The first acquittal was overturned because the trial judge failed to consider the "cumulative effect of the evidence". At the second trial, the trial judge found that Constable Jellinek's opinion that Bingley's driving was impaired due to drugs was inadmissible because it lacked a scientific basis. He was once again acquitted. This resulted in a second summary conviction appeal, in which the judge ordered a new trial because Constable Jellinek should not have been subjected to a *Mohan voir dire*. Bingley appealed the second summary conviction order to the Court of Appeal for Ontario, which dismissed the appeal. The appellate court found that both the object and text of section 243(3.1), which provides that the DRE is to "determine" whether a driver is drug impaired, obviates a *voir dire*. A new trial was ordered. *Ibid* at paras 6–31, 44–56. Bingley appealed to the Supreme Court.

23. *Bingley*, *supra* note 5 at paras 13–33.

24. *Ibid* at paras 23–28.

25. *Ibid* at para 27.

26. *Ibid* at para 24.

27. *Ibid* at paras 56–59.

28. *Supra* note 8, s 258(1)(c); *ibid* at paras 44–59.

and thus approved its reliability for that purpose.<sup>29</sup> Accordingly, “courts retain discretion to require — through evidence or precedent — confirmation that the science behind DRE evaluations meets the necessary level of reliability before admitting the evidence at trial”.<sup>30</sup>

Understanding the divergence between the majority and dissent requires some background into the science of the DREP. In the following section, we will situate the DREP within forensic science, a field that has received recent criticism for failure to provide empirical support for its claims. The DREP is a subjective forensic science; unlike more objective methods, its conclusions are profoundly affected by human judgment.

## I. The DREP as a Subjective Forensic Science

The DREP is a forensic scientific method. Forensic science (i.e., science as it applies to criminal law) has been deeply criticized over the past several decades for the use of unreliable methods and the failure of organizations of practitioners to self-police.<sup>31</sup> These criticisms have come from both scientific bodies and the legal community. In 1998, former justice Fred Kaufman reported the results of his investigation into the wrongful conviction of Guy Paul Morin.<sup>32</sup> He found that flawed forensic science, particularly hair and fibre identification, contributed to the jury’s decision to convict Morin.<sup>33</sup> Ten years later, Stephen Goudge J, as he then was, detailed numerous grave failures in the practice of forensic pediatric pathology in Ontario.<sup>34</sup> Both jurists recommended trial judges play a stronger role as gatekeepers of forensic science.<sup>35</sup> In other words, they should refuse to admit such evidence when it lacks a valid scientific foundation.

---

29. *Bingley*, *supra* note 5 at para 40.

30. *Ibid.*

31. See e.g. Barack Obama, “The President’s Role in Advancing Criminal Justice Reform” (2017) 130:3 *Harv L Rev* 811 at 860–62.

32. Ontario, *The Commission on Proceedings Involving Guy Paul Morin: Report* (Toronto: Ministry of the Attorney General, 1998) vol 1 (The Honourable Fred Kaufman, C.M., Q.C.) [Morin Report].

33. *Ibid* at 83.

34. Ontario, *Inquiry into Pediatric Forensic Pathology in Ontario: Report* (Toronto: Ministry of the Attorney General, 2008) vols 1–4 (The Honourable Stephen T Goudge) [Goudge Report].

35. Morin Report, *supra* note 32. The report states: “In my respectful view, however, it is appropriate for trial judges to undertake a far more critical analysis of the admissibility of this kind of evidence. My own view is that hair comparison evidence of the kind introduced in the *Morin* case should rarely be admitted for inclusionary purposes.” *Ibid* at 323. See also Goudge Report, vol 3, *supra* note 34 at 496–502.

The Supreme Court responded to these miscarriages of justice by changing expert opinion evidence law. The Court expanded the role of trial judges, directing them to review the validity of scientific evidence (i.e., how it was tested, its error rate, whether it was peer reviewed and published, and whether it is generally accepted) when that evidence is novel science.<sup>36</sup> Later in *R v Trochym*<sup>37</sup> and *White Burgess Langille Inman v Abbott and Haliburton Co (White Burgess)*, the Court added that “contested science” should also receive this scrutiny.<sup>38</sup>

Since these inquiries into specific convictions and practices, empirical research has confirmed that invalid forensic science is present in about sixty percent of wrongful convictions.<sup>39</sup> These sobering findings have inspired renowned scientific bodies to investigate the shortcomings in forensic scientific theory and practice that may be accountable. For example, in 2009, the NAS published a report (NAS Report), led by prominent judges and scientists, describing a three-year study into the state of forensic science.<sup>40</sup> This report was remarkably critical: “[T]he forensic science system exhibits serious shortcomings in capacity and quality; yet the courts continue to rely on forensic evidence without fully understanding and addressing the limitations of different forensic science disciplines”.<sup>41</sup> The report provided a framework for change, including recommending increased funding to the forensic sciences and further research.

In response to the NAS Report, President Obama directed his President’s Council of Advisors on Science and Technology (PCAST) to conduct additional research on how forensic science should be improved. PCAST published a report in 2016 (PCAST Report).<sup>42</sup> The PCAST Report found that

---

36. See *R v J-LJ*, 2000 SCC 51 at paras 33–36, [2000] 2 SCR 600; *Daubert v Merrell Dow Pharmaceuticals Inc*, 509 US 579 (1993) [*Daubert*].

37. 2007 SCC 6 at paras 32–36, [2007] 1 SCR 239.

38. *Supra* note 1 at para 23.

39. See Brandon L. Garrett & Peter J. Neufeld, “Invalid Forensic Science Testimony and Wrongful Convictions” (2009) 95:1 Va L Rev 1 at 9.

40. NAS Report, *supra* note 6. For two cogent discussions of the NAS Report, see Jennifer L. Mnookin, “The Courts, the NAS, and the Future of Forensic Science” (2010) 75:4 Brook L Rev 1209; David H. Kaye, “Probability, Individualization, and Uniqueness in Forensic Science Evidence: Listening to the Academies” (2010) 75:4 Brook L Rev 1163.

41. NAS Report, *supra* note 6 at 53.

42. US, President’s Council of Advisors on Science and Technology, *Forensic Science in Criminal Court: Ensuring Scientific Validity of Feature-Comparison Methods* (Washington, DC: Executive Office of the President, 2016) [PCAST Report]. For additional reviews of the forensic sciences, see Edmond & Roach, *supra* note 4; Gary Edmond et al, “Contextual Bias and Cross-Contamination in the Forensic Sciences: The Corrosive Implications for Investigations, Plea Bargains, Trials and Appeals” (2015) 14:1 L Probability & Risk 1 [Edmond et al, “Contextual Bias”].

deep flaws still exist in several forensic methodologies, both with respect to their foundational validity (i.e., can the method accurately make the identification it purports to make?) and their validity-as-applied (i.e., can practitioners consistently and accurately apply the methodology in the field?).<sup>43</sup>

Particularly relevant to the DREP, the PCAST Report found that deficiencies in validity ran deeper for subjective methods, which rely on human judgment to a greater extent than objective methods.<sup>44</sup> For example, bitemark analysis, like the DREP, is highly subjective, relying on a practitioner's comparisons of a found bitemark to known dentition.<sup>45</sup> On the other hand, DNA analysis, which relies on automated comparisons performed by computer algorithms, is predominantly objective.<sup>46</sup>

Foundational validity requires two things: (1) consistent and reproducible methods, and (2) a measure of accuracy.<sup>47</sup> For the first requirement, the methodology must be systematic and return the same result across different examiners. A high degree of methodological rigour is especially important for subjective methods because human judgment is vulnerable to numerous cognitive biases.<sup>48</sup> Examiners must follow a procedure that does not vary based on their subjective judgments. For instance, the fingerprint misidentification of the perpetrator of the 2004 Madrid train bombing case is widely considered an instance of a subjective methodology led astray by motivated reasoning.<sup>49</sup> In that case, FBI investigators reported a match with one hundred percent certainty, but were mistaken.<sup>50</sup> Subsequent investigations found the investigator took advantage of an overly flexible methodology to come to the outcome he was seeking.<sup>51</sup>

---

43. PCAST Report, *supra* note 42 at 7–14.

44. *Ibid* at 46–47.

45. Michael J Saks et al, “Forensic Bitemark Identification: Weak Foundations, Exaggerated Claims” (2016) 3:3 *JL & Biosciences* 538 at 540–41.

46. PCAST Report, *supra* note 42 at 69–75.

47. *Ibid* at 47–54.

48. See Edmond et al, “Contextual Bias”, *supra* note 42; NAS Report, *supra* note 6 at 122–24, 184–85; Goudge Report, *supra* note 34 at 387–90; PCAST Report, *supra* note 42 at 31; Itiel E Dror, “A Hierarchy of Expert Performance” (2016) 5:2 *J Applied Research in Memory & Cognition* 121; Itiel Dror & Robert Rosenthal, “Meta-Analytically Quantifying the Reliability and Biasability of Forensic Experts” (2008) 53:4 *J Forensic Science* 900. For a review of the dangers posed by expert testimony from a psychological perspective, see Michael J Saks & Barbara A Spellman, *The Psychological Foundations of Evidence Law* (New York: New York University Press, 2016) at 202–41.

49. See PCAST Report, *supra* note 42 at 28; US, Federal Bureau of Investigation, *Report on the Erroneous Fingerprint Individualization in the Madrid Train Bombing Case* (2005) 7:1 *Forensic Science Communications* 1.

50. PCAST Report, *supra* note 42 at 90.

51. *Ibid*.



The DREP, while providing a useful systematic procedure, is still subjective. Akin to fingerprint identification, there is a procedure to follow. Each step, however, engages human judgment. Several steps are particularly subjective. For instance, in the horizontal gaze nystagmus test (HGNT),<sup>52</sup> the suspect is asked to follow a stimulus (e.g., the officer's penlight), with the DRE looking for signs of "a lack of smooth pursuit".<sup>53</sup> That judgment—what is smooth and what is not—admits a great deal of subjectivity. Indeed, research has demonstrated that several contextual factors affect the reliability of the HGNT.<sup>54</sup> Similar subjectivity is found in the DRE's assessment of balance and the suspect's ability to divide his or her attention.<sup>55</sup>

Moreover, a DRE's ultimate determination is a discretionary balancing of the totality of the steps.<sup>56</sup> As admitted by Constable Jellinek at Bingley's trial, there is no standardized scoring system or mechanism to weight the steps.<sup>57</sup> When examiners have such latitude, there is a greater chance that they will focus on the criteria that confirm their suspicions.<sup>58</sup> In fact, even highly trained scientists have admitted to preferentially relying on and reporting the data points that support their hypothesis when they have the latitude to do so<sup>59</sup>—what one might refer to as "researcher degrees of freedom".

The accuracy component of foundational validity can be measured in multiple ways. In criminal law, which seeks to avoid wrongful convictions,<sup>60</sup> the false positive rate (FPR),<sup>61</sup> or the chance of declaring a hit when there was

---

52. *Impaired Operation Regulations*, *supra* note 9, ss 2(a), 3(b)(i).

53. US, Transportation Safety Institute & National Highway Traffic Safety Administration, *Drug Evaluation and Classification Training: "The Drug Recognition Expert School"* (No HS172 R01/11) (Washington: US Department of Transportation, 2011), Session IV at 12 [DRE Manual].

54. For instance, time of day because the eye fatigue may contribute to the ability to control one's gaze. See Charles R Honts & Susan L Amato-Henderson, "Horizontal Gaze Nystagmus Test: The State of the Science in 1995" (1995) 71:3 NDL Rev 671 at 694–95.

55. See *Impaired Operation Regulations*, *supra* note 9, s 3(c).

56. See DRE Manual, *supra* note 53 at 85.

57. See Bingley Factum, *supra* note 12 at para 11(g–h).

58. See Gary Edmond et al, "Thinking Forensics: Cognitive Science for Forensic Practitioners" (2017) 57:2 Science & Justice 144 at 146–47 [Edmond et al, "Thinking Forensics"]; Itiel E Dror, David Charlton & Ailsa E Péron, "Contextual Information Renders Experts Vulnerable to Making Erroneous Identifications" (2006) 156:1 Forensic Science Intl 74.

59. See Leslie K John, George Loewenstein & Drazen Prelec, "Measuring the Prevalence of Questionable Research Practices with Incentives for Truth Telling" (2012) 23:5 Psychological Science 524.

60. See *R v Hart*, 2014 SCC 52 at para 8, [2014] 2 SCR 544. See also Edmond & Roach, *supra* note 4.

61. In other sciences, the FPR is known as "Type I Error". See Jacob Cohen, "A Power Primer" (1992) 112:1 Psychological Bull 155 at 156.

no hit, is especially important.<sup>62</sup> In the context of drug testing, the FPR is the probability of detecting a drug when there was actually no meaningful amount present in the suspect's system. The PCAST Report notes: "Methods with a high FPR are scientifically unreliable for making important judgments in court . . . To be considered reliable, the FPR should certainly be less than 5 percent and it may be appropriate that it be considerably lower, depending on the intended application."<sup>63</sup>

Alternatively, consider a method's sensitivity,<sup>64</sup> which is the chance of reporting a hit when there is a hit in reality. In the DREP context, it is the probability of identifying a person under the influence of drugs when they are, in fact, under the influence of drugs. Sensitivity is a very important metric for investigative tools. For instance, imagine if DREs only identified 450 of 900 cases of actual inebriation, a sensitivity of 50%. In such cases, half of those drivers who were actually under the influence would evade the bodily fluid test and go free. Still another measure is the overall accuracy rate: the number of true positives and true negatives divided by the total number of cases.<sup>65</sup> It is important to note that these measures of accuracy are quantitatively distinct—a sensitive and overall accurate measure may still carry a high FPR.

To make this point more concrete, consider this hypothetical example, the results of which are summarized in Table 1. Assume that in a population of 1,000 people tested, 900 have actually taken drugs and 100 have not. Now, assume that DREs identify 820 as having taken the drug and 180 as not. On these assumed facts, the DREs are demonstrating an overall accuracy of 88% and sensitivity of 88.89%. As shown in greater detail in Table 1, this is because overall accuracy is the number of true positives (800) and true negatives (80) divided by the total number of observations (1,000). And sensitivity is the number of true positives (800) divided by the total number of drug-influenced individuals (900).

An overall accuracy of 88% and sensitivity of 88.89% are impressive, but these metrics do not tell the whole story. On these assumed numbers, the same DREs also have a FPR of 20%. Recall that FPR is the number of false positives (20) divided by the number of people not actually under the influence (100). And in the criminal law context, a FPR of 20% is a cause for concern.

---

62. See PCAST Report, *supra* note 42 at 151–52.

63. *Ibid.*

64. In other sciences, sensitivity is known as statistical power. See Cohen, *supra* note 61 at 155.

65. See Douglas J Beirness, Erin Beasley & Jacques LeCavalier, "The Accuracy of Evaluations by Drug Recognition Experts in Canada" (2009) 42:1 *Can Society Forensic Science J* 75 at 78.

**Table 1**

In this hypothetical study of DRE accuracy, sensitivity is the number of true positives (800) divided by the total number of drug-influenced individuals (900) tested and thus 88.89%. Overall accuracy is the number of true positives (800) and true negatives (80) divided by the total number of observations (1000) and thus 88%. The FPR is the number of false positives divided the number of people not under the influence of a drug (100) and thus 20%.

		<u>DRE's Opinion</u>	
		<u>Positive</u> (820 total)	<u>Negative</u> (180 total)
Actual presence of drug	<u>Positive</u> (900 total)	800 (true positive)	100 (false negative)
	<u>Negative</u> (100 total)	20 (false positive)	80 (true negative)

After determining that a methodology is foundationally valid, validity-as-applied must be established. In other words, it must be verified that the person performing the method in the current case is capable of accurately doing so in circumstances resembling that case. Subjective methods like the DREP are particularly at risk for failures of applied validity because of what psychologists call “individual differences”: humans vary on several dimensions related to drug recognition, like natural ability and motivation to study and work harder.<sup>66</sup> Measuring applied validity is typically accomplished through “proficiency testing”: the assessment of individual analysts to see if they can accurately employ the method.<sup>67</sup>

Proficiency tests must themselves be performed rigorously.<sup>68</sup> The PCAST Report recommends that proficiency testing should be administered by arm’s length parties, noting substantial deficits in proficiency testing in the fields of latent fingerprint analysis and firearms analysis (i.e., determining if ammunition is associated with a certain firearm).<sup>69</sup>

---

66. See Bradford T Ulery et al, “Accuracy and Reliability of Forensic Latent Fingerprint Decisions” (2011) 108:19 Proceedings National Academy Sciences 7733 at 7733–38; Jonathan J Koehler, “Fingerprint Error Rates and Proficiency Tests: What They are and Why They Matter” (2008) 59:5 Hastings LJ 1077; David White et al, “Perceptual Expertise in Forensic Facial Image Comparison” (2015) 282:1814 Proceedings Royal Society B.

67. PCAST Report, *supra* note 42 at 56–59.

68. See *ibid* at 58.

69. *Ibid* at 57–58.

Assuming that the DREP is foundationally valid, it still may be that police officers are not good at applying it.<sup>70</sup> They may outweigh the results of certain parts of the test, or rely too much on a preconceived notion of whether the subject was drug impaired. The latter, which is also known as confirmation bias—forcing incoming information to fit a pre-existing theory<sup>71</sup>—could flow from a variety of sources, such as the type of neighbourhood the suspect was detained in, the suspect’s race, or any other idiosyncratic contextual cue.<sup>72</sup> Moreover, like any skill, the ability to employ the DREP may erode over time. For example, the DRE may see his or her opinion corroborated by a subsequent urine analysis several times in a row and rely on this limited sample to become overconfident and thus not follow the method as rigorously as before. Or similarly, the DRE may begin to rely only on a subset of the twelve DREP steps.

## II. Scientifically Informed Legal Analysis

The above review demonstrates that, from a scientific standpoint, the DREP is a fundamentally subjective methodology. As a result, it raises a different set of concerns than more objective tests, such as breathalyzers. Therefore, we contend that it was a mistake for *Bingley*’s majority to place these types of evidence on such similar footing for two general reasons. First, due to its subjectivity, the DREP’s applied validity (i.e., are the DREs accurately and reliably applying the method) is much more tenuous and thus requires more judicial oversight. And second, while the DREP may be good at detecting drug use (i.e., high sensitivity), its FPR appears to be considerably worse. This suggests that it is a useful investigative method, but a dangerous form of inculpatory evidence in court.

As an objective method, the results of a breathalyzer demand a different type of scrutiny by courts. Like DNA analysis, breathalyzer data is analyzed in a predominantly automated way, with little opportunity for subjective judg-

---

70. See *R v Bingley*, 2017 SCC 12, [2017] 1 SCR 170 (Factum of the Intervener Canadian Civil Liberties Association). Although they did not use the term “proficiency test”, the importance of such tests accords with the Canadian Civil Liberties Association’s (CCLA) factum at the Supreme Court of Canada. In their factum, the CCLA questioned: “What if, in the officer’s experience evaluating individuals who are suspected of being impaired by drugs, that officer’s assessment is regularly not borne out the toxicological tests?” *Ibid* at para 27.

71. See Raymond S Nickerson, “Confirmation Bias: A Ubiquitous Phenomenon in Many Guises” (1998) 2:2 *Rev General Psychology* 175 at 175.

72. See Edmond et al, “Contextual Bias”, *supra* note 42.

ment.<sup>73</sup> As a result, once foundational validity is established—perhaps through Parliament’s review of empirical studies—the primary adjudicative issue is whether the methodology was properly followed. For example, was there a failure in the chain of custody?<sup>74</sup> Indeed, courts may exclude breathalyzer results if the device was not properly calibrated.<sup>75</sup> By analogy, DNA analysis of single-source samples are widely considered foundationally valid,<sup>76</sup> but can be unreliable if the methodology is not properly followed such as via “sample mix-ups, contamination, incorrect interpretation, and errors in reporting”.<sup>77</sup> For instance, in March 2017 the Attorney General for Western Australia reported that twenty-seven criminal convictions were in doubt for failure of a police lab to follow proper DNA testing procedures.<sup>78</sup>

Subjective methods, which rely on human judgment, are qualitatively different. Whereas computer programs can be expected to run identically across crime labs and countries, human judgment varies a great deal by individual. Therefore, because Parliament deemed breathalyzers “conclusive proof”<sup>79</sup> and did not do the same with the DREP, it stands to reason that this was not mere inadvertence but a deliberate choice to treat different types of science differently. In other words, empowering the police to use the DREP should not be seen as endorsement of its unlimited use in court. The DREP is a much more fragile form of evidence.

The majority did not seem to recognize the importance of applied validity. In particular, they made a lack of applied validity a matter of weight, as opposed to admissibility: failing to “conduct the drug recognition evaluation in accordance with his or her training” is not a grounds for excluding the DREP evidence.<sup>80</sup> As peak scientific bodies have noted, applied validity is just as essential to the truth of an expert’s claim as foundational validity; the results of a foundationally valid

---

73. See PCAST Report, *supra* note 42 at 46–47.

74. See *ibid* at 50–51.

75. See e.g. *R v Carles*, 2003 CarswellOnt 4962 (WL Can) at para 53–60, [2003] OJ No 4860 (QL) (Ct J) (in this instance the Court found the breathalyzer test reliable).

76. See PCAST Report, *supra* note 42 at 71–73.

77. *Ibid* at 73.

78. See Graeme Powell, “Leading DNA Scientist Sacked, 27 Criminal Convictions in Doubt, WA Attorney-General Says”, *ABC News* (31 March 2017), online: <[www.abc.net.au/news/2017-03-31/sacking-of-was-leading-dna-scientist-27-criminal-cases-in-doubt/8403618?pfmredir=sm](http://www.abc.net.au/news/2017-03-31/sacking-of-was-leading-dna-scientist-27-criminal-cases-in-doubt/8403618?pfmredir=sm)>.

79. *Criminal Code*, *supra* note 8, s 258(1)(c).

80. *Bingley*, *supra* note 5 at para 32. Rather, the majority held that the DRE’s evidence could only be excluded when its probative value was outweighed by its prejudicial effect. They suggested this may occur when a DRE was completely unable to explain how he or she came to the conclusion. *Ibid* at para 30.

procedure mean nothing if the expert cannot or will not faithfully apply it.<sup>81</sup> Moreover, as a matter of law, failure or inability to follow a methodology is a matter of admissibility, not just weight.<sup>82</sup>

Finally, the absence of a known FPR associated with the DREP suggests that trial judges ought not abdicate their authority to assess its foundational validity and reliability. The Crown, in its submissions to the Court, noted that Parliament appeared to accept a study finding the DREP possessed a 98.6% accuracy rate.<sup>83</sup> A review of the Parliamentary Committee hearing transcript leaves it unclear whether that is an estimate of overall accuracy or sensitivity.<sup>84</sup> But in any case, the Committee's focus on this metric and lack of any evidence that they turned their minds to the FPR further supports a continuing gatekeeper role for trial judges in the context of DREs. This analysis coincides with the dissent's broader point: Parliament seems to have only vetted the DREP's reliability as an investigative method,<sup>85</sup> for which sensitivity and overall accuracy are the salient measures. They did not consider whether it should provide confirmatory evidence to be presented directly to judges and juries, for which FPR is the most relevant measure.

Furthermore, and drawing again on Table 1, the state of the DREP science is, in fact, largely consistent with that table. Several studies have shown that DREs identify drug impairment with generally high overall accuracy and sensitivity.<sup>86</sup> But, at the same time, they do so with a false positive rate that ranges from 8%,<sup>87</sup> to 20%,<sup>88</sup> to as high as 57%.<sup>89</sup> One review of the existing

---

81. See PCAST Report, *supra* note 42 at 56–59. See also NAS Report, *supra* note 6 at 8, 206–08.

82. See *R v Abbey*, 2009 ONCA 624 at para 119, 97 OR (3d) 330 [*Abbey* 2009]; *R v Srikananda*, 2016 ONCJ 667. In the fresh evidence application following the second *Abbey* trial, the Court of Appeal for Ontario considered fresh evidence casting doubt on the methodology of the Crown's expert sociologist. The Court determined that the fresh evidence was cogent enough to have affected the verdict. It would have likely led to the exclusion of the sociologist's evidence for failure to follow proper methodology. See *R v Abbey*, 2017 ONCA 640 at paras 107–32, 350 CCC (3d) 102. Thorough scientific vetting of the expert's method at first instance would have produced substantial judicial economy and saved *Abbey* from years in prison as his case progressed through two lengthy trials and appeals. *Ibid.*

83. *Bingley*, *supra* note 5 at para 56.

84. House of Commons, Standing Committee on Justice and Human Rights, 39th Parl, 1st Sess, No 077 (14 June 2007) at 13.

85. See *Bingley*, *supra* note 5 at para 46.

86. See e.g. Beirness, LeCavalier & Singhal, *supra* note 16 (reviewing several studies).

87. See US, National Highway Traffic Safety Administration, *Identifying Types of Drug Intoxication: Laboratory Evaluation of the Subject Examination Procedure* (DOT HS 806 753) (Washington, DC: US Department of Transportation, 1985) at 12.

88. See Beirness, Beasley & LeCavalier, *supra* note 65 at 78.

89. See David Shinar & Edna Schechtman, "Drug Identification Performance on the Basis of Observable Signs and Symptoms" (2005) 37:5 *Accident Analysis & Prevention* 843 at 847.

research suggests that FPR rates are typically higher in studies when officers have reason to believe “that most participants had ingested drugs”.<sup>90</sup> The same can be said for the method as currently employed because DREs will only be called upon when the investigating officer has a reasonable suspicion of drug impairment. In fact, some Canadian trial courts have rejected the DRE evidence because of its high false positive rate.<sup>91</sup> A statutory interpretation exercise that was sensitive to the science of the DRE might have taken a closer look at whether Parliament considered the method’s FPR, and concluded Parliament did not intend such a fallible (but sensitive) method to evade judicial scrutiny.

### III. What Happened to Contested Science?

Implicit in the above analysis was a normative argument: not only did the majority fail to account for several scientific distinctions in its interpretive exercise, but in doing so they may have allowed unreliable science into trial courts across the country. This move counters a decades-long trend that has increased the scope of scientific evidence that trial judges are directed to scrutinize: “The unmistakable overall trend of the jurisprudence, however, has been to tighten the admissibility requirements and to enhance the judge’s gatekeeping role.”<sup>92</sup> This trend has moved in tandem with and taken notice of meta-scientific insights from works like Goudge J’s Inquiry and the NAS Report. We will now briefly review this trend in law and suggest the majority’s decision in *Bingley* represents a step backwards.

The move towards enhanced gatekeeping of scientific evidence can be traced back to *R v Mohan*, which still stands as Canada’s leading expert evidence decision.<sup>93</sup> Most are familiar with *Mohan*’s four well-trod requirements for expert evidence.<sup>94</sup> But Sopinka J, writing for the Court in *Mohan*, also started Canadian jurisprudence down a new path for scientific evidence by holding that novel science should be subjected to “special scrutiny” as to its reliability.<sup>95</sup> A year earlier, the US Supreme Court had made the same move in *Daubert v*

---

90. Beirness, LeCavalier & Singhal, *supra* note 16 at 372.

91. See e.g. *R v Wood*, 2007 ABQB 503, 426 AR 335; *R v Thomas*, 2012 BCPC 215.

92. *White Burgess*, *supra* note 1 at para 20.

93. *Supra* note 2.

94. They are: logical and legal relevance; necessity in assisting the trier of fact; absence of an exclusionary rule; and a properly qualified expert. *Ibid* at 20.

95. *Ibid* at 25.

*Merrell Dow Pharmaceuticals Inc (Daubert)*.<sup>96</sup> It rejected deference to the general acceptance of science among experts, and instead demanded that trial judges assess the reliability and validity of the science prior to admitting it.<sup>97</sup> Justice Sopinka's decision, at least implicitly, aligned Canadian jurisprudence with that case.

Any doubts about whether Canada had adopted *Daubert* were put to rest by the Supreme Court in *R v J-LJ*.<sup>98</sup> In that case, the majority of the Court employed *Daubert*'s four factors as special scrutiny to exclude novel scientific evidence that purported to identify pedophilia on the basis of a psychological and physiological testing procedure.<sup>99</sup> Those factors are (1) whether and how the science has been tested; (2) whether and how the science has been peer reviewed and published; (3) the error rate associated with the science; and (4) general acceptance of the science in the field in which it is situated.<sup>100</sup>

*R v Trochym* expanded *Daubert*'s ambit in Canada, noting that given science's evolutionary nature, even once-accepted theories and practices (i.e., non-novel) may be unreliable: "Therefore, even if it has received judicial recognition in the past, a technique or science whose underlying assumptions are challenged should not be admitted in evidence without first confirming the validity of those assumptions."<sup>101</sup>

The majority then applied the *Daubert* factors to the science of hypnotically recovered memories and ruled the memories derived from such procedures inadmissible, despite their use in previous decisions and the theoretical plausibility of the method that produces them.<sup>102</sup> Put simply, modern psychological science had found the method unreliable in application. In *White Burgess*, the Supreme Court reaffirmed the applicability of the *Daubert* factors "in the case of an opinion based on novel or *contested science* or science used for a novel purpose".<sup>103</sup>

The majority's view in *Bingley* seems to undercut several decades of appellate court guidance directing trial judges to closely scrutinize scientific evidence. In particular, in *Bingley*, the Chief Justice stated that the DREP represented novel science, but exempted it from judicial scrutiny due to its

---

96. *Supra* note 36.

97. *Ibid.* For a concise review of *Daubert* and the US cases that preceded and followed it, see David L Faigman, John Monahan & Christopher Slobogin, "Group to Individual (G2i) Inference in Scientific Expert Testimony" (2014) 81:2 U Chicago L Rev 417 at 427–31.

98. *Supra* note 36.

99. *Ibid* at para 33.

100. *Daubert*, *supra* note 36 at 593–95.

101. *Supra* note 37 at para 32.

102. *Ibid* at paras 55–61.

103. *Supra* note 1 at para 23 [emphasis added].



place in the *Criminal Code*.<sup>104</sup> This is troubling because the DREP fits more clearly into the category of *contested science*.<sup>105</sup> Despite the original DREP studies dating back to 1985, several recent studies have demonstrated that the DREP suffers from a high FPR.<sup>106</sup> Furthermore, trial courts in Alberta, New Brunswick, and British Columbia have engaged in thoughtful examinations of the DREP and found it lacks a sufficient scientific foundation.<sup>107</sup> In particular, the Alberta Court of Queen's Bench, in excluding DREP evidence, stated:

The low accuracy reported in the 1996 study is significant. Whether the cause is the DRE protocol's inability to enable evaluators to accurately identify the physiological reactions caused by particular classes of drugs, the reliability of the science on which it is based, the evaluators' training, or something else is well beyond the scope of this appeal. What is apparent, however, is the fallibility of the DRE, which strikes at the heart of its reliability.<sup>108</sup>

But perhaps more troubling than the failure to characterize the DREP as contested science is the fact that the majority seemed to altogether remove the contested science language from its enunciation of the type of science that will trigger *Danbert* scrutiny: "Such knowledge is *required only where the science is novel*."<sup>109</sup> While the dissent did recognize the above case law as "unsettled"<sup>110</sup> and would have subjected the DREP to a *voir dire*, it also failed to pick up on the notion of contested science from *Trochym* and *White Burgess*.

This is not the time to narrow the scope of judicial gatekeeping. Recent meta-scientific insights from the NAS and PCAST have demonstrated that many long-standing assumptions about the scientific validity of the forensic sciences were simply wrong. And more generally, new findings from the basic sciences demonstrate that several fields have not adequately policed themselves and thus many established findings may be spurious.<sup>111</sup> Even old science can be invalid. This is all compounded by the fact that the "novel science" distinction proved unworkable. Some courts defined it as science not generally accepted

---

104. *Bingley*, *supra* note 5 at para 23.

105. See *Trochym*, *supra* note 37 at para 32. Justice Deschamps stated that: "While some forms of scientific evidence become more reliable over time, others may become less so as further studies reveal concerns. Thus, a technique that was once admissible may subsequently be found to be inadmissible." *Ibid.* Accord *White Burgess*, *supra* note 1 at para 23.

106. See e.g. Beirness, LeCavalier & Singhal, *supra* note 16.

107. See note 91, *above*, for examples of such cases.

108. *Wood*, *supra* note 91 at para 82.

109. *Bingley*, *supra* note 5 at para 22 [emphasis added].

110. *Ibid* at para 57.

111. See note 7, *above*, for the sources of these findings.

by scientists,<sup>112</sup> while others thought it was science not yet accepted by courts.<sup>113</sup> Other courts defined it as temporally new science.<sup>114</sup> The United States does not have the contested science standard distinction.<sup>115</sup>

## IV. A (Scientific) Path Forward

Although never expressly stated, the majority's pullback on the trial judge's gatekeeper role seems to have been driven by concerns about judicial economy. This consideration played a key role in the Crown's submissions, which worried that exposing DREs to *voir dire* would promote "litigation in trial courts across the country".<sup>116</sup> And, in fact, the procedural history in *Bingley* itself was long and included two trials.<sup>117</sup>

In what was likely a response to these concerns about judicial economy, the majority stated that before determining admissibility of expert evidence, "the trial judge must determine the nature and scope of the proposed expert opinion".<sup>118</sup> As there was no issue of scope in *Bingley*, McLachlin CJC's statements suggest that carefully scoping expert evidence represents one of the Court's preferred solutions to proliferating expert evidence. In fact, this approach traces back to Doherty J's guidance in *Abbey*<sup>119</sup> and has now been twice highlighted by the Supreme Court.<sup>120</sup> We agree that this is a step in the right direction and believe it can be modified and applied to DRE evidence in a way that retains the trial judge's gatekeeper function for such evidence.<sup>121</sup>

We propose that, assuming the DREP must be considered foundationally valid pursuant to the majority decision in *Bingley*, DREs could be questioned on a limited and predetermined set of questions focused on *validity-as-applied* to determine if their evidence should be admitted.<sup>122</sup> In fact, the DREP presents

---

112. See *Wolfen v Shaw* (1998), 43 BCLR (3d) 190 at para 17, 1998 CanLII 15046 (SC).

113. See *R v Pearce (ML)*, 2014 MBCA 70 at para 69, 318 CCC (3d) 372.

114. See e.g. *P(SF) v MacDonald*, 1998 ABQB 855 at para 14, 238 AR 175.

115. See David M Paciocco, "Context, Culture and the Law of Expert Evidence" (2001) 24:1 Adv Q 42 at 46.

116. *R v Bingley*, 2017 SCC 12, [2017] 1 SCR 170 (Factum of the Respondent at para 40).

117. For a description of this procedural history, see note 22, *above*.

118. *Bingley*, *supra* note 5 at para 17. See also *ibid* at paras 16, 29.

119. *Abbey* 2009, *supra* note 82 at paras 62–70.

120. See *Sekbon*, *supra* note 3 at paras 46–48; *Bingley*, *supra* note 5.

121. See Helena Likwornik, "Overstepping and Sidestepping: The Expert Evidence Dance" (2017) 35:4 Adv J 24 (which suggests that a requirement for the court to frame key questions for experts in advance of the trial may be useful in many contexts).

122. Indeed, the majority left open the possibility that DREs could still be excluded under the trial judge's final discretionary weighing of the costs and benefits of the evidence. See *Bingley*, *supra* note 5 at para 30.

what may be an ideal context for trial judges to perform such a function because DRE evidence is controlled by specific legislative provisions.<sup>123</sup> In other words, it is much more predictable than typical expert evidence. These questions could be put to the DRE either at the preliminary stage of delineating the scope and nature of the expert evidence or as part of pretrial case management. If no *voir dire* is held in relation to this type of evidence, the questions could even form part of the trial.

A standard set of these applied validity questions for DREs might include the following:

- When were you last tested on your DREP proficiency by an independent third party?
- What were your results on this test?
- Did you follow all the required steps in applying the DREP?
- Do you tend to rely more on a particular aspect or aspects of the test?
- Were there any conditions that were not ideal for applying the procedure?
- Were you aware of any impediments to visibility?
- Were you aware of any extraneous noises or distractions?<sup>124</sup>

This suggestion represents an economical method that is sensitive to the science of the DREP. Perhaps more importantly, it does not abandon decades of jurisprudence focused on strengthening the trial judge's gatekeeping power. In the DREP context, it also prevents professional groups (e.g., the IACP) from dictating what is reliable enough for Canadian courts. Of course, active involvement by the court continues the trend away from a strictly adversarial process. But this is a trend already underway through the gatekeeping process itself. The need for the trial judge to act as gatekeeper, both initially and throughout the process, acknowledges that the adversarial process alone does not always offer sufficient protection against misleading scientific evidence.

Well before *Bingley* was decided, McLachlin CJC provided the following quote in the introduction to the National Judicial Institute's *Science Manual for Canadian Judges*: "Without the proper tools, the justice system is vulnerable to unreliable expert scientific evidence."<sup>125</sup> While the Chief Justice was discussing scientific education when she wrote this introduction, the willingness and confidence to scrutinize investigatory methods is also an important tool. Prior

---

123. See *Criminal Code*, *supra* note 8; *Impaired Operation Regulations*, *supra* note 9.

124. See Honts & Amato-Henderson, *supra* note 54.

125. Canada, National Judicial Institute, *Science Manual for Canadian Judges* (Ottawa: National Judicial Institute, 2013) at 14.

to *Bingley*, trial courts around Canada were building a DREP jurisprudence, carefully considering whether it was reliable enough to serve as inculpatory evidence in criminal proceedings. This is a culture of judicial skepticism that should be fostered rather than tamped down, especially in this time of uncertainty within the forensic sciences.