Re-Thinking Expert Evidence

by

Joshua Assaph Moshe Freedman

A thesis submitted in conformity with the requirements
for the degree of Masters in Law (LLM)

Faculty of Law
University of Toronto

© Copyright by Joshua Assaph Moshe Freedman 2019
Re-Thinking Expert Evidence

Joshua Assaph Moshe Freedman

Masters of Law (LLM)

Faculty of Law
University of Toronto

2019

Abstract

The jurisprudence surrounding expert evidence in Canadian law has evolved since the 1994 case, *R. v. Mohan*, but it is still replete with numerous problems. This thesis challenges the current admissibility test for expert evidence in Canadian law, and suggests a new, more focused test, which places reliability as the predominant factor in admitting expert evidence. This thesis will argue that changing the admissibility test in this direction will make the jurisprudence in this area more coherent, and lead to better admissibility outcomes.
Acknowledgments

This thesis would like to thank my LLM supervisor, Professor Simon Stern, who provided excellent advice throughout the writing of this paper. The author also extends thanks to all other members of the faculty, the hard-working LLM team, other LLM colleagues, and family.
## Table of Contents

Acknowledgments.............................................................................................................. iii

Table of Contents................................................................................................................. iv

1 Introduction:...................................................................................................................... 1

2 Part I: The Framework For Admitting Expert Evidence: ............................................. 3

   2.1 General Rules Of Evidence:..................................................................................... 3

   2.2 A Developing Approach Towards Expert Evidence:.............................................. 4

      2.2.1 Abbey: .............................................................................................................. 4

      2.2.2 Beland: .......................................................................................................... 5

      2.2.3 Lavallee: ......................................................................................................... 6

      2.2.4 Marquard: ...................................................................................................... 7

      2.2.5 Conclusion: ..................................................................................................... 7

   2.3 The Framework In Mohan: ....................................................................................... 8

      2.3.1 Relevance: ....................................................................................................... 8

      2.3.2 Necessity: ....................................................................................................... 9

      2.3.3 The Absence Of Any Exclusionary Rule:......................................................... 9

      2.3.4 A Properly Qualified Expert: .......................................................................... 10

      2.3.5 Conclusion: .................................................................................................... 10

   2.4 Developing Mohan: ................................................................................................. 11

      2.4.1 R. v. J. (J.): .................................................................................................... 11

      2.4.2 R. v. D. (D.): .................................................................................................. 12
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.4.3</td>
<td><em>R. v. Trochym</em></td>
<td>13</td>
</tr>
<tr>
<td>2.5</td>
<td>The <em>Abbey</em> Test And A New Framework:</td>
<td>14</td>
</tr>
<tr>
<td>2.5.1</td>
<td><em>R. v. Abbey</em></td>
<td>14</td>
</tr>
<tr>
<td>2.5.2</td>
<td><em>R. v. Sekhon</em></td>
<td>16</td>
</tr>
<tr>
<td>2.6</td>
<td><em>White Burgess</em> And Its Aftermath:</td>
<td>16</td>
</tr>
<tr>
<td>2.6.1</td>
<td><em>R. v. Awer</em></td>
<td>17</td>
</tr>
<tr>
<td>2.6.2</td>
<td><em>R. v. Bingley</em></td>
<td>18</td>
</tr>
<tr>
<td>2.6.3</td>
<td><em>R. v. Abbey</em> (The Sequel):</td>
<td>19</td>
</tr>
<tr>
<td>2.6.4</td>
<td>Conclusion:</td>
<td>20</td>
</tr>
<tr>
<td>3</td>
<td>Part II: Evaluating the Approach Towards Expert Evidence:</td>
<td>21</td>
</tr>
<tr>
<td>3.1</td>
<td>The Good:</td>
<td>21</td>
</tr>
<tr>
<td>3.1.1</td>
<td>1-It is Tailored Towards Reaching Criminal Justice Goals:</td>
<td>21</td>
</tr>
<tr>
<td>3.1.2</td>
<td>2-It Has Improved Over Time:</td>
<td>21</td>
</tr>
<tr>
<td>3.2</td>
<td>The Bad:</td>
<td>24</td>
</tr>
<tr>
<td>3.2.1</td>
<td>1-Reliability:</td>
<td>24</td>
</tr>
<tr>
<td>3.2.2</td>
<td>2-The Characterization Problem:</td>
<td>25</td>
</tr>
<tr>
<td>3.2.3</td>
<td>3-Will <em>White Burgess</em> Really Eliminate Bias?:</td>
<td>26</td>
</tr>
<tr>
<td>3.2.4</td>
<td>4-The Gatekeeper Needs More Help:</td>
<td>27</td>
</tr>
<tr>
<td>3.2.5</td>
<td>5-The Test Is Too Flexible:</td>
<td>28</td>
</tr>
<tr>
<td>3.2.6</td>
<td>Conclusion:</td>
<td>29</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>6.1.3</td>
<td>Revisiting D. (D.):</td>
<td>46</td>
</tr>
<tr>
<td>6.1.4</td>
<td>Conclusion:</td>
<td>47</td>
</tr>
<tr>
<td>6.2</td>
<td>Conclusion:</td>
<td>48</td>
</tr>
<tr>
<td>6.3</td>
<td>Bibliography:</td>
<td>49</td>
</tr>
<tr>
<td>6.3.1</td>
<td>Articles:</td>
<td>49</td>
</tr>
<tr>
<td>6.3.2</td>
<td>Books:</td>
<td>50</td>
</tr>
<tr>
<td>6.3.3</td>
<td>Cases:</td>
<td>51</td>
</tr>
</tbody>
</table>
1 Introduction:

This thesis will engage in a comprehensive dialogue with the current admissibility test in Canadian law for expert evidence, and argue that the predominant factor in the admissibility of expert evidence in Canadian law should be the reliability of the proposed evidence. In doing so, this thesis will critique the scattershot approach currently employed in Canadian law, which seems to equally focus on a variety of factors, in turn providing a lack of guidance for trial judges to produce coherent admissibility decisions. To that end, this thesis will propose a new three-part framework for the admission of expert evidence, where the first part will be solely focused on reliability, and ask: 1) Is this field capable of being reliable, 2) Is this person qualified to give an opinion in this field, and 3) Is their opinion based on reliable methods. In the remaining two parts of the test, which in many ways will resemble the current approach, reliability will still remain the guiding ethos. In advocating for these changes, this thesis will argue that a turn to reliability will create a more legally sound criterion for admissibility, and better protect the criminal justice goals of seeking the truth, preventing wrongful convictions, and trial efficiency. In addition, these changes would curtail the all too common practice of erring on the side of admitting expert evidence, with any concerns about the evidence’s reliability going to its evidentiary weight instead of its exclusion.

While making the broader argument that the admissibility test used in Canadian law on expert evidence should be modified, this thesis will often focus on how social science evidence (and other non-scientific pieces of evidence) should be treated in particular. This thesis will do so, because it believes that the often indeterminate nature of this type of evidence most strongly illuminate the tensions in the current admissibility test.

An illustrative example of the current pathologies in this area of jurisprudence, which the new framework would seek to correct, is R. v. Gager. There, despite recognizing the numerous flaws in a gang expert’s qualifications, the trial judge allowed their testimony, finding that the “essential

---

criteria for the admission of expert evidence are met in this case.”\(^2\) Some of these flaws included the fact that the expert referred to himself in braggadocious terms in his resume,\(^3\) had not studied the gangs involved in the trial,\(^4\) destroyed much of the data underlying his opinion,\(^5\) inflated his sample size to bolster his opinion,\(^6\) altered his work to conform to the defense theory,\(^7\) and did not seek out available information which could affect his testimony.\(^8\) While it is arguable that perhaps the trial judge merely misapplied the current admissibility test, a recurring critique that this thesis will make of the current test is that its vagueness allows the justification of nearly any result. Indeed, as this thesis will argue, any admissibility test which allows such wide leeway for flawed expert opinions, should be modified towards one that will create a stricter hierarchy of values, and provide judges with more guidance in reaching their admissibility decisions.

To dive into this topic, this thesis will be structured into several sections. In the first section, this thesis will look at the evolution of the admissibility framework for expert evidence over time, and show that the jurisprudence has infrequently been focused about the reliability of such evidence. In examining this framework, this section will also delve into how the admissibility framework has been modified when applied to social science evidence. In the second section, this thesis will look at academic and legal commentary surrounding the admissibility of expert evidence, and talk about the merits and demerits of the current approach. The takeaway of this section will be that a reliability component should become preeminent in admitting expert evidence. To that end, in the third section, this thesis will advance a new admissibility test for expert evidence, and explain how it will change according to context. Moreover, this section will advance a set of principles that will guide the framework in different scenarios. In the fourth section, this thesis will show how weaknesses in the law’s current framework has led to the admission of dubious forms of evidence, and how using a test focused on reliability may avoid these problems. In the fifth section, this thesis will re-examine previously dismissed forms of evidence, in light of the changing focus of the new framework, and discuss if these types of evidence should now become admissible.

\(^2\) Ibid., at para 28
\(^3\) Ibid., at paras 34-36
\(^4\) Ibid., at para 41
\(^5\) Ibid., at para 44
\(^6\) Ibid., at paras 47-48
\(^7\) Ibid., at para 69
\(^8\) Ibid., at para 76
Part I: The Framework For Admitting Expert Evidence:

2.1 General Rules Of Evidence:

All evidence is generally admissible if it is relevant to a fact in issue, and not subject to an exclusionary rule. Historically, certain types of evidence like hearsay, character evidence, and expert evidence are presumptively inadmissible. Judges have also retained the residual discretion to exclude otherwise relevant evidence, if for example, it is more prejudicial than probative. The presumption against admitting the subject matter of this thesis, expert evidence, can be rebutted in situations where an expert witness can satisfy certain requirements established in Mohan and White Burgess. Namely, relevance, necessity, the lack of any other exclusionary rule, a properly qualified expert, and recently, the expert not being biased.

Another important component of evidence law is the idea that the trial judge is there to play a role in assessing the threshold reliability of a statement, and not its ultimate reliability. This distinction lies between the evidence’s admissibility and the reliance placed on it. Thus, although at the threshold stage “a trial judge of necessity engages in an evaluation that shares some of the features with the evaluation ultimately performed by the jury if the evidence is admitted,” a trial judge’s job is only to decide if the evidence is sufficiently reliable to be put to the jury. Then, whether the evidence is actually to be believed, namely its ultimate reliability, is supposed to be decided by the jury. Historically, judges conceived their judicial task as presenting expert evidence to the jury when the jury was unlikely to know anything about a certain subject matter, regardless of the expert’s bona fides. As a result, the law often allows evidence with significant deficiencies to be

---

10 See for example, Part III of, S. Casey Hill, Louis P. Strezos & David M. Tanovich. Canadian Criminal Evidence, 5th ed. (Toronto: Thomson Reuters Canada, 2013, loose-leaf), Online: WestlawNextCanada [McWilliams’]
11 Anderson, supra note 9, at pg. 37
12 Ibid., at pg 38
14 Abbey I, infra note 84, at para 142
16 Lederman, infra note 160, at pg 221
put before the jury, with the belief that the jury will be able to assign the proper weight to this evidence.\textsuperscript{17}

\section*{2.2 A Developing Approach Towards Expert Evidence:}

To understand why \textit{Mohan}, the leading Canadian Supreme Court case on expert evidence, gives such short shrift to reliability as an central admissibly factor, it is necessary to briefly study several precursor cases to \textit{Mohan} which would influence its 4-part test. In these cases, which tentatively set the determining factors in admitting expert evidence, this thesis will argue that the Supreme Court set its jurisprudence towards this topic on the wrong path by missing several opportunities to establish reliability as an important admissibility factor.

\subsection*{2.2.1 \textit{Abbey}:}

For example, in the 1982 case \textit{Abbey},\textsuperscript{18} the Supreme Court discussed the utility and danger of admitting expert evidence, but failed to make any mention about the importance of reliable evidence. There, the Court was dealing with a situation where a trial judge acquitted an accused (upheld by the B.C. Court of Appeals), who through the use of a psychiatric expert was able to establish an insanity defense.\textsuperscript{19} The potential utility stemming from this expert was the fact that the expert could “provide the judge and jury with a ready-made inference which the judge and jury, due to the technical nature of the facts, are unable to formulate.”\textsuperscript{20} In contrast, the danger which the trial judge strayed into was accepting the expert’s opinion “as going to the truth of the facts stated in it,”\textsuperscript{21} without supporting evidence. Having done so, the Court ruled that the trial judge erred in assigning too much weight to an opinion largely based on the accused’s own conversation with the expert.\textsuperscript{22}

\begin{flushleft}
\textsuperscript{17} David M. Tanovich. “\textit{R. v. Hart}: A Welcome New Emphasis on Reliability and Admissibility” (2014) 12 CR-ART 298 [Tanovich]
\textsuperscript{19} Ibid., at paras 5-15
\textsuperscript{20} Ibid., at para 44
\textsuperscript{21} Ibid., at para 49
\textsuperscript{22} Ibid.
\end{flushleft}
Though it was not incorrect for the Court to draw the conclusion that it was dangerous for the trial judge to accept an insanity defense based on assertions made by the accused through his expert witness, the Court’s focus on weight obscures the much clearer reason to exclude the evidence based on its unreliability. While this distinction does not seem so great, it amounts to the court preferring an overly technical answer (the admission of this evidence was in error because too much weight was given to an opinion not based in an established factual foundation), over the far more generalizable and simpler answer (the admission of this evidence was in error because there is no way to establish how reliable it is). As will be discussed going forward, one of the benefits of this thesis’ proposed new approach will be the relative simplicity of it as compared to the current test.

2.2.2 Beland:

In Beland,23 the Supreme Court again touched on reliability when ultimately disallowing the testimony of a polygraph expert on the basis that it was a form of oath-helping, and violated the rule against providing character evidence.24 For the majority, Justice McIntyre stated that the exclusion of this type of evidence did not lie in the fact it may be unreliable, as they were not presented evidence about its accuracy, but concluded that even if the polygraph test was riddled with errors, this alone would not lead to its exclusion.25 Similarly, the dissent (that would have allowed the polygraph expert’s testimony) brushed off reliability concerns raised by the Crown, saying that it was “open to the opposing party to cross-examine the operator as to the weaknesses inherent in the process and to call an opposing expert to dispute the validity or interpretation of the results.”26 As will be talked about further on, this strain of thought that the adversarial process is well equipped to weed out unreliable evidence is dangerous, and will be addressed in the proposed admissibly test.

24 Ibid., at paras 65-67
25 Ibid., at para 19
26 Ibid., at para 61
2.2.3 Lavallee:

A clearer missed opportunity to address reliability came in *Lavallee*,\(^{27}\) where the Supreme Court ruled that expert evidence on the mindset of a battered woman went beyond the typical knowledge and experience of a layperson, and thus, was necessary.\(^{28}\) In discussing this evidence, the Court uncritically referred to the work of Dr. Lenore Walker who wrote a book about the battered women phenomena, and who created a theory about the cycle of violence that is common to these relationships.\(^{29}\) While this thesis does not desire, nor is equipped, to evaluate the merits of this research, many legal scholars have questioned the court’s acceptance of this theory without delving into its reliability.\(^{30}\) And as will be talked about later, because of *stare decisis*, experts in this area are still able to testify in this area which has been deemed as necessary,\(^ {31}\) whereas far more demonstrably reliable psychological evidence has been ruled unnecessary.

This case is also important because the Court touched on a recurring point of contention in subsequent cases dealing with psychological evidence. Namely, the belief on one side of the debate, that there are certain human experiences which the trier of fact cannot draw inferences from, based on their inexperience with living through certain circumstances.\(^{32}\) On the other side, there is the competing belief “that judges and juries are thoroughly knowledgeable about ‘human nature.’”\(^ {33}\) In the former scenario, expert evidence is seen as necessary to illuminate these unique life circumstances, whereas in the latter scenario it is seen as unnecessary for the judge and jury to hear the expert’s evidence because they would not learn anything new.

\(^{27}\) *R. v. Lavallee*, [1990] 1 SCR 852 [*Lavallee*]

\(^{28}\) Ibid., at para 55

\(^{29}\) Ibid., at para 50


\(^{31}\) See for example, *R. v. Meecham*, 2019 ONSC 494, [*Meecham*]

\(^{32}\) *Lavallee*, *supra* note 27, at para 32

\(^{33}\) Ibid., at para 33
2.2.4 *Marquard*:

Last, in *Marquard*,\(^{34}\) the Supreme Court again showed a lack of concern about reliability by allowing two expert witnesses to testify far outside the scope of their expertise, raising doubts about the relevance and accuracy of their testimony. In *Marquard*, the accused was charged with assaulting her three-year-old granddaughter by placing her face on a hot stove door.\(^{35}\) The Crown called experts in pediatrics and child abuse to opine about statements made by the child after being burned, and the source of the burns. A Dr. Mian, who was only qualified as an expert in child abuse, nevertheless testified that the child’s burns could have only come from the oven door.\(^{36}\) Similarly, a Dr. Zuker, who was only qualified as an expert on burns, testified that the passivity of the child during his examination was consistent with being abused.\(^{37}\) Despite recognizing that the scope of the doctors’ testimony went beyond what they were qualified, the Court was satisfied that as practicing physicians they still had more experience than general members of the public to testify in these areas, and that any deficiencies in their expertise would go to weight.\(^{38}\)

Later, in the Goudge Inquiry, the report was sharply critical of the decision in *Marquard*, noting that allowing medical experts to testify beyond their scope was part of the reason why Charles Smith’s testimony led to a number of wrongful convictions.\(^{39}\)

2.2.5 Conclusion:

In total, these four cases cited by *Mohan*, provide context for why *Mohan* would be focused on things like relevance, necessity, and technical categorization, and not reliability or the proper scope of an expert’s testimony.

\(^{34}\) *R. v. Marquard*, [1993] 4 SCR 223 [*Marquard*]
\(^{35}\) Ibid., at para 2
\(^{36}\) Ibid., at para 31
\(^{37}\) Ibid., at para 33
\(^{38}\) Ibid., at para 35
\(^{39}\) Goudge Inquiry, *infra* note 179, at pgs 473-474
2.3 The Framework In *Mohan*:

In *Mohan*, the Supreme Court looked to many of the factors cited in the abovementioned cases, and created a four-part test to determine the admissibility of expert evidence. These four factors look to: the relevance of the expert’s opinion, its necessity in assisting the trier of fact, the absence of any other exclusionary rule, and whether the expert is properly qualified. In what this paper will argue is a crucial omission in *Mohan*, the Court did not explicitly ask whether the reliability of the evidence is a chief factor to consider, but only considered it as a secondary consideration in the reliability analysis and when the expert opinion is novel. In failing to do so, the Supreme Court started the modern jurisprudence on this issue on a troublesome course, and left it for later judgments to put some teeth into the reliability requirement.

The dispute in *Mohan* involved whether the defense could call a psychiatric expert to testify that the character traits of the defendant did not fit the psychological profile of someone who would have committed a string of sexual assaults. The expert proposed to testify that only a sexual psychopath could have committed the 4 sexual assaults in question, and that Mohan (himself a doctor), did not fit that profile. The trial judge disallowed the evidence on the basis that the expert did not establish that only a distinct group could have committed these offences, whereas the Court of Appeal for Ontario held that the trial judge should have allowed the jury to determine the weight to be given to the expert’s relevant evidence. Before delving into whether the evidence should have been admitted, the Supreme Court went into detail about each of the relevant factors in their proposed framework.

2.3.1 Relevance:

At the relevance inquiry, the Court stated that a judge must find any proposed evidence to be logically relevant, and to engage in a cost-benefit analysis of whether the evidence should be

---

41 Ibid., at paras 17-21
42 Ibid., at paras 22, 32
43 Ibid., at paras 9-11
44 Ibid., at paras 12-15
included or excluded. 45 By logically relevant, the Court meant whether the evidence related to a fact in issue, and by cost-benefit analysis, the Court meant whether the evidence’s “probative value is overborne by its prejudicial effect, if it involves an inordinate amount of time which is not commensurate with its value or if it is misleading in the sense that its effect on the trier of fact, particularly a jury, is out of proportion to its reliability.” 46 If a judge determines that the evidence does not pass this cost-benefit analysis, then they should exclude the evidence. 47

2.3.2 Necessity:

To be necessary, the proposed evidence must be more than helpful, and go beyond the experience and knowledge of the judge and jury. 48 Usually this happens when the evidence is technical in nature, and there is no realistic chance that it will be within the trier of fact’s experience. 49 While not an absolute prohibition, the Court also stated that the expert’s evidence should not typically weigh in on the ultimate issue (and would be treated strictly if it did), namely, whether they believe the accused is innocent or guilty. 50 For example, expert evidence about whether an accused was in a gang, based on various indicia, could go to the ultimate issue, whereas general testimony about the nature of gangs in Toronto would not. 51

2.3.3 The Absence Of Any Exclusionary Rule:

As mentioned above, certain types of evidence are presumptively inadmissible, so if the expert’s testimony were to conflict with the rule against hearsay or character evidence, this could lead to its exclusion. For example, if the Crown cross-examined the accused’s expert witness about the propensity of the accused to commit the crime, and the accused had not yet put their character into issue, the answers in this cross-examination would violate the character evidence rule. 52

45 Ibid., at para 22
46 Ibid.
47 Ibid.
48 Ibid., at para 26
49 Ibid.
50 Ibid., at para 29
51 See Gager, supra note 1
52 Mohan, supra note 40, at para 30
2.3.4 A Properly Qualified Expert:

Last, to be qualified as an expert, the Court held that an expert must have “acquired special or peculiar knowledge through study or experience in respect of the matters on which he or she undertakes to testify.”53 The Court also said that if the expert proposes to testify about a novel area of science or is using a new technique, the Court would apply more scrutiny to the reliability of the proposed evidence.54 Unfortunately, the Court did not provide much guidance about what exactly would constitute novel science, although cases in the future did attempt to provide some clarity on this issue.55

2.3.5 Conclusion:

Ultimately, the Court rejected the evidence of the proposed expert having regard to a combination of the above 4 factors. First, the Court agreed with the trial judge that the expert could not establish that “either the perpetrator of the crime or the accused has distinctive behavioral characteristics such that a comparison of one with the other will be of material assistance in determining innocence or guilt.”56 Here, that was the case, because a person who commits sexual offences against young women could not be said to belong to a distinctive enough group.57 Therefore, the expert’s testimony was seen to be not sufficiently reliable, and therefore unnecessary.58 Having said that, the Court did leave open the possibility that if the circumstances of a crime could only have been committed by a unique perpetrator, expert evidence might be relevant in that situation.59 For example, if the perpetrator of a crime most likely would need to show necrophilia like tendencies, expert evidence might be relevant for an accused to show they do not belong within this distinct group.60

53 Ibid., at para 31
54 Ibid., at para 32
56 Mohan, supra note 40, at para 49
57 Ibid., at para 50
58 Ibid.
59 Ibid., at para 49
60 R. v. J. (J.), 2000 SCC 51, at para 42 [J. (J.)]
2.4 Developing *Mohan*:

2.4.1 *R. v. J. (J.)*:

In *J. (J.)*, the Supreme Court further elaborated on when reliability is an important consideration, especially in the context of novel expert evidence. As in *Mohan*, the defense in *J. (J.)* wanted to call a psychiatrist to testify that only a serious sexual deviant could have committed a sexual assault against two male children, and that the defendant in this case did not fit that profile. The expert wanted to introduce results from penile plethysmograph tests that purported to show that the defendant was not aroused by images of young boys, and thus could not have committed the offence. The trial judge excluded the evidence on the basis that the expert’s testimony did not fit within the distinct group exception laid out in *Mohan*, nor did the expert provide support for his conclusions. In contrast, the majority at the Quebec Court of Appeal believed that the expert was using recognizable diagnostic tools, and that the jury should decide how much weight should be given to the expert’s contention that the accused did not have the characteristics of someone who would have committed sodomy on a child.

In reviewing whether the testimony should have been admitted, the Supreme Court advised looking at four factors, borrowed from *Daubert* (the leading American case on expert evidence), to use when deciding if a novel scientific technique (penile plethysmography) was reliable enough. These 4 factors include, whether: 1) the theory or technique can or has been tested, 2) whether the theory or technique has been subject to peer review, 3) the rate of error and the standards controlling the technique’s operation, and 4) whether there is acceptance in the relevant scientific community. The Court also described trial judges as gatekeepers, who must consider that “the criteria for reception are relevance, reliability and necessity measured against the counterweights

---

61 Ibid.
62 Ibid., at para 1
63 The device involves attaching a gauge to the penis to measure a physiological response. See, at para 11
64 Ibid., at para 15
67 *J. (J.)*, supra note 60, at para 33
of consumption of time, prejudice and confusion." Balancing all of these factors, the Court rejected the expert’s evidence for a variety of reasons. First, because the expert was not forthcoming in their testimony or willing to provide the data that underlay their conclusions. Second, because the penile plethysmograph was not generally accepted as a reliable forensic tool to determine whether someone fit a profile to commit a crime. Third, the Court found that the expert’s attempt to explain what sort of distinct group would commit these crimes was overly vague, and that the test had a high error rate.

While this decision was helpful in saying what type of reliability analysis should be used in cases where the expert is testifying about a novel science, the Court provided little guidance about what novel science is. Here, the novel science constituted using an accepted diagnostic tool in a new way, so by implication, proposed expertise that attempts to do the same would likely fall under the novel science banner.

2.4.2 R. v. D. (D.):

In D. (D.), the Supreme Court further explained the necessity stage of the Mohan test, by differentiating between types of evidence that should be provided by an expert witness, and those that should be given through a jury instruction. There, the Crown called a child psychologist to testify that just because a child complainant waited two years before reporting her abuse, this did not indicate an inference of false reporting. In overturning the trial judge’s decision to allow this expertise, Justice Major (for the majority) quoted an article from Paciocco who noted that expert evidence is only necessary “when lay persons are apt to come to a wrong conclusion without expert assistance, or where access to important information will be lost unless we borrow from the learning of experts.” In this case, where the expert’s evidence was that, “In diagnosing cases of

---

68 Ibid., at para 47
69 Ibid., at para 50
70 Ibid., at para 35
71 Ibid., at para 55
72 Ibid., at paras 51-55
74 Ibid., at para 44
75 Ibid., at para 57
child sexual abuse, the timing of the disclosure, standing alone, signifies nothing,” Justice Major found that a jury instruction could just as easily convey the expert’s testimony, and thus, that an expert was unnecessary. The problem with this conclusion is that it is unclear how any judge is supposed to know when an expert’s proposed testimony can be dealt with through a jury instruction, aside from a bare assertion that it can be. What one judge may see as necessary evidence requiring an expert witness, could be merely helpful for another, and only require a jury instruction. Similarly, it is also unclear why necessity outweighed other considerations in the cost-benefit analysis, such as the relevance of the evidence.

In contrast, Chief Justice McLachlin (as she then was) in dissent believed that it might not be within the knowledge of the jury as to why a child complainant might delay their reporting of abuse, and would have allowed the expert’s evidence. Later, this thesis will revisit the many problems with using necessity as a ground for excluding expert testimony in these kinds of cases.

2.4.3 R. v. Trochym:

In perhaps the Supreme Court’s high water mark in assigning prominence to reliability concerns, Trochym dealt with whether post-hypnosis evidence, which had generally been accepted in the past, should now be rejected in light of new evidence. Significantly, hypnosis evidence was not a novel science, yet it was still treated with a high level of scrutiny. Writing for the majority, Justice Deschamps started off her decision by noting that there had recently been many public inquires about wrongful convictions, many of which featured dubious expert witnesses. Seemingly influenced by that, Justice Deschamps ruled that a party wishing to rely on hypnosis evidence must show that it is reliable. Here, the literature surrounding hypnosis evidence suggested that it had a high rate of error, could create distorted memories, and overall could not be shown to be accurate. As a result of this determination, the Majority ruled that hypnosis evidence was no

---

76 Ibid., at para 59
77 Ibid., at para 24
78 R. v. Trochym, 2007 SCC 6 [Trochym]
79 See dissent, at paras 132-133
80 Ibid., at para 1
81 Ibid., at para. 33
82 Ibid., at paras 38-54
longer sufficiently reliable to use in Court. As an addendum to J. (J.), Trochym suggests that once accepted expert subjects can be treated like novel science if there are changing circumstances to suggest the evidence is no longer reliable.

2.5 The Abbey Test And A New Framework:

2.5.1 R. v. Abbey:

In Abbey, the Court of Appeal for Ontario re-conceptualized how the Mohan test should be structured, and also dealt with how non-scientific experts should be scrutinized. At a first degree murder trial, the Crown wanted to call a sociologist to testify about the meaning of a teardrop tattoo on the accused’s face. In short, the expert was set to testify that if a gang member got a teardrop tattoo, it meant that they had either: killed a rival gang member, lost a close family member, or had spent time in jail. Since the shooting, the accused had not lost a family member or spent time in jail. At trial, the trial judge determined that the evidence was inadmissible because Dr. Totten’s proposed testimony did not meet the standard of reliability for novel science as set out in Daubert and J. (J.). Among other things, the trial judge was concerned that Dr. Totten could not provide an error rate for his conclusions, worried that Totten had never actually interviewed anyone from the accused’s gang, and skeptical that Totten was able to verify the truth of the answers given to him by gang members.

Before deciding that Dr. Totten’s testimony should have been allowed, at the Court of Appeal for Ontario, Justice Doherty attempted to provide clarity to Mohan by creating a two-stage test for the admissibility of expert evidence. At stage one, he proposed that there be four preconditions to admissibility: whether the proposed opinion relates to subject matter that is properly the subject of expert opinion evidence, whether the witness is qualified to give the opinion, whether the opinion runs afoul of another exclusionary rule, and whether the evidence is logically relevant. Unlike Mohan, the relevance at this stage does not engage in a cost-benefit analysis of whether the

---

83 Ibid., at paras 65-66
84 R. v. Abbey, 2009 ONCA 624 [Abbey I]
85 Ibid., at para 34
86 Ibid., at paras 55-59
87 Ibid., at paras 55-56
evidence should be admitted. At the second stage then, the trial judge performs their gate-keeper function, and looks to see if the evidence is legally relevant. In other words, does the probative value of the evidence outweigh its prejudicial effects? This cost-benefit analysis also takes into account how reliable the proposed evidence is, and if it is necessary for the trier of fact to hear it.

In deciding that Dr. Totten’s evidence should have been admitted, Justice Doherty found that the trial judge erred in applying a test of scientific validity to a non-scientific area of expertise. Instead, the trial judge should have focused on whether the expert had "specialized knowledge gathered through experience and specialized training in the relevant field." Then, the trial judge could have looked to certain factors to prove reliability, like whether the expert sought to testify in a recognized field, what quality controls they have on their work, how their data is recorded and stored, to what extent their data methods are recognized as being appropriate by other experts within that field, etc. Having regard to these reformulated factors, Justice Doherty found that Dr. Totten was testifying in a recognized field, was using appropriate methods in that field, and therefore, that his testimony was sufficiently reliable.

While cases like J. (J.) and Trochym talked about reliability in the context of novel science and evolving views towards a previously recognized science, Justice Doherty’s opinion provides a comprehensive account of how to assess reliability in the non-scientific context. As will be discussed later, while many subsequent courts have embraced the holding in Abbey that non-scientific evidence should be treated differently, not all of these cases have asked if the proposed evidence meet the indicia of reliability proposed by Justice Doherty. The decision is also important in that Justice Doherty emphasized that trial judges have a strong role to play in monitoring the scope of an expert’s testimony, often to avoid touching on the ultimate issue.

---

88 Ibid., at para 76
89 Ibid., at paras 84-96
90 Ibid., at para 109
91 Ibid., at para 119
92 Ibid., at para 126
93 Ibid., at paras 102-103
2.5.2  *R. v. Sekhon*:

In *Sekhon*,\(^{94}\) the Supreme Court took up the finding from *Abbey* that the trial judge must play an active role in ensuring that the expert witness does not testify outside the scope of their expertise. There, the accused was tried with trafficking cocaine, after police found fifty kilograms of drugs hidden within the accused’s truck. The accused’s defense was that he was an unwitting drug mule and did not know that the cocaine was there.\(^{95}\) The Crown called a police officer to testify about the customs and practices of the drug trade, and in the course of his testimony the police officer stated that he had never seen a case where someone with drugs in their car had been a blind courier.\(^{96}\) Defense counsel did not object to the statement being admitted, but merely stated that they trusted the trial judge to give this evidence the proper weight.\(^{97}\) In overturning the defendant’s conviction, the Court said that notwithstanding the fact that defense counsel did not object to the officer’s statement, the trial judge had the responsibility to prevent the expert witness from testifying outside the scope of their expertise.\(^{98}\) Namely, that the actions of past couriers had bearings on the accused’s guilt or innocence. Therefore, the Court excluded the evidence on the basis that it was not reliable or necessary, and was extremely prejudicial by implying the defendant must be guilty.\(^{99}\)

2.6 *White Burgess And Its Aftermath:*

In *White Burgess v Abbot*\(^{100}\) the Supreme Court adopted much of the two stage approach suggested by Justice Doherty in *Abbey*, and gave the clearest statement since *Mohan* of what factors to take into account when deciding the admissibility of an expert witness. The Court also made it mandatory for trial judges to consider if the expert witness was biased. Writing for the Court, Justice Cromwell said that at the first stage of the admissibility inquiry, the proposed evidence must meet four threshold conditions (plus a fifth one if the technique is novel). First, it needs to be

---

\(^{94}\) *R. v. Sekhon*, 2014 SCC 15 [Sekhon]

\(^{95}\) Ibid., at paras 22-24

\(^{96}\) Ibid., at para 20

\(^{97}\) Anderson, *supra* note 9, at pg 93

\(^{98}\) *Sekhon*, *supra* note 94, at paras 46-48

\(^{99}\) Ibid., at paras 49-51

\(^{100}\) *White Burgess Langille Inman v. Abbott and Haliburton Co.*, 2015 SCC 23 [White Burgess]
logically relevant to a fact in issue. Second, it must be necessary to assist the trier of fact. Third, it cannot run afoul of another exclusionary rule. Fourth, the expert must be properly qualified, and show themselves able to be impartial, independent, and unbiased. Fifth, if the proposed evidence is novel, it must be shown to be reliable. If the evidence meets these threshold requirements, then the trial judge turns to the second stage. At the second stage of the inquiry, the trial judge plays a gatekeeper role and balances the costs and benefits of admitting the evidence, having regard to its relevance, necessity, reliability, and there being absence of bias. In an improvement on Mohan, the Court more explicitly mentioned reliability as a consideration, both when the evidence is novel and at the cost-benefit stage, but still retained it as one admissibility factor among many. As the Court never referred to the type of reliability analysis conducted in Abbey, presumably Abbey remains good law in Ontario in those circumstances.

2.6.1 R. v. Awer:

Decided shortly after White Burgess, Awer revealed the often artificial nature surrounding whether evidence goes to ultimate reliability or not, and that the Supreme Court can be complacent in letting the adversarial process sniff out reliability concerns. Arising from Alberta, the disputed expert testimony in Awer related to whether or not the Crown’s DNA expert could properly testify that the existence of the complainant’s DNA on the accused’s penis came from a wet source (a person), and not through another sort of secondary transfer (a dry source like clothing). The expert’s testimony was based on trade journal articles, and his own experience of having never seen so much DNA transfer from a dry source. The majority for the Alberta Court of Appeal found that the trial judge did not err in relying on the expert’s experience, to establish expertise, in coming to their conclusion about the source of the DNA transfer. In contrast, the dissent concluded that the expert’s conclusions were anecdotal, and not shown to be reliable.

101 Ibid., at para 23
102 Ibid., at para 24
104 R. v. Awer, 2017 SCC 2 [Awer SCC]
106 Ibid.
107 Ibid., at paras 85-88
108 Ibid., at paras 128-131
At the Supreme Court, the Court ruled that the matter should be remanded for a new trial on the basis that the trial judge subjected different standards of scrutiny to the Crown and Defense’s expert witnesses, but they were agnostic about whether the DNA expert’s testimony should be seen as reliable or admissible. Instead, they were content to let the evidence be put forward at another voir dire, and have the defense cross-examine the expert, presumably under the assumption that this would adequately show if the evidence was reliable. Regarding the issue of ultimate reliability, the Court stated that this evidence differed qualitatively from the impugned testimony in Sekhon by not going directly to the ultimate issue. While perhaps technically correct, it seems that accepting the testimony that the DNA came from a wet source could only lead to a conclusion of the accused’s guilt.

2.6.2 R. v. Bingley:

Similar to Awer, the Supreme Court in Bingley again did not pay much heed to reliability questions, this time in context of a statutory scheme dealing with Drug Recognition Experts (DRE’s). In Bingley, the dispute essentially centered on whether s. 254(3.1) of the Criminal Code negated the need for DRE’s to go through a Mohan voir dire. Both Crown and Defense agreed that the proposed evidence was relevant, necessary, and did not run afoul of another exclusionary rule, but it was contested whether a DRE was a properly qualified expert. At trial, the trial judge focused on the novelty of the techniques practiced by the DRE’s, and therefore, required that these experts had to explain the underlying reliability of the science behind the technique in a voir dire. At the Supreme Court however, the Court found that “the reliability of the 12-step evaluation comes from the statutory framework itself. Parliament has determined that the 12-step evaluation performed by a trained DRE constitutes evidence of drug impairment.” That being the case, the DRE did not need to explain the underlying science behind the techniques legislated in the Criminal Code, thus negating the need for any Mohan voir dire in the future because the

---

109 Awer SCC, supra note 104, at paras 2-6
110 Ibid.
111 Ibid.
113 Ibid., at paras 28
114 Ibid., at para 23
115 Ibid., at para 24
expert was now interpreted to be a qualified expert. While it is understandable to some extent why the Court would defer to Parliament’s legislative scheme, it is also troubling that by virtue of the fact that Parliament deemed this 12-step evaluation to be reliable, that the Court took a hands-off approach. As much as one would hope, the fact that Parliament legislates something does not sufficiently establish whether or not a particular technique is reliable.

2.6.3  *R. v. Abbey (The Sequel):*

After the first *Abbey* appeal, the Crown retried Abbey for murder, and this time, the expert’s evidence about teardrop tattoos was allowed into evidence. Abbey was convicted of murder, and the defense on appeal brought a fresh evidence application that they said called into question the reliability underlying Dr. Totten’s opinion. Citing the new admissibility framework from *White Burgess*, and applying some of the questions identified in *Abbey I*, the Court of Appeal in *Abbey II* now found many troubling aspects to Dr. Totten’s testimony. First, it appeared that Dr. Totten inflated the number of gang members that he had talked to, calling into question the provenance of the sample size from which he drew conclusions about the meaning of tattoos. It also appeared that he talked to far fewer people who had specifically committed a homicide, and could explain what a tattoo meant in that context. Combined, the fact that there was now a much smaller sample size to draw generalizations from made his conclusions less reliable. Second, in digging into his data the Court discovered (and found it implausible) that Totten had recorded that everyone who had a teardrop tattoo did it because they had murdered another gang member, and not because they had a death in the family or had served a prison stint. Third, and like in *J. (J.)*, the Court found it disqualifying that Dr. Totten did not keep much of the underlying data supporting his claims, leaving the Court unable to determine if his data was correct and reliable. As more of a general point of interest, the Court also noted the fact that in a case after *Abbey I*, the defense attempted to have Dr. Totten testify as an expert on gangs, and the Crown there harshly questioned his credentials. The Court of Appeal found it unsavory that the Crown would say that Dr. Totten’s

---

116 *R. v. Abbey*, 2017 ONCA 640 [*Abbey II*]
117 Ibid., at paras 72-82
118 Ibid., at paras 83-85
119 Ibid., at paras 98-100
120 Ibid., at paras 95-97
evidence for the defense in one trial was unreliable, but then rely on his opinion to secure a conviction when he was their witness.\textsuperscript{121}

2.6.4 Conclusion:

As of now Mohan and White Burgess remain the leading cases in setting out the framework for the admission of expert evidence, with the other cases mentioned speaking to different aspects of the admissibility test. In the next section, this thesis will evaluate the body of case law talked about in this section, and discuss the strengths and weaknesses of the jurisprudence in this area.

\textsuperscript{121} Ibid., at paras 140-141
3 Part II: Evaluating the Approach Towards Expert Evidence:

3.1 The Good:

3.1.1 1-It is Tailored Towards Reaching Criminal Justice Goals:

After *White Burgess*, a trial judge now needs to conduct a multi-faceted inquiry when admitting an expert witness. This inquiry focuses on some of the important questions that a justice system focused on trial efficiency, preventing wrongful convictions, and seeking the truth, would want to answer. These considerations are looked at in both stages of the newly modified two-step *Mohan* framework.\(^{122}\) Regarding trial efficiency, the inquiry asks whether the proposed expert evidence is necessary or whether the information can be introduced in a more efficient matter. In preventing wrongful convictions, the inquiry is now more focused on reliability and if the expert witness is biased. Last, the focus on whether the evidence is logically and legally relevant to the case at hand, gets at the trial judge’s goal to seek the truth. As will be argued later, this thesis’ new admissibility framework will better rebalance these goals.

3.1.2 2-It Has Improved Over Time:

From the Supreme Court case in *Abbey* in 1982 to the Court of Appeal case of *Abbey II* in 2017, the jurisprudence towards expert witnesses has become more fleshed out over time. First, the role of the trial judge as a gatekeeper has become more emphasized, although not necessarily explained very well. In *Mohan*, the Supreme Court does not refer to the trial judge as having a gatekeeper role, and it is unclear from the case how active a role a trial judge is supposed to play aside from applying a set of criteria towards the evidence.\(^{123}\) In *J. (J.)*, the Court referred to the trial judge as being a gatekeeper, and said that they had an active role in flushing out novel science that was unreliable.\(^{124}\) The law further evolved in *Abbey I* (and confirmed in *White Burgess*) to include a

\(^{122}\) *White Burgess*, *supra* note 100, at paras 23-24
\(^{123}\) *Mohan*, *supra* note 40, at paras 16-21
\(^{124}\) *J. (J.)*, *supra* note 60, at para 35
second stage to the admissions inquiry, where the trial judge in their role as gatekeeper conducts a cost-benefit analysis of whether the evidence is ultimately more helpful than harmful.\textsuperscript{125}

Second, the scope of an expert’s testimony has become more tightly restricted. In \textit{Marquard}, a pediatric expert’s comments about the origins of the victim’s burns was accepted by the Court,\textsuperscript{126} despite not having expertise in that area, on the basis that the expert was a Doctor. Post-\textit{Abbey I}, it is likely a trier of fact would have restricted the expert’s testimony to what they were actually an expert in, and excised the additional comments about burn origins. Similarly, despite not being trained in identifying shaken baby syndrome, Dr. Charles Smith caused multiple wrongful convictions by relying on his general medical knowledge to testify about this specialized area.\textsuperscript{127}

Indeed, in \textit{Abbey I} the Court of Appeal stated that judge should play an active role in paring down the scope of an expert’s evidence,\textsuperscript{128} stating that, “Admissibility is not an all or nothing proposition.... The trial judge may admit part of the proffered testimony, modify the nature or scope of the proposed opinion, or edit the language used to frame that opinion.”\textsuperscript{129} Thus, the Court of Appeal in \textit{Abbey I} restricted Dr. Totten’s testimony to testifying what teardrop tattoos could mean in general, and not what they meant on the accused’s face in particular.\textsuperscript{130} Similarly, while finding no fault in \textit{Sekhon} with the police officer’s general testimony about the drug trade, the Court found that the trial judge erred in admitting the officer’s opinion that he had never seen an unwitting drug carrier before.\textsuperscript{131}

Third, there has been a greater focus on the reliability of evidence, although this thesis argues the jurisprudence does not go far enough. In \textit{Mohan}, the Court briefly adverted to the idea that the reliability of an expert’s evidence is a consideration during the relevance stage, and also mentions that novel science should be scrutinized more than accepted fields.\textsuperscript{132} However, this focus on

\begin{itemize}
\item\textsuperscript{125} \textit{Abbey I, supra} note 84, at para 76
\item\textsuperscript{126} Kent Roach & Gary A. Edmond. “Contextual Approach to the Admissibility of the State’s Forensic Science and Medical Evidence” (2011) 61 U Toronto LJ 343, at pg 382 [Roach & Edmond]
\item\textsuperscript{128} Lisa Dufrainmont. “New Challenges for the Gatekeeper: The Evolving Law on Expert Evidence in Criminal Cases” (2012) 58 CLQ 531, at pgs 547-549 [Dufrainmont 2012]
\item\textsuperscript{129} \textit{Abbey I, supra} note 84, at para 63
\item\textsuperscript{130} Ibid., at paras 67-70
\item\textsuperscript{131} \textit{Sekhon, supra} note 94, at para 49
\item\textsuperscript{132} \textit{Mohan, supra} note 40, at para 32
\end{itemize}
reliability only became more apparent in *J. (J.)*, where the Court looked to standards from *Daubert* to see if the novel use of a penile plethysmograph could be used for forensic purposes.\(^{133}\) Similarly, in *Trochym*, the Court emphasized that the reliability of a certain technique could change over time, and that hypnosis, a previously accepted form of expert evidence, was now no longer reliable.\(^{134}\) As stated earlier, *Abbey I* then explained how non-scientific evidence should be treated, and provided a list of factors that may be useful in determining its reliability.\(^{135}\) Finally, *White Burgess* stated that novel science must be shown to be reliable as a threshold consideration, and that the reliability of the underlying evidence must be considered in the cost-benefit balancing analysis as well.\(^{136}\)

Fourth, there is now an explicit requirement that the expert should be independent, impartial, and unbiased. In *Mohan*, the Court does not mention bias as being relevant to whether an expert should be allowed to testify. In *Abbey I* however, Justice Doherty stated that when looking at the reliability of the expert’s evidence, their ability to be impartial and objective should be considered.\(^{137}\) Finally, in *White Burgess*, the Court clarified that the lack of independence and impartiality in a witness can be grounds for excluding the witness at the threshold inquiry, and can also be considered in the overall cost-benefit analysis.\(^{138}\)

Fifth, the Courts have become somewhat more cognizant of the role that expert witnesses have played in wrongful conviction cases. While not dealing with expert evidence, the Supreme Court in the *United States of America v. Burns*\(^ {139}\) determined that a Canadian citizen could not be extradited to the United States if they faced the death penalty, because of the concern that someone would be wrongfully executed.\(^{140}\) Thus, in *Trochym*, the first case dealing with expert evidence after *Burns*, the case linked problems with hypnosis to the wrongful conviction of Thomas

\(^{133}\) Dufraimont, *supra* note 128, at pgs 541-542  
\(^{134}\) *Trochym*, *supra* note 78, at para 67  
\(^{135}\) *Abbey I*, *supra* note 84, at para 119  
\(^{136}\) *White Burgess*, *supra* note 100, at paras 23-24  
\(^{137}\) *Abbey I*, *supra* note 84, at para 87  
\(^{138}\) *White Burgess*, *supra* note 100, at para 34  
\(^{139}\) *United States of America v. Burns*, 2001 SCC 7 [Burns]  
\(^{140}\) Roach & Edmond, *supra* note 126, at pg 384
Sophonow, who was convicted in part because of the hazards of human memory. More explicitly, they also said the reliability analysis is needed to prevent wrongful convictions.

3.2 The Bad:

Despite the slow improvements made to the test for admitting expert evidence, the inquiry is still bedeviled by: its failure to emphasize reliability, an increasing focus on how to characterize evidence, and a lack of focus and structure. After discussing these problems, this thesis will propose a solution to address many of these issues, by offering a new framework for admissibility in Part III.

3.2.1 1-Reliability:

While recognizing the growing judicial awareness that the reliability of expert evidence is important, this thesis and many scholars do not think that this inquiry goes far enough. For one thing, the reliability of the evidence remains an individual factor amongst many others, often leaving it as an item to check off in an admissibility analysis at the expense of a fulsome inquiry. Indeed, if not in a novel science context, the case law often treats reliability as an unimportant factor, despite Justice Doherty’s directions in *Abbey I*. Given that, and as will be discussed more below, some cases do not make any effort into scrutinizing the reliability of proposed evidence at all. Relatedly, even when judges delve into a reliability analysis, the inquiry often devolves into looking at proxy factors such as the expert’s academic qualifications, instead of figuring out whether they can actually provide reliable testimony. Indeed, Roach and Edmond believe that the Courts should be more concerned with validation studies of the proposed evidence. This concern that a reliability analysis is not pursued in every case, and the question of how a reliability analysis should be conducted, will be answered by this thesis’ new framework for admitting this evidence.

---

141 *Trochym, supra* note 78, at para 46
142 Ibid., at para 60
144 Roach & Edmond, *supra* note 126, at pgs 400-404
3.2.2 2-The Characterization Problem:

Another problem endemic in the case law is that often the scrutiny given towards reliability hinges on the trial judge making a sometimes arbitrary decision about how to characterize evidence. For example, Roach and Edmond point out that there is a common trend to sidestep more probing inquires into reliability by characterizing the expert’s opinion as non-scientific. As discussed in Abbey I, evidence being characterized as scientific or non-scientific will have an impact on how the proposed evidence is scrutinized. Whereas in Abbey I, reliability was still important in addressing both types of evidence, just the nature of the questions changed, now this characterization exercise has emerged as a way to skirt reliability at all. This has caused some scholars to opine that Abbey has opened “the door to a new type of controversy over the proper characterization of a particular technique.” Indeed, Roach and Edmond fear that there is now a growing incentive for expert witnesses to attempt to bypass more rigorous scientific standards by describing their work as non-scientific and being based on their special expertise. In focusing on this classification, this obscures the fact that trial judges should be more concerned with the actual reliability of the evidence.

Similarly, the characterization of whether the evidence is novel or not, or whether the opinion goes directly to ultimate reliability or not, also has a disproportionate impact on how the reliability inquiry is conducted. For example, in Trochym, the majority treated the evidence as novel, and subjected it to a strict reliability analysis, while Justice Bastarache in dissent did not see the evidence as novel and would have admitted it on the basis of hypnosis evidence’s past admission. As will be discussed later when talking about Awer, the characterization of the evidence as not going to the ultimate issue, despite its large role in determining the accused’s guilt, allowed the evidence to be treated with lesser scrutiny.

146 Roach & Edmond, supra note 126, at pg 394
147 Ibid., at pgs 399-400; Chin 2018, supra note 103, at pgs 422-423
148 Trochym, supra note 78, at para 131
149 See Awer, supra notes 104-105
Another characterization problem can take place in identification cases. There, police officers often provide opinions, but instead of being considered opinion experts subject to a *Mohan voir dire*, they are called “witnesses with special knowledge purporting to make an identification.”\(^{150}\) For example, in *R. v. McGean*\(^{151}\) the supervisor of a wire tap room, named Officer Knisley, was called to provide an opinion of whether the voice captured on the wire tap was the same voice as the accused.\(^{152}\) Officer Knisley’s testimony was considered to be the opinion of a lay person, and was accepted despite him not having shown an ability to accurately match voices, vitiating the need for the Crown to call an expert to perform this voice identification analysis. This type of ‘lay evidence’ offered by the police has also been accepted in cases where police officers testified that they observed evidence that attempts had been made to remove blood stains in a home.\(^{153}\) Taken together, these attempts at avoiding the more rigorous reliability associated with scientific evidence misinterprets *Abbey I*, and turns the focus towards who is giving an opinion instead of how reliable it is.

### 3.2.3 3-Will *White Burgess* Really Eliminate Bias?:

While on paper *White Burgess* directed trial judges to consider whether an expert is biased, it is unclear how much of an impact this will have preventing the dangers that stem from a biased witness, namely, obscuring the truth and causing wrongful convictions. First, the decision does not take into account more veiled forms of bias, like contextual bias. Contextual bias takes place when the expert is exposed to “information, procedures and other influences that may influence analysis or interpretation in undesirable ways.”\(^{154}\) For example, in a case called *Aitken* (which will be talked about later), the expert’s identification was made in the context where “he was comparing the gait from the incident video with that of a single individual whom he knew the police believed to be

---

\(^{150}\) *R. v. Williams*, [1995] OJ No 1012 (CA), at para 17 [*Williams*]

\(^{151}\) *R. v. McGean*, 2016 ONSC 5572 [*McGean*]

\(^{152}\) Ibid., at paras 41, 56


Aside from the question if this sort of comparison can produce a reliable result, it is not difficult to see why this could bias an expert to reach a certain conclusion, even if they otherwise had no partisan interest in the case.

Another problem that this thesis foresees with the new focus on bias is that defense witnesses may be more likely to be harmed than Crown witnesses. Crown witnesses, like Dr. Charles Smith, often have impressive resumes and the imprimatur of a reputable institution (like the Centre of Forensic Sciences). On paper, there is little reason to doubt they are unbiased, even though prior inquiries into wrongful convictions have shown that “there may be a tendency [among government employed forensic scientists] to consider themselves aligned with the Crown ... and that their function is to support the police theory.”\textsuperscript{156} In contrast, defense experts may lack a form of institutional prestige, and advance more novel theories that seem improperly tied to the defense’s theory of the case. If the effect of \textit{White Burgess} is to stymie an accused persons’ ability to raise a defense, then the case’s focus on bias might be more harmful than helpful.

3.2.4 4-The Gatekeeper Needs More Help:

The two \textit{Abbey} cases present a sobering reminder that even if a Court engages in the proper questions about an expert’s reliability, it is always possible that they will still allow unreliable evidence. While this may be an intractable problem, as it is not realistic that an admissibly test will always reach the right answer, this pursuit of truth is not helped by a current admissibility test that is focused on too many disparate things. As discussed in \textit{Abbey I}, the Court of Appeal overturned the trial judge’s conclusion that Dr. Totten’s testimony was inadmissible,\textsuperscript{157} it was admitted in \textit{Abbey II}, and Abbey was convicted.\textsuperscript{158} On appeal Abbey’s lawyers then brought fresh evidence that really dug into Dr. Totten’s data and methods, which had the effect of showing that Dr. Totten’s opinion could not be relied on.\textsuperscript{159} While not saying that using a different admissibility test

\begin{footnotesize}
155 Ibid., at pg 365
157 \textit{Abbey I, supra} note 84, at para 4
158 \textit{Abbey II, supra} note 114, at paras 1-10
159 Ibid., at paras 117-125
\end{footnotesize}
framework would have discovered this evidence sooner, it is useful to ask whether asking counsel to focus on a whole number of factors (under the current test) prevented them from really honing in on Totten’s reliability. In contrast, focusing the admissibility test on reliability may make the gatekeeper role easier, by possibly having the effect of incentivizing counsel to put more effort into dissecting an expert’s methodologies and conclusions in front of the trial judge.

3.2.5 The Test Is Too Flexible:

Another problem with the admissibility test is that it speaks to too many concerns without providing a hierarchy of values. The threshold questions alternatively speak to concerns about court efficiency, how the proposed evidence fits within the larger rules of evidence, if the expert is qualified, and sometimes reliability concerns, without stating which factors are more important. While there is some benefit to addressing several different concerns at the same time, doing so also leaves it difficult to ascertain the relative importance of each one, leading to court decisions that seem to oscillate between different predominant factors. For example, in D. (D.) and other cases dealing with behavioral evidence, necessity is often the dominant factor, whereas in Trochym the court was focused on reliability even though the subject of the expert’s expertise was not technically novel. In contrast, the new admissibility test is focused on creating a framework that is more targeted to the one factor it finds most important.

Similarly, the different threshold questions often seem to conflict with one another, and there are no tools offered to address these conflicts. For example, in D. (D.), the expert’s proposed testimony about delayed complaint was highly relevant to the case at hand, but this apparently conflicted with the necessity part of the Mohan inquiry. Besides the fact that the necessity inquiry is much more subjective than every other threshold question, and seemingly arbitrary, it is unclear from D. (D.) why necessity won the day in that case. While it would not be desirable for judges to have no flexibility in picking which factors are more important in each individual case, it is also undesirable for otherwise relevant evidence to be turned into a jury instruction based on the highly subjective determination made under the necessity test. Indeed, Lederman hypothesizes that especially in the context of psychological evidence, trial judges say the evidence is not necessary because that is
easier than deciding how valid these fields are.\textsuperscript{160} Thus, while believing that judges still should have a lot of power in their gatekeeping role, this thesis claims that this power can be better served through a framework with more structure.

3.2.6 Conclusion:

While this thesis recognizes that the current admissibility does a somewhat adequate job of representing criminal justice interests, and has improved over time, it is still beset by a host of issues. As mentioned above, the admissibility inquiry is: not focused enough on reliability, too often focuses on characterizing evidence, and too unstructured. In Part III, the proposed admissibility framework will attempt to redress these concerns.

Part III: A New Test:

Just as Justice Doherty reformulated the Mohan test in Abbey (since slightly modified in White Burgess), this thesis proposes that the test for admitting expert evidence should again be changed, this time adding reliability as its own separate part in a three-part framework. By making these changes, the aim is to make the admissibility framework more logical, cohere better with other aspects of evidence law, prevent wrongful convictions, give trial judges more structure in their gatekeeping role, address troubling jurisprudential trends, and properly shift some of the burden of proving reliability on the experts themselves. Having explained the new test, and the improvements it would offer, Part IV will then apply this framework to two cases much discussed in the academic literature, and see if the admissibility analysis would have been different.

4.1 The Framework:

As envisioned, the only threshold question at the first stage of this new admissibility analysis would be focused on whether the evidence in question is reliable. This first stage would have several component parts, borrowing from existing case law, and ask: 1) Is this field capable of being reliable, 2) Is this person qualified to give an opinion in this field, and 3) Is their opinion based on reliable methods. In addressing both the field of expertise and the expert’s qualifications, these questions would be posed to generally ensure that a qualified expert is using proper tools in an area amenable to expertise, and that can produce reliable results. If the proposed expert cannot satisfy these 3 questions, the evidence should be excluded, and the other factors from Mohan and White Burgess need not be addressed. However, if the evidence is reliable, then a trial judge can move on to the second stage. At the second stage of this new test, the trial judge would then look at factors such as whether the expert is acting independently, impartially, and in an unbiased manner and if the evidence: runs afoul of other exclusionary rules, is logically relevant, and necessary. At the last stage, the trial judge must still then engage in a cost-benefit analysis of how

---

161 For example, questions 1 and 2 in part resemble questions asked by Abbey I in part 1 of its proposed framework. Question 3 then reflects later paragraphs of Abbey I that talk about the importance of reliability, whether using Daubert or asking other questions.
reliable the evidence is, whether the expert is unbiased, and if the evidence is legally relevant, necessary, or would have a significant impact on the accused’s guilt or innocence.

Recognizing the difficulty that judges have in this admissibility task, several scholars have suggested ways to make this admissibility process easier, which this framework would also encourage. For example, Lederman suggests that there should be a greater onus before trial for Counsel to provide disclosure of the general testimony their expert is going to give, including “the methodology upon which the expert anchored his or her conclusions.”162 Similarly, Chin suggests that the expert should provide full transparency to the Court, before trial, on how they reached their conclusions.163 Combined, these two suggestions would fit with the framework’s goal of having experts show their work, and it would also focus the mind of all the criminal justice participants to the task of assessing reliability.

4.2 Guiding Principles:

4.2.1 Type Of Evidence:

In this framework, the standard of reliability needed to pass the first stage will change depending on the type of evidence being presented, but in general, there is an expectation that the expert should be able to validate their work through the transparent disclosure of their work product. The distinction made in Abbey between scientific and other types of evidence still remains apt, and different types of questions should be asked depending on the type of evidence tendered.

For example, if a DNA expert is proffered to provide an opinion that an accused’s DNA matches evidence found at a crime scene, having regard to the Daubert factors would continue to make sense. That is to say, the court would enquire whether the expert used techniques that have been tested, whether this technique was the subject of peer review, what the error rate of this technique is, and whether the DNA community generally accepts this technique. Because this subject area of expertise is amenable to a clearer right or wrong determination, the aim of these questions would be to provide the trial judge with a clear mandate to admit or dispose of the evidence.

162 Lederman, supra note 160, at pg 239
163 Chin 2018, supra note 103, at pg 463
On the opposite end of the spectrum, the questions that a trial judge would need to ask in a case like *Shafia*,\(^{164}\) where the expert testified about Afghan culture and honor killings,\(^{165}\) would be different. In that sort of case, where the expert’s opinion is essentially informational, more attention would need to be paid about the expert’s academic credentials, what methodology they used to reach their conclusions, and if there are competing theories in their field, etc. The point of these questions would be to allow the trial judge to determine if the expert’s testimony would be generally reliable, or whether they are representing a fringe opinion.

In cases like *Abbey I*, where the expert is making conclusions bolstered by their specialized knowledge gathered through experience and training, the standard should be that the expert must have some validating data to turn to when providing their opinion.\(^{166}\) This requirement is needed because although Totten’s testimony in *Abbey I* might not have been based on pure science, it is also not a field inimical to safeguards in the form of data collection. While this thesis does not suggest how robust a sample size is needed to satisfy the validating data requirement in every case, obviously factors such as a greater sample size, combined with complying with proper methods, would lean towards admissibility. As well, the questions that Justice Doherty poses in *Abbey I*, like how their data was collected and what quality controls they used, would also be helpful in assessing the strength of this evidence. Thus, although it turned out in *Abbey II* that the data which Dr. Totten relied on may have been fabricated, it was reasonable for the Court of Appeal in *Abbey I* to believe that his opinion was reliable enough for the purpose of threshold reliability. In contrast, had Dr. Totten offered to provide his opinion on the meaning of teardrop tattoos based on his experience, but in the absence of any research, his opinion would not have passed part 1 of the new framework. While not a panacea, this expectation that an expert basing conclusions on their experience must have some data to underlie their opinion, creates a simple barrier to prevent the most dangerous types of evidence from being admitted.

Another type of evidence that would be treated with special concern under the new framework would be in areas where wrongful convictions have happened in the past. For example, forensic

---

\(^{164}\) *R. v. Shafia*, 2016 ONCA 812

\(^{165}\) Ibid., at paras 203-205

\(^{166}\) Referred to as a ‘testedness’ requirement in the Goudge Report, this standard would ask experts in fields such as handwriting analysis and forensic odontology to prove their work. See Goudge Report, *infra* note 179, at pg 482
pathology would be treated with heightened attention, given the concerns the Goudge Inquiry revealed with this type of evidence. A case the framework would attempt to emulate would be *France*,\(^{167}\) where Justice Molloy harshly criticized the forensic pathologist’s testimony in that case. Indeed, having referred to the Goudge Inquiry, Justice Molloy concluded that the expert evinced bias, used improper language, was not prepared to testify, and spoke outside of his expertise.\(^{168}\) By using this heuristic, the framework is attempting to ensure that wrongful convictions in certain areas do not repeat themselves.

4.2.2 Role Of The Evidence:

Recognizing the sometimes artificial distinctions drawn between evidence that does or does not go to the ultimate issue, this new framework will be more focused about the role that an expert’s evidence can be expected to play in determining the accused’s guilt or innocence. In general, an expert’s testimony should be treated with more caution if it can reasonably be expected to play a large role in determining guilt or innocence, regardless if it actually touches on the ultimate issue or not. Going back to the prior examples, in a case where DNA evidence may be very consequential in proving the defendant’s guilt, it makes sense to hold the expert to a higher standard. Similarly, a high level of scrutiny would be attached to Dr. Totten’s testimony in *Abbey*, where although he was restricted to saying that a tattoo could have one of three meanings, it was not hard to draw the very damaging inference that Abbey’s tattoo signified killing a gang member. In contrast, in a case where an expert is providing mostly informational testimony, the role they play in the conviction might be more muted, and their evidence would be treated with more flexibility. For example, if the expert had been allowed to testify in *D. (D.)*, the goal of their testimony from the point of the view of the Crown would have been to rebut the defense inference that delayed disclosure from a child witness meant that an assault never happened. Had the expert been called, the jury accepted their testimony, and the defendant was found guilty, this testimony may have been helpful in reaching that conclusion, but it seems unlikely that it would be a major or significant cause.

---

\(^{167}\) *R. v. France*, 2017 ONSC 2040 [*France*]

\(^{168}\) Ibid, at paras 68-69
4.2.3 How Reliable The Evidence Is:

Another consideration under this framework, when it gets to the balancing stage, is asking how reliable the proposed evidence is. For example, while under the proposed framework the type of evidence offered by Dr. Totten could pass stage 1, the fact that it is not the most reliable type of evidence should be considered under stage 3. In contrast, in areas of social science where there is robust evidence attesting to its reliability, like eyewitness identification (which has been deemed unnecessary and inadmissible), this would be looked upon favourably in the cost-benefit analysis. Another consideration that may be useful at stage 3 would to ask about the circumstances of how the evidence was created. For example, was the expert’s data created independent of the matter in court, were they aware of who the accused was when they came to their opinion, etc. Another suggestion raised by Chin is to ask is this field capable of reasonably creating proficiency, by for example, providing immediate feedback to a practitioner of it.169 This consideration of how reliable the evidence is, which seems to be a more generally objective query, is seen by this essay as a better standard for admissibility at the cost-benefit stage, instead of the current case law which often looks to necessity.

4.2.4 Who Is Offering The Evidence:

Last, this new framework is also concerned with who is presenting the evidence. As pointed out by Roach and Edmond, two fundamental criminal justice values are the presumption of innocence, and the state’s obligation to prove the defendant guilty beyond a reasonable doubt.170 Given those factors, and the fact that many wrongful convictions have been caused by experts tied to the state, this thesis agrees with them that experts put forward by the Crown should be treated with greater scrutiny than defense experts.171 In practice, this could mean that a Crown witness would need to have a more robust body of validation data than a defense witness, or that theories propagated by Crown experts must have more support in the scientific literature. Put simply, there is greater danger in a Crown expert providing damning, but unreliable evidence, that leads to a finding of guilt, than in a defense expert attempting to present a more speculative theory on behalf of the

---

169 Chin 2018, supra note 103, at pg 474
170 Roach and Edmond, supra note 126, at pgs 348-349
171 Ibid.
accused’s innocence. This position would also seem to align with the Supreme Court’s commentary in *Seaboyer*\(^{172}\) that judges should be “extremely cautious in restricting the power of the accused to call evidence in his or her defence, a reluctance founded in the fundamental tenet of our judicial system that an innocent person must not be convicted.”\(^{173}\)

### 4.3 Benefits Of The New Test:

#### 4.3.1 Making The Test More Coherent:

While this new framework would still consider various other factors, the focus on reliability, which relates to truth-seeking and the desire to prevent wrongful convictions, would provide a clearer roadmap to what is important in this admissibility analysis. Thus, a benefit of this change would be that judges would still be expected to play an important gatekeeper function, but would have more structure to guide their admissibility decision. As stated above, this thesis believes that the current framework lends itself to too subjective an analysis, which instead of providing trial judges with appropriate flexibility, has instead given them no guidance at all. With this newfound clarity and structure, it is hoped that the jurisprudence surrounding this topic will become more coherent, easier to judge on its merits, and more amenable to appellate review.

Similarly, focusing the evidence in this way would eliminate the need for the trial judge to engage in a number of discretionary decisions that affects how they conduct their reliability inquiry. For example, because the reliability framework is present in every case, the trial judge does not need to as strongly focus on whether the evidence is novel, whether it goes to the ultimate issue, and what different steps in the *Mohan* framework should be relatively more important.

#### 4.3.2 Fitting In With Other Areas Of The Law:

A new focus on reliability would also align this test with other evidentiary tests, and by doing so, create more coherence in evidence law. For example, the principled exception to the hearsay rule

\(^{172}\) *R. v. Seaboyer*, [1991] 2 SCR 577 [*Seaboyer*]

\(^{173}\) Ibid., at para 48
only focuses on two factors, whether the evidence is necessary and reliable.\textsuperscript{174} Like opinion evidence, the focus on reliability is concerned with the threshold reliability of the statement, and not its ultimate reliability.\textsuperscript{175} While the dangers of relying on hearsay evidence differ from expert evidence, it seems odd to be greatly concerned with reliability in the hearsay context, while treating it as a secondary consideration in the opinion context. In the hearsay context, reliability is established when there are either procedural or substantive safeguards. Procedural safeguards include an out of court statement being given under oath, it being on video, the declarant being warned about lying, and/or the declarant facing contemporaneous questioning.\textsuperscript{176} Substantive reliability, on the other hand, is established when the statement is given in such circumstances that it makes it highly unlikely the evidence is false.\textsuperscript{177} For example, in \textit{Khan} the child’s statement that she had been sexually assaulted by a doctor was substantively reliable due to the presence of semen being found on her shirt.\textsuperscript{178}

While the framework does not perfectly line up with these procedural and substantive safeguards, it does attempt to replicate it in some respects. For example, the questions in part 1 of the three-part inquiry can be seen as touching on procedural issues, like if the expert is qualified to give testimony, and on the substantive side, whether the expert’s testimony is created in such a way that their opinion can be seen as reliable. Similarly, factors like whether the expert was able to identify an accused based on their work, but without knowing who the accused was before, would also provide some substantive guarantee of reliability.

\subsection*{4.3.3 Preventing Wrongful Convictions:}

As pointed out in \textit{Trochym}, dubious types of expert evidence have contributed to many wrongful convictions, and it only stands to reason that a more rigorous examination about the reliability of this type of evidence could prevent this phenomenon in the future. Indeed, in the Goudge Inquiry, Justice Goudge points out that the jurisprudence has often ignored the importance of reliability, 

\begin{flushend}

\textsuperscript{174} Bradshaw, \textit{supra} note 13, at para 1  
\textsuperscript{175} Ibid., at para 26  
\textsuperscript{176} Ibid., at para 28  
\textsuperscript{177} Ibid., at paras 30-31  
\textsuperscript{178} R. v. Khan, [1990] 2 SCR 531 [\textit{Khan}]
and supports the notion that evidence must pass a certain threshold of reliability. Aside from proactively trying to prevent dubious experts through substantive questioning, it is also possible that creating a more rigorous test will incentivize justice system actors to approach expert evidence in a more critical light. Namely, by focusing the trial judge, the Crown, and Defense on the importance of reliability, they will be incentivized by the adversarial system to find weaknesses in an expert’s testimony, instead of focusing their arguments on things like if the expert’s testimony is necessary. As mentioned above, it is not realistic to always expect that trial judges will be able to discern the threshold reliability of expert evidence on their own, but that they will instead need to rely on motivated Counsel to point out flaws in the proposed testimony.

4.3.4 Addressing Troubling Trends:

As will be discussed more in Part IV, two major concerns this new framework will address, is a sometimes laissez-faire attitude towards reliability, and a growing focus on characterizing evidence as scientific or not. Regarding the former issue, making reliability its own first step in a new proposed framework, regardless of whether the evidence is novel or not, creates a significant new requirement for expert evidence to meet. As to the latter point, while this framework does treat evidence differently, depending on whether it is scientific or not, this does not necessarily have any effect on the level of scrutiny shown towards that evidence (as it does now). For example, if a forensic expert is giving a non-scientific opinion which will have an important effect on the defendant’s guilt or innocence, the framework expects they will have some support aside from their experience to bolster their opinion. Namely, that the expert can present evidence that they diligently followed a method that can capably provide reliable results, or if original research, provide their own work that does the same.

---


180 Ibid., at pg 481
4.3.5 Shifting the Burden:

Last, whereas the current test places a large burden on Counsel to conduct effective cross-examination, and the jury to determine whether the evidence in ultimately reliable, this framework shifts more of the burden onto the experts themselves. This appears to make more sense than the current system, as the expert is better positioned than Counsel and the jury to defend the reliability of their field, and the basis of their opinion. As mentioned above, this would entail the expert to show their work to Counsel and the trial judge before the voir dire, creating an expectation that they need to defend their work.\(^{181}\) Moreover, given that “faced with an expert's impressive credentials and mastery of scientific jargon, jurors are more likely to abdicate their role as fact-finders and simply attorn to the opinion of the expert in their desire to reach a just result,”\(^{182}\) it seems safer to relieve the jury of some of the burden of assessing reliability. Thus, instead of asking the jury to largely determine whether a type of expert practice is reliable or not, in the absence of a lot of information, the expert should now be expected to play a greater role in defending the contents of their own opinion. Once this greater burden is satisfied by the expert, the jury will then continue to play its traditional role of deciding how much importance they will give this evidence.

4.3.6 Conclusion:

This section presented a new framework for admitting expert evidence, one that was more focused on the reliability of the expert’s evidence. Having presented the new framework, the section then presented some principles that can help guide a trial judge’s analysis in this new system. Last, this section showed some of the benefits of this new focus, in essence showing that the framework is both more ideologically more coherent and practically beneficial. In the next section, the framework will be applied to two cases where evidence of dubious quality was admitted, and discuss how the new framework would treat this evidence.

---

\(^{181}\) Chin 2018, *supra* note 163

\(^{182}\) D. (D.), *supra* note 73, at para 53
5 Part IV: Applying The New Framework:

While the latter section of Part III claimed that the new framework would create a better admissibility test for expert evidence, the purpose of this section is to put these claims into action by reviewing two cases that would be differently decided under the new framework.

5.1.1 Reliability and Awer:

As briefly touched on earlier, the Awer case from the Alberta Court of Appeal is a good example of an Appeal Court not being overly concerned with reliability.183 There, the accused was charged with sexually assaulting a sleeping woman at a party. An important piece of evidence presented by the Crown was the results of a penile swab which showed that the victim’s DNA was on the defendant’s penis.184 The Crown’s DNA expert testified that based on the amount of DNA found on the complainant, it must have been transferred by a wet source (such as a vaginal secretion or saliva)185 and not a dry one (such as fabric).186 This expert’s conclusion was largely based on trade publications that suggested wet DNA transfers in greater quantity than dry DNA, and on their own personal experience that they had never seen a dry DNA transfer that large.187 In contrast, the defense expert testified that one could not reliably conclude whether the source of the DNA came from a wet or dry source.188 Moreover, the expert said that the Crown witness misinterpreted the articles they cited in support of their position.189 Ultimately, the trial judge preferred the evidence of the Crown witness, citing his experience of never having seen such a large DNA transfer from a wet source, and along with other evidence, convicted the defendant.190

At the Court of the Appeal, the majority upheld the admission of the evidence on a number of grounds. First, they did not find any error in the trial judge’s analysis, and therefore would give

---

184 Ibid.
185 Ibid.
186 Awer ABCA, supra note 105, at para 9
187 Ibid.
188 Ibid., at para 10
189 Ibid.
190 Ibid., at paras 16-22
their reasoning deference.\textsuperscript{191} Second, they found that the statement did not go the ultimate issue, like the impugned evidence in \textit{Sekhon}, and was therefore not that prejudicial.\textsuperscript{192} While perhaps technically true, if the trial judge accepted the idea that the DNA transfer was wet, it seems difficult to draw any other conclusion than guilt. Last, the majority did not find any problem with the expert relying on trade publications to support his opinion, and were fine with the trial judge largely discounting the defense expert’s testimony.\textsuperscript{193} Notably, the majority did not ever really dwell on whether the evidence was reliable, instead saying that the trial judge must have decided it was.\textsuperscript{194} In contrast, the dissent found fault with the Crown expert relying on his own experience, especially because DNA experts typically have scientific grounding for their conclusions, pointing out that “numerous cases demonstrate the potential for anecdotes and unsystematic observations of experts to be followed.”\textsuperscript{195} In absence of empirically validated studies that would support the expert’s postulations, the dissent said that it was incumbent for the trial judge to use his gatekeeper role to exclude the evidence.\textsuperscript{196}

Having regard to the framework and the principles of the new test, it is likely that the expert’s testimony in \textit{Awer} would be excluded for being unreliable. Turning to step 1 of the inquiry, the field of DNA transfer seems to be capable of being reliable, and the Crown expert was qualified. However, it does not appear that the opinion in this case was based on reliable methods, as the expert’s opinion was mostly a product of personal anecdotes. As pointed out by the dissent, basic questions like how many wet samples the expert had seen, what sizes were they typically, and what sizes were dry transfers in comparison, were never asked of the expert.\textsuperscript{197} Similarly, Likwornik points out that if the expert had only been asked to testify on this point, and could not bootstrap off his real expertise as a general DNA expert, it is unlikely that his testimony would have been accepted.\textsuperscript{198} Given that this area of expertise would be amenable to having those

\textsuperscript{191} Ibid., at paras 58-61
\textsuperscript{192} Ibid., at para 59
\textsuperscript{193} Ibid., at paras 62-84
\textsuperscript{194} Ibid., at para 57
\textsuperscript{195} Ibid., at para 122
\textsuperscript{196} Ibid., at paras 127-132
\textsuperscript{197} Ibid.
\textsuperscript{198} See Chin 2018, supra note 103, at footnote 207
questions answered, but none were provided by the expert, the framework would exclude the evidence at this stage.

Some of the principles laid out also support the exclusion of this evidence. First, the DNA expert had no data to point at to support his opinion except for contested trade publication studies, even though it would not seem particularly difficult to bring data to support the expert’s anecdotal observations if the field was well-established. Alternatively, if there was little published information in this area, perhaps it would have sufficed for the expert to have provided detailed records of their past interactions with dry or wet DNA to ground their conclusions. Second, although perhaps not directly touching on the ultimate issue, this evidence was very damning and should have been treated with a high degree of skepticism without highly reliable evidence. Indeed, this type of evidence seems like the poster child for creating a wrongful conviction. Namely, extremely probative evidence that may be based on highly unreliable science, which the trial judge chooses to accept with the hope that other adversarial tools will come to the rescue.

5.1.2 Characterization and Aitken:

Another problem this new framework hopes to solve is of courts being overly focused on classifying evidence into different categories, which can come at the expense of reliability concerns. As alluded to before, while the proper application of the current test can lead to sound admissibility decisions, a benefit of the new framework is that eliminates the need for trial judges to engage in discretionary decisions that may have a significant outcome on whether reliability is an important focus of their analysis.

For example, in Aitken\(^{199}\) the British Columbia Court of Appeal upheld the admission of expert evidence given by a podiatrist, largely based on the fact that the evidence was classified not to be novel or scientific. Frustratingly, the current framework allows for this type of parsing, and as the trial judge came to that decision, the evidence was not treated with the level of scrutiny outlined in J. (J.)\(^{200}\) nor was much focus put on reliability at all. There, a Mr. Kelly was tasked with viewing security footage of a suspected shooter, and 6 known video images of the defendant, to determine

\(^{199}\) R. v. Aitken, 2012 BCCA 134 [Aitken]
\(^{200}\) Ibid., at para 74
if the gait of the 2 parties matched. Using a 6 point scale developed by a forensic science group (the reliability of which was never proven), Mr. Kelly concluded that there was a ‘very strong likeness’ between the shooter and the defendant’s gait. In addition, the expert opined that the defendant shared gait characteristics with only 1 percent of the population, but the trial judge excluded the expert’s evidence on this point, as the expert had not conducted a proper statistical analysis that would allow him to reliably reach that conclusion. Despite the seeming lack of factual foundation to the expert’s opinion, the trial judge accepted the expert’s expertise in reaching his ‘very strong likeness’ conclusion.

At the Court of Appeal, the Court agreed with the trial judge that podiatry is not a novel science, and therefore not subject to increased scrutiny. The Court also said that Mr. Kelly’s opinion fell under the type of expert whose opinion is based on specialized knowledge and experience, and was not scientific. Therefore, considerations like the Daubert factors had “limited relevance in a case like the one at hand where a witness's expertise is gained over a period of years through observation and experience in the professional realm.” While it may be true that podiatry as a general field is not novel, this conclusion misses the fact that forming opinions about a suspect’s gait based on security footage could be. Similarly, it may be true that Daubert would not have applicability in this case, but it does not mean that the expert should have not had to meet any reliability standards at all. Last, despite opining that there was a very strong likeness between the shooter and the defendant, the Court found that the expert did not touch on the ultimate issue, despite the inculpatory power of the expert’s testimony.

Again, under the new proposed framework, it is likely that this evidence would not be admitted. It is unclear whether gait analysis is capable of being a reliable field, and certainly the expert in this case did not bring any evidence to prove it was. The Crown witness did seem to have significant

---

201 Ibid., at para 63  
202 Cunliffe & Edmond, supra note 154, at pg 363  
203 Aitken, supra note 199, at para 64  
204 Ibid., at para 65  
205 Ibid., at para 74  
206 Ibid., at para 73  
207 Ibid., at para 80  
208 Cunliffe & Edmond, supra note 154, at pgs 359-361  
209 Aitken, supra note 199, at para 84
experience as a podiatrist, but it is unclear whether that also made them an expert in this specific area. Similarly, on the third question in part 1 of the new framework, the expert also failed to substantiate their opinion through using reliable methods. Although the expert claimed to be working on a system that could more accurately determine whether a gait matched, this was not yet seen to be reliable. Taken together, these factors would point towards the exclusion of this evidence.

Like in *Awer*, the expert also did not point to any validating studies to support their opinion. Without these, it is impossible to say if the defendant had a gait similar to a large percentage of the population, or that it was unique in some way. While there was a lot of circumstantial evidence implicating the defendant’s guilt, this evidence was still important and should have been treated with more scrutiny. Last, the circumstances in which the evidence was presented to the expert should be cause for concern. The expert was given video of the suspect and several videos of the defendant, and asked to compare the two to see if there was match.\footnote{Ibid., at para 15} As pointed out earlier, this can create a form of contextual bias where the expert is subtly pressured to identify an accused as the guilty party. In contrast, in *Woodcock*\footnote{R. v. *Woodcock*, 2010 ONSC 671 [*Woodcock*]} the Crown sought to avoid the contextual bias problem when looking for an expert who would be able to discuss what someone carrying a gun would look like on video. There, they provided the expert with footage from a mall, and without providing the expert “any details of the investigation…asked him whether he could identify any person displaying the characteristics of armed persons.”\footnote{Ibid., at para 1} The expert picked the defendant out of the video and was allowed to provide his testimony about the characteristic body movements someone with a gun would make.\footnote{Ibid.}

**Conclusion:**

This section showed that under the new framework, experts who attempt to present opinions based on their experience, in the absence of validating data will be treated skeptically. All the more so, when the opinion the expert makes is amenable to empirical testing.
6   Part V: Re-Examining Excluded Areas Of Expertise:

Given the focus on reliability in the proposed new framework for admitting expert evidence, and this thesis’ skepticism with necessity as a particularly useful factor within this new framework, a likely outcome of this new framework will be that evidence once seen as unnecessary will now become admissible. That is, if experts in areas such as eyewitness testimony, false confessions, and delayed disclosure, can prove that their field is reliable, the new framework would almost compel the acceptance of their testimony, instead of saying it is better left to a jury instruction. This seems like a better way to treat this evidence, because it removes the subjective determination of what a trial judge believes is necessary, and it also allows Counsel to call evidence in the way that they prefer. That being said, given that it is unlikely that experts in these areas will be called whenever this type of evidence is present, jury instructions will still remain as a way to convey this information to the jury.214

6.1.1 Eyewitness Identification:

In McIntosh,215 the Court of Appeal for Ontario rejected a defence expert’s testimony about the frailties of eyewitness identification, instead ruling that these frailties could be better expressed through a jury instruction.216 More recently, this decision was upheld by the same court in Frimpong.217 The Court’s four reasons in McIntosh to reject expert testimony in this area were that: it was unclear how recognised an area of science this was, the weaknesses of this identification were already within the jury’s knowledge, accepting that jurors could not assess identification evidence could raise questions about the integrity of the jury system, and it was easier to to use a jury instruction.218

On the first point, numerous legal scholars have pointed out that contrary to the decision, eyewitness identification is a well-developed area of study, which is very conducive to

215 R. v. McIntosh, [1997] OJ No 3172 (CA) [McIntosh]
216 Ibid.
217 R. v. Frimpong, 2013 ONCA 243 [Frimpong]
218 McIntosh, supra note 215
experimentation to test such things as cross-racial identification, the accuracy of recalling faces, etc.\(^{219}\) On the second point, numerous experiments have shown that ordinary people have many misguided views about the accuracy of eyewitness instructions.\(^ {220}\) For example, there is a significant body of research which shows that juries can be easily swayed by such factors as witness certainty, which in reality does not correlate with identification accuracy.\(^ {221}\) The third point is puzzling, because the Court’s judgement seems to express the idea that the legal system should pretend that juries can accurately assess the adequacy of a particular eyewitness’ identification, even if that is not the case, because the alternative would call into question the integrity of juries. But this is silly, as the point of expert witnesses is to assist the jury in areas where they could not be expected to have knowledge, and in these areas nobody questions the integrity of the jury system. On the last point, this thesis has already expressed that it does not find it particularly convincing that a judge can decide when something is better presented through a jury instruction. Last, a body of research suggests that jury instructions are ineffective at disabusing juries’ trust in the reliability of this identification.\(^ {222}\) Given all of these factors, and the role that mistaken eyewitness testimony has played in wrongful convictions, it seems appropriate to allow this type of expert evidence when the expert can prove their reliability.

6.1.2 False Confessions:

Another area with strong research backing, but which experts have not been allowed to testify, is in the area of false confessions. In Osmar,\(^ {223}\) while not discounting the future helpfulness of evidence about false confessions, the Court of Appeal said that the expert’s testimony was unnecessary because it “was not about matters on which ordinary people are unlikely to form a correct judgment.”\(^ {224}\) While the Court of Appeal may hope this to be the case, mock jury experiments show that confessions are considered very impactful, even if the mock jury is aware

\[^{219}\text{Dufraimont 2008, supra note 214, at pgs 267-270}\]
\[^{220}\text{Ibid.}\]
\[^{221}\text{Ibid., at pgs 267-268}\]
\[^{222}\text{Ibid., at pgs 306-307}\]
\[^{223}\text{R. v. Osmar, 2007 ONCA 50 [Osmar]}\]
\[^{224}\text{Ibid., at para 68}\]
of pressure tactics used in the police interrogation.\textsuperscript{225} Even in situations where jurors recognize in the abstract that false confessions exist, it is unclear whether a jury instruction will help them suss out a false confession as effectively as an expert witness.\textsuperscript{226} Again then, this is another area where the new framework may allow previously dismissed expertise.

6.1.3 Revisiting $D. (D.)$:

As talked about earlier, in $D. (D.)$, Justice Major ruled that a jury instruction could sufficiently express the idea to the jury that an adverse inference could not be drawn from a complainant’s delayed disclosure of abuse. In reaching this decision, Justice Major said that a jury instruction should align exactly with what the expert’s actual testimony was,\textsuperscript{227} indicating he believed it to be relevant and reliable, but unnecessary. While expertise in this area does not lend itself to the type of experimentation available in the eyewitness context, there are undoubtedly front-line workers who have experience dealing with sexual assault complaints, and can dispel myths that are forbidden in Canadian law. Thus, while not saying this type of evidence would always be seen as reliable, the assumption from $D. (D.)$ that a jury instruction is better suited to this type of evidence, should no longer be the going presumption.

Other considerations that may lean towards the admissibility of this type of evidence is the possibility that a jury instruction is not as effective as dispelling ‘rape myths’ as an expert responding to live questioning is.\textsuperscript{228} As pointed out by Craig, law reform to eliminate such things like the hue and cry assumption were not made as a result of the general public realizing this assumption was wrong, but from a concerted effort made by feminist legal scholars to eliminate these misconceptions from the law.\textsuperscript{229} Indeed, Craig still sees decisions made after $D. (D.)$ that still seem to accept the idea that if a complainant does not make a complaint right away, this reflects

\begin{footnotes}
\item[226] Dufraimont 2008, supra note 214, at pgs 271-273
\item[227] $D. (D.)$, supra note 73, at para 68
\item[229] Elaine Craig. “The Relevance of Delayed Disclosure to Complainant Credibility in Cases of Sexual Offence” (2011) 36 Queens LJ 551, at pgs 582-584 [Craig]
\end{footnotes}
badly on them. These facts challenge the idea that a simple instruction to the jury can counteract long-engrained assumptions, and would suggest as McLachlan does in the minority in D. (D) that expert evidence in this area is needed. Moreover, given the complexity of peoples’ behaviour after sexual assaults, a jury instruction also assumes that a trial judge will be able to craft a sufficient explanation about this behaviour, or if relying on precedent, know enough about this area of psychology to see if their model instruction remains accurate.

6.1.4 Conclusion:

Given the new framework’s focus on reliability, this section discussed how several social science areas which have been deemed as unnecessary under the current framework may now be admissible. This section argued that this approach is more theoretically sound than a judge making a very subjective decision about a piece of evidence’s necessity.

---

230 Ibid.,
231 Bala, supra note 228
6.2 Conclusion:

This thesis has discussed the modern expert evidence rule, and has chronicled its shifts over time. After having discussed how the jurisprudence on this issue has evolved, this thesis looked at the strengths and weaknesses of the current admissibility framework. Finding many faults with it, this thesis then proposed a new framework, which would be more focused on reliability and the validating of expert testimony. Having laid out this framework, this thesis then applied it to two disputed cases, and showed how the evidence in these cases would have been treated differently. Last, this essay concluded by taking some principles of the new framework, and argued that certain types of testimony that have been typically disallowed in favour of jury instructions should now be admissible.
6.3 Bibliography:

6.3.1 Articles:


Roach, Kent & Gary A. Edmond. “A Contextual Approach to the Admissibility of the State’s Forensic Science and Medical Evidence” (2011) 61 U Toronto LJ 343


6.3.2 Books:


6.3.3 Cases:


*R. v. Abbey*, 2009 ONCA 624

*R. v. Abbey*, 2017 ONCA 640

*R. v. Aitken*, 2012 BCCA 134

*R. v. Awer*, 2016 ABCA 128

*R. v. Awer*, 2017 SCC 2


*R. v. Bingley*, 2017 SCC 12

*R. v. Bradshaw*, 2017 SCC 35


*R. v. France*, 2017 ONSC 2040

*R. v. Frimpong*, 2013 ONCA 243

*R. v. Gager*, 2012 ONSC 1472

*R. v. J. (J.*)*, (1998) RJQ 2229 (CA)

*R. v. J. (J.*)*, 2000 SCC 51


*R. v. Khelawon*, 2006 SCC 57
R. v. Lavallee, [1990] 1 SCR 852


R. v. McIntosh, [1997] OJ No 3172 (CA)

R. v. McGeen, 2016 ONSC 5572

R. v. Meecham, 2019 ONSC 494


R. v. Osmar, 2007 ONCA 50


R. v. Sekhon, 2014 SCC 15

R. v. Shafia, 2016 ONCA 812

R. v. Trochym, 2007 SCC 6

R. v. Williams, [1995] OJ No 1012 (CA)

R. v. Woodcock, 2010 ONSC 671

United States of America v. Burns, 2001 SCC 7

White Burgess Langille Inman v. Abbott and Haliburton Co., 2015 SCC 23