MEMORANDUM

To The Law Commissions for England & Wales and Scotland

From Allen & Overy LLP

Our ref RJAF/PERSONAL-FARNHILR LT:21498226.1

Date 7 February 2019

Subject Automated Vehicles: Response to Consultation Paper 240

Introduction

1. Consultation Paper 240 (the Consultation) is a very thorough and well considered review of a substantial body of law. Inevitably, any response to such a paper focuses on areas of difference or points for further consideration. That should not in any way detract from our view as to the quality and comprehensive nature of the Consultation.

2. Many of the areas that the Consultation addresses are not ones that arise in the field of commercial litigation, in which we practice, and so we express no view on them. Our response is focussed on questions 17, 18 and 19 only, which do raise issues as to which we have experience and, in certain cases, some reservations as to how the law might operate both in practice and as a matter of principle.

3. We note at the outset that the views set out in this response are those of Allen & Overy LLP. They are not necessarily the views of our clients, either individually or collectively.

4. For reasons set out more fully below, our position is:

   a. Question 17(1): The operation of the Automated and Electric Vehicles Act 2018 (the 2018 Act) is, in our view, sufficiently clear. Subject to the broader points we make on the imposition of liability on autonomous systems we consider that an automated vehicle should be held to the standard of a human driver equipped with any additional data available to the automated vehicle through its sensors. For reasons developed more fully in response to question 17(2), we do not consider that the defence of contributory negligence arises where another party is 100% liable because we do not consider that in such circumstances liability arises under the 2018 Act at all.

   b. Question 17(2): In principle we consider that issues of causation can properly be left to the courts: the rules themselves have been addressed repeatedly and at length and the application
of those rules is fact specific and something that the courts are well qualified to determine. We are concerned, however, that the examples given in the Consultation appear to suggest that the normal rules on causation would not apply in the case of automated vehicles. To the extent that is what the Consultation proposes, we would not support such an approach. We consider that it is wrong as a matter of principle and extremely difficult to apply in practice.

c. Question 17(3): We consider that data retention is potentially problematic, and not simply for the reasons raised in the Consultation. As we address in more detail in connection with question 18, the relevant window for the accumulation of data is arguably the lifetime experience of the automated vehicle. Such volumes of data are unquestionably unmanageable. Giving notice immediately after an accident is unlikely to help when much of the data that influences how the artificial neural network that underpins the automated vehicle’s decision making operates will have been acquired before the accident.

d. Question 18: We consider that the product liability regime in the Consumer Protection Act 1987 (the 1987 Act) is unlikely to work well (or, in some cases, at all) in the context of automated vehicles. There are significant evidential issue in establishing the existence of a defect in an artificial neural network, both due to the volumes of relevant data involved and due to the impossibility of properly analysing the relevant algorithms. We further consider that the distinction between consumers and non-consumers may, in the context of automated vehicles, no longer be justifiable, at least on the grounds originally put forward in support of that distinction.

e. Question 19: We see a number of potential issues in this regard. First, just as there are evidential issues with showing a defect under the 1987 Act, there are the same issues with establishing breach of the duty of skill and care in connection with an automated vehicle’s decision making. Secondly, artificial neural networks will have applications in a whole range of industries. We consider that, at least as a starting point, such technology should be treated in the same way regardless of the use to which it is put. As such, the approach adopted should not be driven by factors that are relevant to automated vehicles but which may not be relevant in other fields of application (for example, mandatory insurance). We regard this as highly significant for the law of obligations; indeed, in some ways it seems to us that this is the most significant question in the Consultation. Thirdly, we are concerned that a different approach for automated vehicle liability may create an uneven playing field, creating an economic drag on the introduction of this technology. We consider that tort law should avoid such an outcome where, as here, there is no justification for it.

Question 17(1): Are sections 3(1) and 6(3) on contributory negligence sufficiently clear?

5. We consider that the provisions are sufficiently clear to be workable.

6. We note the point made in paragraphs 6.35 to 6.37 of the Consultation that an automated vehicle may, by virtue of its greater range of sensors, be better placed to detect a hazard than a human driver would be in the same conditions. However, we do not take the reference to “person” in section 3(1) to be limited to a human driver. That term is followed by the words “other than the insurer”, which in almost every case will be a corporate entity rather than an individual. That reflects the use of “person” in section 2 and, therefore, section 6(3). As such, we consider the term is being used widely and does not imply that the courts should treat the issue as if a human driver had been involved.

7. As to the specific point regarding the superior ability of automated vehicles to operate in low light conditions (or, indeed, any other conditions that would be adverse to human perception), subject to what we say below about the appropriate standard, we consider that this specific issue could be addressed through the use of a variable standard, comparing the automated vehicle’s conduct to the
hypothetical reasonably competent driver with access to the same information as the automated vehicle. A similar approach is adopted in the law around directors’ duties where the particular director has specialist skills or qualifications.¹

8. As a matter of principle, we do not consider that contributory negligence can be a 100% defence to a claim involving the 2018 Act where the other road user is 100% responsible. For reasons developed more fully below, we do not consider that in such circumstances the accident would have been caused by an automated vehicle at all and so section 2(1) of the 2018 Act is never triggered. Legally, the distinction is clear: the cause of action never arises, rather than there being a defence to it (as would be the case for contributory negligence). Economically we consider that the result is the same, whichever route is used.

Question 17(2): Do you agree that the issue of causation can be left to the courts, or is there a need for guidance on the meaning of causation in section 2 of the 2018 Act?

9. In principle we consider that issues of causation can and should be left to the courts. The rules on causation, in respect of breach of contract, negligence and breach of statutory duty, have recently been revisited by the Supreme Court and clearly restated. The application of those rules will, by necessity, be fact dependant and can best be handled by the judge addressing the facts of the specific dispute.

10. We do have some concern, however, arising out of the two examples given at paragraphs 6.47 and 6.48 of the Consultation. The first example involves the automated vehicle being shunted by another vehicle into the car in front. As the Consultation makes clear: “In these circumstances, the automated vehicle is in the chain of causation, but not at fault.” The second example involves the automated vehicle taking evasive action “to avoid an erratic cyclist” and hitting a parked car. Although it is not explicitly stated that the automated vehicle is not at fault, that conclusion is implicit in the description of the cyclist’s riding as “erratic”.

11. In both examples we consider that the law is entirely clear: if the automated vehicle is not at fault (that is to say, it is not defective under the 1987 Act and did not fail to act with reasonable skill and care) there should be no prospect of insurers being found liable. There are three connected reasons for this.

12. First, we consider this to be the clear meaning section 2(1) of the 2018 Act. That provision only applies where “an accident is caused by an automated vehicle when driving itself”. As the explanatory notes to the 2018 Act explain:²

“This Part extends compulsory motor vehicle insurance to cover the use of automated vehicles in automated mode, so that victims (including the ‘driver’) of an accident caused by a fault in the automated vehicle itself will be covered by the compulsory insurance in place on the vehicle.”

13. In neither example is there any suggestion of a fault in the automated vehicle. The threshold for applying the 2018 Act is simply not met.

14. Secondly, in each case the true cause of the accident, at least applying the normal rules of causation, is the other road user, not the automated vehicle.

15. In Hughes-Holland v BPE Solicitors³ the Supreme Court made a number of points regarding causation:

¹ Companies Act 2006, section 174
² 2018 Act Explanatory Notes, paragraph 12, emphasis added
³ [2017] UKSC 21
a. The “but for” test is generally a necessary condition for recovery but is not always in itself sufficient. The law is concerned with assigning responsibility for the consequences of the breach; a defendant may not be deemed responsible in law for everything that flows from their act, even if that act is wrongful.4

b. This principle applies equally to claims for breach of contract, negligence and breach of statutory duty.5

c. Where a defendant is under a duty to perform a specific act, such as the provision of information: “He is responsible only for the consequences of that information being wrong. A duty of care which imposes upon the informant responsibility for losses which would have occurred even if the information he gave had been correct is not in my view fair and reasonable as between the parties.”6

d. The burden of proof in showing causation is on the claimant.7

16. In the two examples it is hard to see how the claimant could ever discharge its burden as regards the automated vehicle or its insurers. There is no sense in which the automated vehicle could be said to be responsible for what happened.

17. In the first example, the automated vehicle has no control over events at all. The outcome is purely a consequence of Newton’s second law of motion. Any object of equivalent mass and rigidity would behave in precisely the same way. Yet presumably there would be no question of liability if the automated vehicle were parked and switched off or being driven by a human operator but stationary (for example, at traffic lights) or, indeed, if the object were not a car (or even capable of being propelled) at all. The fact that the vehicle is automated is wholly irrelevant to the sequence of events. It is therefore hard to see how the automated nature of the vehicle could ever be said to be “causative” in English law.

18. The second example produces outcomes that, while less arbitrary, remain difficult to justify. The implication is that the automated vehicle responded in a reasonable way (hitting a car rather than a cyclist). In those circumstances we would not normally say that a human driver caused the accident. Again, therefore, the fact that the vehicle is driving itself is irrelevant; had a human been operating the vehicle, the same thing would have happened. The “but for” test, which as noted above is a necessary condition for recovery, is not satisfied.

19. To impose liability in such circumstances would not in any sense be connected to responsibility. As such it would appear, to quote the House of Lords, not to be fair and reasonable as between the parties. To the extent that there were compelling policy reasons, Parliament could of course have created a special rule of causation in the 2018 Act displacing the common law test. We do not, however, read section 2(1) as being such a rule, which in our view would require very clear language.

20. Finally, we are concerned that such a broad approach to causation would undermine the regime of recovery under the 2018 Act. Under section 5 the insurer will be able to recover against “any other person liable to the injured party in respect of the accident.” If the automated vehicle is found to have caused the accident, however, there has arguably been a break in any causative chain, the automated vehicle representing a novus actus interveniens. In such circumstances, the parties who set in train the relevant chain of events that resulted in the loss (the driver initially responsible for the collision and the cyclist, respectively) would not be liable to the injured party and so there would be

---

4 At [20]
5 At [47]. See also [28] and the reference to Banque Bruxelles Lambert SA v Eagle Star Insurance Co Ltd [1997] AC 191 at 211H
6 At [29], quoting Banque Bruxelles at 214D
7 At [53]
no basis of recovery against them. The issue is particularly problematic in our view because once something is found to be a *novus actus interveniens* there is no issue of apportionment: the automated vehicle would be treated as the sole cause of the damage to the car in front.

21. It could be argued that the situation is more accurately characterised as one of joint tortfeasors or several tortfeasors causing the same loss. In such cases, the claimant would be entitled to sue either tortfeaster for the full amount and it is for the defendant to claim contribution against the other tortfeaster or tortfeasors. That is not straightforward on the facts contemplated in the examples, however, because it is unclear what tort the maker or owner of the automated vehicle can be said to have committed in circumstances where there is no suggestion of fault. As the Consultation notes at paragraph 6.26, the 2018 Act is not designed to allocate responsibility, so cannot give rise to a tortious liability; the vehicle is not defective under the 1987 Act; nor has there been negligence. There can be no joint or several liability for a tort or torts where one party has not committed a tort at all.

22. If, under section 2 of the 2018 Act, the automated vehicle is to be considered the cause of the accident in the absence of any fault, it must therefore comprise a *novus actus interveniens*, where no issue of fault on the part of the latter actor need arise. Clearly, it would be an unattractive argument to run for the party setting in motion the chain of events. At the same time, section 5 only becomes relevant if section 2 is engaged, meaning there must have been a finding that the automated vehicle caused the accident, meaning in turn that, under *Hughes-Holland*, the law considers the automated vehicle to be responsible for the loss.

23. In our view the result ought not to arise because there ought, in both examples, to be a finding that the automated vehicle did not cause the loss. If, however, it is to be found to have caused the loss even where there is no suggestion of fault, then the language of the 2018 Act strongly suggests that it is to be treated as having done so for all purposes. That would be to create some form of no fault liability for the insurers of automated vehicles while simultaneously creating issues for those insurers in recovering against the parties that are at fault. We do not consider either outcome to have been the intention of the 2018 Act.

24. There is a suggestion at paragraph 6.50 of the Consultation that the courts could or should look to find that the automated vehicle caused the loss because it may be the only way to secure a meaningful award of damages to the ultimate victim. We do not agree, either as a matter of law or as a matter of economic principle. The latter point gives rise to broader considerations and we address them in response to Question 19 at paragraphs 72 to 78 below. As a simple legal point, however, the cases on causation make no mention of any such an exception to the general approach to causation and, while it was open to Parliament to create such an exception in the 2018 Act, it would have required clear wording. In fact, as we note at paragraphs 12 to 13 above, the explanatory notes suggest strongly that Parliament’s concern was to apply the regime of the 2018 Act only where there was fault on the part of the automated vehicle.

**Question 17(3): Do any potential problems arise from the need to retain data to deal with insurance claims?**

25. The Consultation focuses on the data acquired by the automated vehicle in the run-up to, during and immediately after any accident (whether or not a collision is involved). What the consultation does not address is the data acquired by the automated vehicle during its lifetime. We consider that this is also a highly relevant consideration.

26. The decision making system at the heart of automated vehicles is an artificial neural network. Broadly, this is an algorithm or series of algorithms that will analyse large data sets to identify
patterns and then use those patterns to make predictions. The quality of those predictions (and so the quality of the automated vehicle’s “decision making”) will turn on two things: the quality of the algorithm; and the data it has available to it in identifying the relevant patterns. The same algorithm will therefore perform differently depending on the data available to it throughout its operational life (either through direct exposure or through being periodically updated).

27. We address in more detail below the significant difficulties that we consider to exist in ascertaining liability in the case of such algorithms, many of which are already touched upon in the Consultation. Our point here is more of an evidential one: even assuming those transparency issues could be overcome, without an accurate record of the data upon which the algorithm was educated, there is an evidential gap in ascertaining whether that algorithm was defective or acted in a way that was unreasonable. As noted at paragraph 6.54 of the Consultation, an automated vehicle generates terabytes of data every day, any part of which may be relevant. Storage of even a few days’ worth of data would be wholly impractical. We see this evidential gap as a real issue that needs to be considered and addressed in the liability analysis.

Question 18: Is there a need to review the way in which product liability under the 1987 Act applies to defective software installed into automated vehicles?

28. We see significant difficulty with applying the regime under the 1987 Act to the decision making software installed in automated vehicles for three principal reasons.

a. The lack of transparency inherent in artificial neural networks, identified in paragraphs 9.76 to 9.88 of the Consultation, renders it difficult if not impossible to ascertain whether there was a defect in the automated vehicle and, if so, what it was. We do not consider that Wilkes v DePuy International Limited⁹ necessarily solves this issue for the reasons given in Ide v ATB Sales Ltd¹⁰.

b. An artificial neural network operates by recognising patterns in data and using those patterns to make a prediction. Even where there is a very strong pattern, producing a high likelihood of the prediction being correct, there are no guarantees: the action taken by the automated vehicle might have been the “correct” choice (in the sense of minimising harm) in 99% of cases, just not in the case before the court. That means the possibility of causing loss is inherent in such systems. Under the 1987 Act, dangers inherent in the nature of the product are not defects.

c. As we have noted in the context of Question 17(3), the accuracy of the predictions generated by an artificial neural network depends on the data upon which it is trained. This gives rise to evidential issues, which we raised in the context of Question 17(3). It also gives rise to an issue of principle. Transfers of certain types of data are prohibited under both UK and EU law. This means that the same algorithm will perform differently depending on whether the owner has a larger or smaller data-set on which to train it. If the algorithm fails to perform as well as it might have done purely because data was not legally available, is that a defect in the algorithm?

29. We address first the broad questions of principle that we consider give rise to these issues before going on to address the issues themselves.

30. In our view, automated vehicles and the artificial neural networks upon which they rely give rise to two areas of legal divergence from other products.

---

⁹ [2016] EWHC 3096 (QB)
¹⁰ [2008] EWCA Civ 424
31. First, although we agree with what is said at paragraph 6.65 of the Consultation as to liability under the 1987 Act being strict, we do not consider that this is no fault liability in the sense that the term is sometimes used.\textsuperscript{11} Harm alone is not enough; there must be an additional factor (a defect in the case of the 1987 Act; failure to act with reasonable skill and care in the case of common law negligence) to create liability. That involves understanding both what happened, in terms of the sequence of events, and how the product came to perform in the way that it did.

32. Historically, even sophisticated products would not truly “learn” from experience. They would perform in an ascertainable, pre-programmed way. Traditional computer “decision making” used pre-set weightings of defined factors. In a sense it was not computer decision making at all; it was human decision making applied by a computer. Inputs could be directly traced through to outputs. That machine-implemented but human determined decision making process could therefore be interrogated, understood and evaluated by another human.

33. Neural networks are different. Rather than weightings being pre-set, the algorithms constantly adjust them by reference to experience, their own or that downloaded from other devices. The decision making process often operates through a number of layers, meaning a change of inputs can be amplified or muffled as it passes through the layers. Artificial neural networks thereby “learn” from experience. Moreover, they are able to process much larger volumes of data at much higher speeds than is possible for a human brain. As such, they are able to identify patterns that other systems (including humans) cannot. Indeed, that is their attraction and the reason they can perform at the levels they do.

34. Moreover, that learning is done in a context that is very different to the reasoning process of a human brain. As the neurosurgeon Henry Marsh recently noted in an article for the Financial Times:\textsuperscript{12}

> "Brains do not come as isolated entities as do computers. They come with bodies, to which they are intimately connected. The bodies are their interface with the world, through which they learn to move, and, some would argue, there cannot be thinking without embodiment."

35. Artificial neural networks tend to operate in a very linear fashion, therefore. They do not demonstrate qualities such as associative reasoning or “common sense” in checking their answers because they have no way in which to do so. Either there is a sufficiently strong pattern or there is not; that is the only thing that affects their decision.

36. Each of these features presents challenges to understanding why an artificial neural network reached the decision that it did. Taken together the challenges are amplified. Any system for imposing liability will need to address the evidential difficulties to which those challenges give rise.

37. Second, in cases involving artificial neural networks the nature of the product is different to any previous product. Historically it was possible to distinguish a product from its operator; there was a difference between what the product did, or was supposed to do, and how it was used. That distinction is now less clear cut. By way of simple example, a car with cruise control will automatically maintain a certain speed (either an absolute speed or by reference to surrounding vehicles) but the decision to set that speed will be taken by the human driver. The question of whether the speed was achieved would be subject to the test under the 1987 Act; the question of whether the right speed was selected and if not, why, would not. In an automated vehicle both the determination of the appropriate speed and its implementation will be carried out without human intervention.

\textsuperscript{11} See Clerk & Lindsell on Torts, 22nd Ed paragraph 1-69
\textsuperscript{12} https://www.ft.com/content/2e75c04a-0f43-11e9-acde-4d9976f1533b
38. The 1987 Act was not intended to deal with operators or true decision makers.\textsuperscript{13} That manifests itself in what are, in our view, two key shortcomings. It arguably fails to recognise that in decision making there can be a spectrum of what is acceptable: one can disagree with a decision without necessarily saying it is definitively wrong.\textsuperscript{14} The question of defect is far more binary. It also creates a protected class, consumers, the justification of which now arguably needs to be revisited.

39. Turning to the first issue facing the application of the 1987 Act to artificial neural networks, transparency, we see significant potential problems. As the European Group on Ethics in Science and New Technologies noted last year in its Statement on Artificial Intelligence, Robotics and Autonomous Systems:\textsuperscript{15}

“It is unfortunate that some of the most powerful among these cognitive tools are also the most opaque. Their actions are no longer programmed by humans in a linear manner. Google Brain develops AI that allegedly builds AI better and faster than humans can. AlphaZero can bootstrap itself in four hours from completely ignorant about the rules of chess, to world champion level. It is impossible to understand how exactly AlphaGo managed to beat the human Go World champion. Deep learning and so-called ‘generative adversarial network approaches’ enable machines to ‘teach’ themselves new strategies and look for new evidence to analyse. In this sense, their actions are often no longer intelligible, and no longer open to scrutiny by humans. This is the case because, first, it is impossible to establish how they accomplish their results beyond the initial algorithms. Second, their performance is based on the data that have been used during the learning process and that may no longer be available or accessible.”

40. We note in particular two issues identified by the European Group on Ethics in Science and New Technology. The fact that the data used in training will no longer be available in the case of automated vehicles is a point we address above in response to question 17(3): retention of that volume of data (which will, in the overwhelming majority of cases where there is no accident, be entirely pointless) is simply impractical, but without it the algorithm’s decision making process is unintelligible.

41. Of equal concern is the lack of transparency inherent in the nature of the product. Generally, the issue of whether there is a defect is addressed by expert evidence. Such evidence:\textsuperscript{16}

“is admissible under section 3 of the Civil Evidence Act 1972 in any case where the Court accepts that there exists a recognised expertise governed by recognised standards and rules of conduct capable of influencing the Court's decision on any of the issues which it has to decide and the witness to be called satisfies the Court that he has a sufficient familiarity with and knowledge of the expertise in question to render his opinion potentially of value in resolving any of those issues.”

42. If the algorithm’s decision making is no longer open to human scrutiny, no witness could ever satisfy that test.

43. The issue can be highlighted by the example of AlphaGo, used by the European Group on Ethics in Science and New Technologies. AlphaGo was an algorithm developed by DeepMind Technologies. In 2016 it played a series of games of Go against the human world champion, Lee Sedol. AlphaGo won the series 4-1. The games are available on DeepMind’s website with a commentary from a series of high-ranking Go players.\textsuperscript{17}


\textsuperscript{14} Bolam v Frien Hospital Management Committee [1957] 1 WLR 582


\textsuperscript{16} Baring's Plc v Coopers & Lybrand (No2) [2001] FNLR 22 at [45]

\textsuperscript{17} https://deepmind.com/research/alphago/match-archive/alphago-games-english/
44. Taking Game 1 as an example, the human commentators considered AlphaGo’s moves 10 and 12 to be a “problem” with its opening. Presumably, if the game had been frozen at that stage and the question had been put to those commentators (who would typically qualify as experts on the rules and strategy of Go under the Barings test), they would have considered that AlphaGo’s performance was potentially defective. By move 14, however, the mood had shifted considerably: “the feeling in the room suddenly changed.” This highlights the issue: had the experts been asked the question at move 12 they would have found that AlphaGo was underperforming; by move 14 it was over-performing. The fact that AlphaGo’s approach could not be understood did not mean that there was a problem with it, however. To the contrary, once the human experts came to understand the approach, they agreed that it was not simply a valid strategy but a strong one.

45. The same issue arises for automated vehicles. The fact that a human expert cannot understand why the algorithm behaved in the way that it did does not mean that behaviour was defective. The car may have steered into a cyclist to avoid a patch of ice that would have caused it to lose control and hit a bus queue of people. Or it may have a bias against cyclists. Unlike AlphaGo, where the game continues thereby providing more information about the quality of the decision, in the case of an accident there will be no further information available. Without being able to penetrate the decision making system, it is simply impossible to say whether the decision was or was not the result of a defect in the algorithm’s decision making process.

46. The Consultation proposes a potential solution to this issue by reference to the approach of Hickinbottom J in Wilkes. We consider that this addresses a rather different problem, however. In Wilkes the issue was whether the loss had been caused by an identified or identifiable defect.18 In such a case, as noted in Ide v ATB Sales Ltd19, the 1987 Act simplifies the claimant’s task: he or she only has to establish there was a defect and that the defect caused the loss. Often, as was the case in Ide itself,20 if other possible causes can be disproved, the only cause left is a defect.

47. Cases involving artificial neural networks are different, however, because it will always be perfectly possible that the automated vehicle made the decision it did for valid reasons, even if those reasons are not understandable to a human being. As such, there will always be the possibility that there was no defect. The opacity of the decision making process makes the probability of that “no defect” scenario impossible to assess.

48. That makes the situation much more like that identified at paragraph [4] of Ide:

“What is impermissible is for a judge to conclude in the case of a series of improbable causes that the least improbable or least unlikely is nonetheless the cause of the event; such cases are those where there may be very real uncertainty about the relevant factual background (as where the vessel was at the bottom of the sea) or the evidence might be highly unsatisfactory. In that type of case the process of elimination can result in arriving at the least improbable cause and not the probable cause.”

49. In some cases, of course, it will be tempting to conclude that the only possible reason for the accident was a defect. In our view that would impermissibly conflate the concepts of harm and defect. The 1987 Act is clear in requiring both. To move from strict liability to truly no fault liability would be a highly significant change in the law. For reasons we set out at paragraphs 72 to 78 below, we do not believe that would be a sensible direction for the law to take.

50. One response would be to make the reasoning process of the underlying algorithms more transparent and so make it more straightforward to identify the existence or otherwise of a defect. We see two significant issues with this:

---

18 In Wilkes the question was hypothetical because there was no defect: [2016] EWHC 3096 (QB) at [117] to [134].
20 At [20]
a. As the Consultation notes at paragraph 9.80, there is a risk that making the algorithms transparent simultaneously robs them of their value. As the technology improves and the gap between what the algorithms can do and what the human brain can follow grows, that risk will only get worse. By limiting the complexity of these systems to a level that can be understood by humans the benefits they can bring, notably in terms of increased safety, must also be limited. If the deployment of automated vehicles (and, indeed, artificial neural networks more generally) is not at some stage to be stifled, therefore, the legal mechanism for assigning liability needs to address this gap.

b. As the European Group on Ethics in Science and New Technologies notes, in order properly to understand the algorithm’s decision making one would need to have access to the data on which it was trained. As the Consultation notes, and as is noted above, the volumes of data render that impractical. Without that data, however, it is difficult to see how a conclusion can properly be reached as to whether there was a defect and, if there was, whether it was the responsibility of the programmer, the trainer, the data provider or some combination of the three. Those parties may all belong to the same organisation, but they may not. If they are all separate parties it may be impossible to carry out the exercise of establishing who is “liable to the injured party in respect of the accident” contemplated by section 5 of the 2018 Act.

51. In our view these practical issues associated with transparency are potentially insurmountable under the regime of the 1987 Act.

52. Underlying these practical points is a general difficulty of principle. The concept of “defect” is difficult to apply to a true decision maker.21 As is clear from the law of negligence, which would apply to a human driver, in a case of true decision making a range of decisions may be reasonable, even if it later transpires that some of them resulted in avoidable loss. There seems no good reason why an algorithm that is, on average, safer than a human driver should be held to a different standard. To the contrary, as we discuss at paragraphs 72 to 78 below, once automated vehicles are permitted for general use both legal and economic principle suggest that the mechanism for attributing liability should be at the very least neutral as between automated and human driven vehicles. Holding the decision maker to a different standard (i.e. under the 1987 Act), particularly where that standard is so difficult to apply in practice, would not achieve that goal.

53. The second issue we see with applying the 1987 Act arises from the way that artificial neural networks operate. Essentially, they analyse data to detect patterns. Based on the strength of that pattern they make predictions. Patterns, of course, are simple correlations; there may be causative links underlying them but there may not, some may be coincidences. As we note above, unlike the human brain the algorithm operates in an entirely abstract environment. It has no concept of common sense or life experience against which it can check its predictions; it only has the pattern or patterns that it has identified. The algorithm therefore has no way of distinguishing between causation and correlation and so inherent in its reasoning is the chance that however strong the pattern may have been historically, it will not repeat in the future. In such cases, the algorithm’s prediction will be inaccurate.22

54. It seems to be accepted that article 6 of Directive 85/374/EEC, and so section 6 of the 1987 Act, was never intended to apply to risks that are, because of their particular nature, inherent to a product.23

21 We note that in the case of animals, which as Clerk & Lindsell notes are chattels that are animate and automotive, a quite different statutory regime applies under the Animals Act 1971. That regime does not depend on the precise workings of the animal’s decisions, which are unknowable, but rather on whether the animal falls within in a particular class or has demonstrated dangerous tendencies in the past. We are not suggesting that such a regime should apply to automated vehicles (or other autonomous devices). As we note below, in our view the proper liability regime merits significant further consideration. We note the example simply to highlight the issues faced by the 1987 Act when it comes to decision makers.

22 See, for example, https://www.ft.com/content/f14db820-26cd-11e8-b27e-ce62a39d57a0

23 See generally Clerk & Lindsell paragraph 11-61
On the face of it, that should normally therefore provide a robust defence in the case of an automated vehicle: it is in the nature of the product that it makes predictions from patterns and it did just that. Again, though, the difficulty is transparency. It is arguably impossible to know what probabilities the algorithm calculated or how. The defence would potentially be rather too robust if it simply comprised of saying that these systems operate on probabilities and that is a risk you take with them.

55. Moreover, although the defence seems well settled its basis is unclear. It may be that it is based on a willing trade-off by society: to derive the benefits we accept the risk. That trade-off is obvious in the case of, for example, hot drinks: most people understand that if they want to have coffee hot, it is hot for all purposes and can therefore cause burns.24 The same is not necessarily true of artificial neural networks, where many people will not understand the principles of how they make decisions, and even experts will not understand the detail behind a particular decision. The 1987 Act offers no real guidance, which would mean significant uncertainty potentially for a prolonged period. That is, self-evidently, highly undesirable.

56. The final issue is whether an automated vehicle can be said to be defective where:

   a. The same or substantially identical systems operate with a higher level of safety because they have access to larger data sets; and

   b. The only reason why this vehicle did not have access to such data sets is legislation prohibiting or prescribing transfers of data.

57. Addressing this question focuses attention on who precisely is liable and for what. Development and production of the artificial neural networks, at least in some cases, is likely to be based on a model already used in microchip manufacture, which a recent article in The Economist highlighted:25

   “Mr Muller’s “drawings” are anything but simple. They are computer code which give Arm’s customers a blueprint for the construction of microprocessors, information-processing machines so complex that firms are happy for Arm to shoulder the burden of their fundamental design. Those clients—consumer-hardware giants such as Apple and Huawei; chip companies such as Broadcom and Qualcomm—pay Arm one-off licence fees to access the design code, add to it what they will, then pay royalties on every product they ship containing it. Apple’s popular a-series mobile processors, for example, are built in this way.”

58. In such a model the chip manufacturer may well be able to point to an excellent safety record with many, or indeed all, other users of its algorithms. That would tend to suggest that whatever else might have caused the problem, it was not their code.

59. In turn, however, the purchaser of the artificial neural network (the car manufacturer) would potentially have recourse to two defences under the 1987 Act.

   a. Assuming the defect is the relationship between the algorithm and the volume of data available, the lack of transparency means that the defect is difficult, if not impossible to discover. Under section 4(1)(e) of the 1987 Act it is a defence to show that the defect could not have been discovered given the state of scientific and technical knowledge at the relevant time.

   b. Section 4(1)(a) provides a defence if “the defect is attributable to compliance with any requirement imposed by or under any enactment or with any Community obligation”. The car manufacturer may not have been able to obtain any larger data set by virtue of requirements in privacy legislation. The question of how data sets can been expanded

---

24 B (a child) v McDonald’s Restaurants [2002] EWHC 490 (QB) at [80]
25 https://www.economist.com/business/2019/01/05/masayoshi-son-wants-arms-blueprints-to-power-all-tech
without violating privacy so as to improve the performance of autonomous systems is currently under consideration by the EU Commission.\textsuperscript{26} We consider this to be an important issue in any consideration of liability for automated vehicles.

60. The concern must be that if the 1987 Act is intended to apply to automated vehicles, the scope of these defences will be critical. There is a risk that any claim may fall between defendants, resulting in victims being unable to recover. Equally, the various parties in the manufacturing process will need certainty if they are to manage the litigation risk. As matters currently stand we consider that the 1987 Act fails to address either issue.

61. The Consultation raises a further, very specific point, concerning the application of the 1987 Act to consumers.\textsuperscript{27} We note that the justification for the EU Council Directive (85/374/EEC), which underpins the 1987 Act was:\textsuperscript{28}

“liability without fault on the part of the producer is the sole means of adequately solving the problem, peculiar to our age of increasing technicality, of a fair apportionment of the risks inherent in modern technological production”

62. It may be that such a division remains “a fair apportionment of the risks” given the differing ability of parties to spread risk through insurance or otherwise. It is no longer the case, however, that businesses (or indeed manufacturers) are any better placed to identify or understand the risk of these products than are consumers. The lack of transparency applies equally to the programmer of the algorithm that drives the artificial neural network, the end user and every party in the supply chain between them. To the extent that a further review of the 1987 Act is now required, and for the reasons given we believe that it is, this question ought equally to be revisited.

Question 19: Do any other issues concerned with the law of product or retailer liability need to be addressed to ensure the safe deployment of driving automation?

63. Question 18 focussed on the 1987 Act. This question is more general in nature and so we raise here points of broader concern:

a. The application of the points made above to the law of negligence.

b. The importance of treating all uses of artificial neural networks in a way that is consistent and economically rational.

c. The economic risks of a liability regime that unduly favours a finding of liability.

64. The issues that arise under the 1987 Act as to establishing a defect equally arise in any attempt to show a breach of the common law duty to act with reasonable skill and care. That exercise is process driven, not simply outcome driven:\textsuperscript{29}

“It is the duty of an employer, in considering whether some precaution should be taken against a foreseeable risk, to weigh, on the one hand, the magnitude of the risk, the likelihood of an accident happening and the possible seriousness of the consequences if an accident does happen, and, on the other hand, the difficulty and expense and any other disadvantages of taking precaution.”


\textsuperscript{27} Paragraphs 6.67-6.69

\textsuperscript{28} EU Council Directive (85/374/EEC), preamble

\textsuperscript{29} Morris v West Hartlepool Steam Navigation Co Ltd [1956] AC 552 at 574
65. As a matter of legal principle, if the algorithm’s decision making process simply cannot be understood, there is no means for a claimant to discharge the evidential burden.

66. A presumption along the lines of *res ipsa loquitur* cannot work where there is simply no evidence and the possibility of a non-negligent cause exists.\(^{30}\) The points we make above on both the lack of transparency and the evidential gap apply equally here. For similar reasons, shifting the burden of proof (as happens, for example, under the Misrepresentation Act 1967) does not offer a solution: it will likely be equally impossible for the manufacturer to demonstrate the process by which the decision was reached. We are concerned that if resort were had to such approaches it would amount to a move to no fault liability for manufacturers. That is a very significant step. If a no fault system of liability is to be introduced for automated vehicles (a position that, for reasons we address below, we consider to be undesirable) it should at least be done clearly and transparently.

67. An alternative solution would be to move the negligence test from a “process” focussed approach to an “outcomes” focussed approach. This is arguably more feasible for negligence than it is under the 1987 Act, since in the law of negligence a range of non-negligent outcomes may exist. Such an approach is potentially workable but creates the real risk that the benefits of artificial neural networks are lost. By definition, the range of “reasonable” outcomes will be those contemplated by human experts. If the algorithm produces a result that they cannot explain, even if in fact the reasons were valid, liability could ensue.\(^{31}\)

68. Turning to the second point, the artificial neural networks that underpin the decision making of driverless cars are equally being used in other fields. By way of example:

   a. In a speech on 21 May 2018 the UK Prime Minister, Theresa May, sought to launch “a whole new industry around AI-in-healthcare” focussing in particular on early stage cancer detection. The Prime Minister noted that early detection in lung, bowel, prostate and ovarian cancers could reduce deaths by 10% within 15 years.

   b. The Economist reported on 31 March 2018 that “America’s air force and defence department are working with C3 IoT, a startup, to scan maintenance logs and past technical problems for signs that aircraft are wearing out.” Such an approach, possibly used alongside sensors in engines, could largely automate inspection of aircraft.

   c. Gartner has estimated that between 2017 and 2021 the share of customer-service interactions worldwide handled entirely by AI will rise fivefold, to 15%; by 2019 at least 40% of such interactions will involve at least some element of artificial intelligence.

   d. The McKinsey Global Institute has estimated that by 2030 up to 375 million people, representing 14% of the global workforce, could see their jobs automated.

69. While we recognise that the Consultation is limited to automated vehicles, in principle the same approach to civil liability should apply equally to other uses of artificial neural networks. If that is to be achieved we are of the view that any proposal for imposing civil liability should be tested against the operation of the technology generally and then applied to automated vehicles, not vice versa.

70. That should be so even where significant differences exist as regards other legal aspects of the use of this technology. Specifically, automated vehicles, like all other cars used on public roads, will be required to carry mandatory third party insurance, as the Consultation notes. The Consultation goes further, however, in appearing to suggest that this would be a legitimate factor for the courts to consider in establishing causation and, by extension, liability.\(^{32}\)

---

\(^{30}\) Percy & Charlesworth on Negligence 14th Ed paragraph 6-24

\(^{31}\) See paragraph 50(a), above.

\(^{32}\) See, for example, paragraph 6.51.
71. Under the current mechanisms for determining liability for defective products the capacity of the defendant to pay an award of damages is irrelevant to establishing liability. We do not see any basis to change that approach. If the defendant’s ability to pay were relevant it ought to make no difference whether that was a result of personal means or some form of third party indemnity (including insurance). As such, it would seem to follow that a wealthy defendant would be found liable in circumstances where a less wealthy defendant running the same system in the same way would not be found liable. There is no legal basis for such an approach to civil liability. Given that there is no such rule in the case of personal resources, it ought equally not to exist for third party indemnities. Unquestionably, the means of a defendant to pay on any award of damages is a relevant consideration for a claimant in deciding whether to bring a claim; it should not affect the question of whether that claim is a good one, however.

72. Turning to the economic issues, given the difficulties of applying current systems of liability there may be a temptation to use a no-fault system, imposing ultimate liability on manufacturers. Although such a system could be structured in a way that was easy to apply, we see real difficulties with it in other respects.

73. First, if Parliament had intended to put in place such a system it would logically have done so in the 2018 Act, either by making the manufacturer directly liable or by giving the insurer automatic recourse against the manufacturer under section 5. That is not the structure of the 2018 Act. To the contrary, it requires consideration of who ultimately would have been liable to the victim.

74. Second, if a no fault system applies to automated vehicles but the current regime continues for human drivers, the law creates an economic disincentive to move to automated vehicles. A human driver who caused damage would continue to have, various defences open to them, most notably discharging the duty of skill and care. Similarly, the manufacturer of the non-automated vehicle will be able to show there was no defect, inherent dangers in the product, the current state of technology meant the defect was unknown at the time of release and so forth. The defendant in a no fault regime would have no such defences.

75. The result would be that it would be easier to establish liability in the context of automated vehicles than it would be in the context of traditional vehicles. That would be the case even though automated vehicles will only be allowed on the roads when they can be shown to be statistically safer than traditional ones. The lower risk product will therefore be more at risk of liability. As such, it will be a higher risk for insurers and so face higher premiums. The greater risk of liability and the greater cost of insuring that risk create an economic drag on the introduction of automated vehicles.

76. Nor does that drag simply mean an financial loss for the economy. In Pathway to Driverless Cars the UK Government estimated that 94% of road deaths were attributable to human error and that the use of automated vehicles could reduce pollution and social exclusion. Moreover, as we note above, the logical starting point is that the same approach should apply to different applications of the same technology. The four cancers identified by the Prime Minister in her May 2018 speech kill roughly 67,800 people annually in the UK; a liability system that permitted the use of a technology reducing that figure by 10% has obvious benefits.

77. Such factors are material factors in tort law, as was recognised by Hale LJ in Sutherland v Hatton: “The law of tort has an important function in setting standards for employers as well as for drivers, manufacturers, health care professionals and many others whose carelessness may cause harm. But

---

35 Pages 16-17
34 https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type
if the standard of care expected of employers is set too high, or the threshold of liability too low, there may also be unforeseen and unwelcome effects upon the employment market.”

78. In the context of autonomous vehicles, we can see a justification for reducing the likelihood of liability compared to human drivers, given the comments made in the Consultation about the reduced risk of accidents. We can see no basis for increasing the likelihood of liability in the absence of any increased risk.

79. Finally, we consider that a no-fault system could create a highly invidious position for some potential users of artificial neural networks. Using the example of early stage cancer diagnosis, for which we have estimated figures, if the neural network detects 10% more cancers than a human oncologist, it may be difficult, under the law of negligence, for a hospital not to use that technology. Yet by using the technology the hospital moves into a no-fault regime that makes it far harder to defend claims. Either way it loses. Considerable thought needs to be given to such risks before making changes to the current regime.

Conclusion

80. As we note above, we are grateful to the Law Commissions for the extensive work that has obviously gone into the Consultation, and for the opportunity to comment on it.

81. On balance, we are less confident than the Law Commissions about the current system’s capacity efficiently to determine civil liability relating to automated vehicles (and, indeed, decision making systems powered by artificial neural networks more generally). We would therefore encourage a thorough review of the operation of the current systems, both statutory and common law, and the options for reform.

82. Obviously, there are various potential legal issues, which we have sought to address here as briefly as possible. We would be happy to discuss these points or assist further in the development of this area if that would prove useful. If there are any areas on which we could usefully expand or provide clarification, please contact Richard Farnhill (by e-mail at [redacted] or by telephone on [redacted])

Allen & Overy LLP
7 February 2019