Causation in the Constitutional Court: *Lee v Minister of Correctional Services*

Anton Fagan[*]

I Introduction

The central question in *Lee v Minister of Correctional Services* was framed as follows:

> Assuming, first, that the defendant had negligently and wrongfully failed to maintain an adequate system for the management of tuberculosis in a prison in which the plaintiff had been incarcerated and, secondly, that the plaintiff had contracted tuberculosis whilst so incarcerated, had the defendant committed a delict against the plaintiff?[1]

The Supreme Court of Appeal concluded previously that the defendant had *not* committed a delict against the plaintiff.[2] Its grounds for so concluding were three-fold. First, the question as to whether the defendant had committed a delict against the plaintiff had to be determined by applying the South African common law. Secondly, the common law contained the following four rules:

(1) A person who performed negligent conduct can be delictually liable for harm suffered by another person only if the negligent conduct was a factual cause of the harm;[3]

(2) negligent conduct was a factual cause of harm if and only if, but for the negligent conduct, the harm would not have occurred;[4]

(3) a court is justified in concluding that a defendant's negligent conduct was a factual cause of a plaintiff's harm if and only if the plaintiff has proved that the negligent conduct probably was a factual cause of that harm;[5] and

(4) a plaintiff has proved that a defendant's negligent conduct probably was a factual cause of the plaintiff's harm if and only if the plaintiff has proved that, but for the negligent conduct, the harm probably would not have occurred.[6]

Thirdly, the condition stipulated by rule (4) above had not been satisfied. That is, the plaintiff had not proved that, but for the defendant's negligent failure to maintain an adequate system for the management of tuberculosis, the plaintiff probably would not have contracted the disease.[7]

[*] WP Schreiner Professor of Law, University of Cape Town. This work is based on research supported by the National Research Foundation.

[1] *Lee v Minister of Correctional Services* [2012] ZACC 30, 2013 (2) SA 144 (CC)('*Lee CC*').

[2] *Minister of Correctional Services v Lee* [2012] ZASCA 23, 2012 (3) SA 617 (SCA)('*Lee SCA*').

[3] Ibid at 628I.

[4] Ibid at 626E, 628J–629C, 629F.

[5] Ibid at 626E.

[6] This result is entailed by (2). It is also assumed to be true by the *Lee SCA* Court. Ibid at 631F.

[7] Ibid at 631D–G.

The Supreme Court of Appeal's conclusion, namely that the defendant had not committed a delict against the plaintiff, was rejected by a majority of the Constitutional Court. According to the *Lee CC* Court, the Supreme Court of Appeal had erred in its conclusion because it had proceeded from a mistaken premise. That premise was that the South African common law contained the four rules set out above. The majority of the Constitutional Court accepted that the common law contained rules (1) and (3).[8] It denied that the common law contained rules (2) and (4).

Rule (2) makes it a necessary condition, for negligent conduct to be a factual cause of harm, that, *but for the negligent conduct*, the harm would not have occurred. That, said the *Lee CC* Court, was not the common law. Why not? Because the common law accepted that, in exceptional circumstances, it would be a sufficient condition for conduct to be a negligent factual cause of harm if, *but for the conduct*, the harm would not have occurred.[9] In other words, according to the *Lee CC* Court, the common law did not contain rule (2) above, but rather rule (2)* below:

(2)* Negligent conduct was a factual cause of harm if and only if either:
  (a) but for the negligent conduct the harm would not have occurred; or
  (b) the circumstances were exceptional and, *but for the conduct*, the harm would not have occurred.

Of course, if rule (2) is replaced by rule (2)*, rule (4) necessarily has to change as well. Specifically, it has to be substituted by rule (4)* below:

(4)* A plaintiff has proved that a defendant's negligent conduct probably was a factual cause of the plaintiff's harm if and only if:
  (a) the plaintiff has proved that, but for the negligent conduct, the harm probably would not have occurred; or
  (b) the circumstances were exceptional and the plaintiff has proved that, *but for the conduct*, the harm probably would not have occurred.

However, according to the *Lee CC* Court, even rule (4)* does not yet properly reflect the common law. For, insisted the majority, the common law recognised a third condition, in addition to conditions (a) and (b) above, sufficient for a plaintiff to have proved that a defendant's negligent conduct probably was a factual cause of the plaintiff's harm. It was this: the circumstances were exceptional and the plaintiff has proved that, but for the negligent conduct, *the risk of that harm would have been reduced*.[10] In other words, according to the majority of the Constitutional Court, the common law contained neither rule (4) nor rule (4)* above, but rather rule (4)** below:

(4)** A plaintiff has proved that a defendant's negligent conduct probably was a factual cause of the plaintiff's harm if and only if:
  (a) the plaintiff has proved that, but for the negligent conduct, the harm probably would not have occurred; or

[8] *Lee CC* (note 1 above) at 161A–D, 163D, 168C, 168E–F, 168H–I, 169C–D, 173B, 173F.
[9] Ibid at 162D–E, 163C, 163F–H, 164B–C, 166B–D, 168B, 173E–F.
[10] Ibid at 168G–H, 169C–D, 170B–C.

(b) the circumstances were exceptional and the plaintiff has proved that, *but for the conduct*, the harm probably would not have occurred; or
(c) the circumstances were exceptional and the plaintiff has proved that, but for the negligent conduct, *the risk of that harm would have been reduced*.

It was on the basis of conditions (b) and (c) above that the *Lee CC* Court set aside the Supreme Court of Appeal's finding that the defendant had *not* committed a delict against the plaintiff. According to the majority, both condition (b) and condition (c) were satisfied. The plaintiff had proved that, but for his incarceration, he probably would not have contracted tuberculosis.[11] He also had proved that, but for the defendant's negligent failure to maintain an adequate system for the management of tuberculosis, the risk of his contracting tuberculosis would have been reduced.[12] On both grounds, therefore, the plaintiff had shown that 'there is a probable chain of causation between the negligent omissions by the [defendant] and [the plaintiff's] infection with TB'.[13]

This article focuses on several questions raised by the way in which the majority of the *Lee CC* Court dealt with factual causation. The first question, which is the concern of section II below, is whether it is of any practical significance whether factual causation is determined by application of rules (2) and (4) or by application of rules (2)* and (4)**. It certainly mattered to the parties to the *Lee CC* case. But does it matter more generally? Are there other situations in which different outcomes will be yielded, depending upon whether factual causation is determined by applying rules (2) and (4) or by applying rules (2)* and (4)**? If so, how common are those situations?

The second question, which is discussed in section III below, is whether the majority of the *Lee CC* Court got the common law right. Did the South African common law, at the time *Lee* was decided, really contain rules (2)* and (4)**, as the majority believed? Or did it rather contain rules (2) and (4), as the Supreme Court of Appeal had assumed? Also discussed in section III is a third question, which is closely connected to the second. If, contrary to what the *Lee CC* Court asserted, rules (2)* and (4)** were not in fact part of the common law, should they nonetheless be? That is, are there any analytical or moral reasons to prefer a common law containing rules (2)* and (4)** over a common law containing rules (2) and (4)?

The conclusion to this article briefly deals with a fourth question raised by the *Lee* case. If the Constitutional Court majority in *Lee* in fact got the common law wrong, would the common law have been altered by its mistake? That is, would the effect of *Lee* be that, henceforward, factual causation in our law is to be determined by application of rules (2)* and (4)** rather than by application of rules (2) and (4)? This question is of particular interest if there are not in fact any analytical or moral reasons favouring rules (2)* and (4)** over rules (2) and (4).

[11] Ibid at 168B.
[12] Ibid at 168G–H, 169C–170C.
[13] Ibid at 173B.

II The Difference the Rules Make

This section shows that it is of great practical significance whether factual causation is determined by application of rules (2) and (4) or by application of rules (2)* and (4)**. A great many situations exist in which different outcomes will be yielded, depending on whether the former two or latter two rules are applied. This palpable difference is most easily demonstrated by example. Consider then the following four hypothetical sets of facts:

> *The driver and the pedestrian:* A drives down a city road at 90 km/h. Without warning, B runs into the road and is killed by the impact of A's car. It was negligent for A to have driven at a speed in excess of 80 km/h. But it would not have been negligent for A to have driven at 80 km/h or slower. Had A not driven down the road, B would still have been alive. However, even if A had driven at 70 km/h to 80 km/h, B probably would still have been killed by the impact of the car.
>
> *The nurse and the patient:* Nurse A administers 100 ml of a certain drug to patient B. B has a reaction to the drug which kills B. It was negligent for A to have administered more than 50 ml of the drug to B. But it would not have been negligent for A to have administered 50 ml of the drug or less to B. Had A not administered the drug to B, B would still have been alive. However, even if A had administered 40 ml to 50 ml of the drug to B, B probably would still have had the reaction and died.
>
> *The doctor and the patient:* Dr A administers 80 ml of a certain drug to patient B in order to combat a certain condition. B nonetheless dies of the condition. It was negligent for A to have administered less than 100 ml of the drug to B. But A would have acted without negligence, had A administered 100 to 150 ml of the drug to B. Had A administered 140 to 150 ml of the drug to B, B probably would still have been alive (as the condition probably would have been successfully managed). However, had A administered 100 ml to 110 ml of the drug to B, B probably would have died of the condition in any event.
>
> *The engineer and the vintage-car owner:* Engineer A designs a retaining wall with a foundation of 1 metre for a property belonging to C who sells it on to B. As a result of a once-in-500-years earthquake, the wall collapses, destroying four vintage cars belonging to B. It was negligent for A to have designed a retaining wall with a foundation of less than 2 metres. But A would have acted without negligence, had A designed a retaining wall with a foundation of 2 metres or more. Had A designed a retaining wall with a foundation of 3 metres or more, B's cars probably would not have been damaged (as the retaining wall probably would not then have collapsed). However, had A designed a retaining wall with a foundation of 2 metres to 2.5 metres, B's cars probably would have been damaged in any event (as the retaining wall would then still have collapsed).

As I explain in subsection A below, if rules (2) and (4) are applied to these four examples, A could not be delictually liable for B's harm in any of them. As is explained in subsection B below, if rules (2)* and (4)** are applied, then more or less the opposite conclusion holds true: That is, A could be delictually liable for B's harm in all four.

To be strictly accurate, it should be noted that these divergent liability outcomes are not produced by the application of these rules alone – that is, by the application of only, on the one hand, rules (2) and (4) and, on the other, rules (2)* and (4)**. The divergent outcomes are produced, instead, by the application of these rules *in combination with* rules (1) and (3). However, it is a bit of a mouthful to speak every time of 'rules (1), (2), (3) and (4)' or of 'rules (1), (2)*, (3) and (4)**'. When

this article speaks of 'rules (2) and (4)', their combination with rules (1) and (3) will therefore often be assumed. Likewise when it speaks of 'rules (2)* and (4)**'. Sometimes the combination is not assumed. In such instances, the text should make that clear.

A Rules (2) and (4)

To understand how rules (2) and (4) are to be applied to the four hypothetical examples provided at the start of this section, one needs look no further than Corbett JA's dissenting judgment in *Siman & Co v Barclays National Bank*.[14] In the judgment, Corbett JA provided a detailed explanation of the so-called 'but-for' or '*causa sine qua non*' test. The explanation is worth quoting in full. It explains not only the application of rule (2), but also the application of rule (4). Corbett JA writes:

> In order to apply this test one must make a hypothetical enquiry as to what *probably* would have happened but for the unlawful act or omission of the defendant. In some instances this enquiry may be satisfactorily conducted merely by mentally eliminating the unlawful conduct of the defendant and asking whether, the remaining circumstances being the same, the event causing harm to plaintiff would have occurred or not. If it would, then the unlawful conduct of the defendant was not a cause in fact of this event; but if it would not have so occurred, then it may be taken that the defendant's unlawful act was such a cause. This process of mental elimination may be applied with complete logic to a straightforward *positive act* which *is wholly unlawful*. So, to take a very simple example, where A has unlawfully shot and killed B, the test may be applied by simply asking whether in the event of A not having fired the unlawful shot (ie by a process of elimination) B would have died. In many instances, however, the enquiry requires the substitution of a hypothetical course of lawful conduct for the unlawful conduct of the defendant and the posing of the question as to whether in such case the event causing harm to the plaintiff would have occurred or not; a positive answer to this question establishing that the defendant's unlawful conduct was not a factual cause and a negative one that it was a factual cause. This is so in particular where the unlawful conduct of the defendant takes the form of a negligent *omission*. [It has been] suggested that the elimination process must be applied in the case of a positive act and the substitution process in the case of an omission. This should not be regarded as an inflexible rule. It is not always easy to draw the line between a positive act and an omission, but in any event there are cases involving a *positive act* where the application of the 'but-for' rule requires the hypothetical substitution of a lawful course of conduct. A straightforward example of this would be where the driver of a vehicle is alleged to have negligently driven at an excessive speed and thereby caused a collision. In order to determine whether there was factually a causal connection between the driving of the vehicle at an excessive speed and the collision it would be necessary to ask the question whether the collision would have been avoided if the driver had been driving at a speed which was reasonable in the circumstances. In other words, in order to apply the 'but-for' test one would have to substitute a hypothetical positive course of conduct for the actual positive course of conduct.[15]

[14] *Siman & Co (Pty) Ltd v Barclays National Bank Ltd* 1984 (2) SA 888 (A).
[15] Ibid at 951B–H (References have been omitted and emphasis has been added).

The majority judgment in *Lee* quoted almost all of the above passage.[16] But it failed to grasp its central point. It is that, in order to determine whether negligent (or wrongful) conduct was a factual cause of harm, one has to mentally eliminate or 'think away' *as much of the conduct as, but no more of the conduct than*, was negligent (or wrongful). Thereafter one must ask: if *this much, but no more*, of the conduct were eliminated, would the harm still have occurred? If the answer is: yes, the harm would still have occurred, the negligent (or wrongful) conduct was not a factual cause of the harm. But if the answer is: no, the harm would not then have occurred, the negligent (or wrongful) conduct was a factual cause of the harm.

It was precisely to illustrate this point that Corbett JA provided his two examples – both of which involve a positive act. In the first example, A's shooting is, as Corbett JA put it, 'wholly unlawful'. To eliminate as much of A's conduct as, but no more than, is wrongful, one thus has to eliminate the shot in its entirety. In the second example, by contrast, it is not the driver's driving *per se*, but only his driving at an excessive speed, which is wrongful. Hence, to eliminate as much of the driver's conduct as, but no more than, is wrongful, one has to eliminate, not his driving, but only his speeding. And, of course, if one eliminates only the driver's speeding, one is left with him driving at a proper or reasonable speed.

There is yet another way of putting the point that Corbett JA made in the quoted passage above. In order to determine whether negligent (or wrongful) conduct was a factual cause of harm, one always has to posit or imagine a course of conduct that deviates from the actual one. However, and this is critical, the course of conduct which one posits or imagines should deviate from the actual conduct *no more than is necessary* to deprive the conduct of its negligent (or wrongful) character. To put it another way, the imagined conduct should alter the actual conduct *only just enough* to render it non-negligent (or lawful). In Corbett JA's first example, the alteration required one to imagine A not shooting at all. In Corbett JA's second example, by contrast, the alteration required one to imagine the driver still driving, but now at a proper speed.

The foregoing represents Corbett JA's explanation, in *Siman*, of rule (2). However, Corbett JA's use of the word 'probably' in the first line of the quoted passage, as well as his subsequent application of 'these principles' to the facts of the case, make it clear that he meant this to constitute an explanation also of rule (4).[17] All that is required is the insertion of the word 'probably' in the appropriate places. In other words, whether negligent conduct *probably* was a factual cause of harm is to be determined as follows:

(1) Mentally eliminate *as much of the conduct as, but no more of the conduct than*, was negligent.
(2) Ask the question: if *this much, but no more*, of the conduct were eliminated, would the harm *probably* still have occurred?
(3) If the answer is: yes, the harm *probably* would still have occurred, then the negligent conduct *probably* was not a factual cause of the harm.
(4) But if the answer is: no, the harm *probably* would not have occurred, then the negligent conduct *probably* was a factual cause of the harm.

[16] *Lee CC* (note 1 above) at 165D–J.
[17] See, in particular, *Siman* (note 14 above) at 918A–C, 920C, 921H.

The implications of this explanation for the first two examples provided at the beginning of this section should be obvious. After all, the example of *the driver and the pedestrian* just adds detail to Corbett JA's own example of the driver and the collision, which has already been discussed. As for the example of *the nurse and the patient*, the cast and the conduct may have been changed, but the basic plot remains the same. However, since the point that Corbett JA made in the above passage may have been misunderstood not only by the majority in *Lee CC* but also by others, it nonetheless is worth setting out those implications in full.

Consider first the example of *the driver and the pedestrian*. Step one, as has just been explained, is to eliminate as much of the driver's conduct as, but no more of his conduct than, is negligent. That is, we are to imagine a course of conduct that deviates from the driver's actual course of conduct to the minimum degree necessary to remove its negligent character. Thus, since the driver's negligence consisted in his driving too fast, we are to imagine him driving at the very highest speed that would not have been negligent. That speed, we are told, is 80 km/h. Step two is to ask whether, if the driver had driven at that speed – that is, the very highest speed that would not have been negligent: thus 80 km/h – B's harm probably would still have occurred. We are also told what the answer to this question is: it probably would. It follows, according to rule (4), that A's negligent conduct probably was *not* a factual cause of B's harm. It follows, when rule (4) is combined with rules (1), (2) and (3), that A could *not* be liable in delict for that harm.

Consider now the example of *the nurse and the patient*. Again we begin by eliminating as much of the nurse's conduct as, but no more of her conduct than, is negligent. In other words, we are to start by positing a course of conduct that differs from the nurse's actual course of conduct to the smallest extent necessary to cancel out its negligent nature. The nurse's negligence, according to the stated facts, consisted in her administering too much of a certain drug to B. We are therefore to imagine the nurse's giving B the very largest amount of that drug that it would not have been negligent of her to give to him. That amount, the stated facts tell us, is 50 ml. Next we are to ask whether, if the nurse had given B that amount of the drug – that is, the very largest amount that it would not have been negligent for her to give him – B's harm probably would still have occurred. Once again, the facts stated in the example tell us what the answer to this question is: even if the nurse had administered 50 ml of the drug to B, B probably would still have had the reaction in question and died. It follows, according to rule (4), that A's negligent conduct probably was *not* a factual cause of B's harm. And it follows, when rules (1), (2) and (3) are added into the mix, that A could *not* be delictually liable for that harm.

The implications of Corbett JA's explanation in *Siman* for the third and fourth examples provided at the outset of this section may be less obvious than for the first and second. The reason is that, whereas the first and second examples involve negligent *positive acts*, the third and fourth examples involve negligent *omissions*. Corbett JA provided no example of a negligent omission in his explanation. However, he made it clear that his explanation covered negligent omissions no less than negligent positive acts. In other words, irrespective of

whether negligent conduct took the form of a positive act or an omission, one is to determine whether that negligent conduct was, or was probably, a factual cause of harm by mentally eliminating *as much of the conduct as, but no more of the conduct than*, is negligent. To put this in the way it also was put earlier: one has to imagine a course of conduct that deviates from the actual conduct *no more than is necessary* to deprive the conduct of its negligent character.

It is true that, in the passage quoted above, Corbett JA stated that, in the case of unlawful omissions, application of the 'but-for' test usually requires 'the *substitution* of a hypothetical course of lawful conduct' rather than 'merely … *eliminating* the unlawful conduct of the defendant'.[18] However, it would be a mistake to infer from this requirement that Corbett JA meant to suggest that the application of rules (2) and (4) to negligent omissions differs fundamentally from their application to negligent positive acts. For he went straight on to point out not only that the line between positive acts and omissions is vague at best, but also that 'there are cases involving a positive act where the application of the but-for rule demands the hypothetical substitution of a lawful course of conduct'.[19] Moreover, he presented his second example – that is his example of the driver and the collision – as exactly such a case.[20] Corbett JA also discussed several earlier judgments in which, though the negligent conduct in question took the form of a positive act, the 'but-for' test was applied by substituting non-negligent for negligent conduct.[21] This alteration holds true of Corbett JA's own judgment in *Siman*: that is, Corbett JA applied the 'but-for' test by substituting a negligent positive act (the giving of 'incorrect advice') with a non-negligent one (the giving of 'correct advice').[22]

That the distinction between positive acts and omissions is of no consequence to the application of rules (2) and (4) is borne out by consideration of the third example: *the doctor and the patient*. In respect of the third example, as with the first and second examples, one has to start by imagining a course of conduct that differs from the doctor's actual course of conduct to a sufficient degree for it to lose its negligent aspect, yet to no greater degree than is necessary for the conduct to do so. The doctor's negligence, we are informed, consisted in her giving too small an amount of a certain drug to B. The question we must therefore ask ourselves is this: what is the very smallest amount of that drug that the doctor could have given B without being negligent? The answer to this question, we are told, is: a 100 ml. In respect of the third example, as with the first and second examples, the question one has to ask next is whether, if the doctor had given B that amount of the drug – that is, the very smallest amount that the doctor could have given without being negligent: thus a 100 ml – B's harm probably would still have occurred. We also find the answer to this question in the stated facts: it probably would. It therefore follows, because of rule (4), that the doctor's negligent conduct probably was *not* a factual cause of B's harm. And it follows,

[18] *Siman* (note 14 above) at 915E, 915B–C (emphasis added).
[19] Ibid at 915F–G.
[20] Ibid at 915G–H.
[21] Ibid at 916A–917H.
[22] Ibid at 918D–922E.

because of rules (1), (2) and (3), that the doctor could *not* be liable in delict for that harm.

The same analysis applies to the final example: *the engineer and the vintage-car owner*. That should come as no surprise. Its basic structure, if not its superficial details, is identical to the example of *the doctor and the patient*. The engineer was negligent because he failed to design a sufficiently deep foundation. For the engineer not to have been negligent, all that was required of him was to have designed a foundation 2 metres deep. However, even if he had designed a foundation of 2 metres in depth, the wall probably would still have toppled over and damaged B's cars. Consequently, the engineer's negligent conduct probably was not a factual cause of B's harm, by virtue of rule (4). Nor, therefore, could the engineer be delictually liable in terms of rules (1), (2) and (3).

B Rules (2)* and (4)**

The previous subsection showed that the application of rules (2) and (4) to the four hypothetical examples provided at the beginning of this section yields the same result: A's conduct did not factually cause B's harm, and thus A could not in any of them be liable for B's harm. This subsection explains that it is by no means certain that the same results would be yielded by the application of rules (2)* and (4)**. It is possible that, for the purposes of rules (2)* and (4)**, the circumstances in all four examples are exceptional rather than ordinary. In that event, whether A's negligent conduct probably was a factual cause of B's harm could (and possibly would have to) be determined by application of condition (b) or (c), rather than condition (a), in rule (4)**. Moreover, as is explained below, condition (c) in rule (4)** is satisfied in all four examples, while condition (b) in rule (4)** is satisfied in at least two of them.

Whether the circumstances in the four examples are exceptional rather than ordinary is a question that will be dealt with presently. For the moment, however, assume that they are. It follows that whether A's negligent conduct probably was a factual cause of B's harm could, in every example, be determined by application of what follows the conjunction in condition (c) in rule (4)**. That is, it could be determined by asking this question: but for A's negligent conduct, would the risk of B's harm have been reduced? For the reasons given below, this is a question which *necessarily* receives a positive answer.

Negligent conduct by definition is conduct which poses a risk of harm. In respect of any negligent conduct, it therefore is an analytic truth that, but for the negligent conduct, there would have been a reduced risk of harm.[23] Admittedly, it is not an analytic truth in respect of any negligent conduct and *any harm*, that, but for the negligent conduct, there would have been a reduced risk of *that* harm. Thus, for example, where a person performs conduct which is negligent because it poses a risk of harm to another person's property, and the latter person suffers harm to his body, it is not necessarily true that, but for the former's negligent conduct, there would have been a reduced risk of the harm actually suffered by the latter.

[23] This point was in fact clearly enunciated by the minority *Lee CC*. *Lee CC* (note 1 above) at 183C–D.

However, it is different where a person performs conduct which is negligent because it poses a risk of a particular harm to another person, and the latter person suffers that very harm. In that case, it *is* an analytic truth that, but for the negligent conduct, there would have been a reduced risk of the harm actually suffered. To put this more formally: if X's conduct was negligent because it exposed Y to a risk of harm Z, and Y suffered harm Z, then it necessarily is the case that, but for X's negligent conduct, there would have been a reduced risk of Y's suffering harm Z. Moreover, precisely that was the case in every one of the four hypothetical examples. In every one of them, A's conduct was negligent because it exposed B to a risk of the harm which B indeed suffered, namely physical injury or death. Hence, in respect of all four of the examples, it *is* an analytic truth that, but for A's negligent conduct, there would have been a reduced risk of B's harm.

What about condition (b) in rule (4)**? On the assumption that the circumstances are exceptional in all four examples, A's negligent conduct in every one of them probably would be a factual cause of B's harm, if the following question has a positive answer: Is it true that, but for A's conduct, B's harm would probably not have occurred? This question yields a positive answer, at least in the examples of *the driver and the pedestrian* and of *the nurse and the patient*. To determine the answer to this question in respect of the first example, we have to (or at any rate may) eliminate, not A's driving too fast, but simply his driving down the road. To determine the answer to this question in respect of the second example, we have to (or may) eliminate, not A's administering too much of the drug, but simply her administering it at all. Then we must ask: what probably would have happened in that event? The answer, of course, is that in both cases B probably (in fact certainly) would not have suffered the harm he actually did suffer. In respect of both examples, therefore, it is the case that, but for A's conduct, B's harm probably would not have occurred.

Because the third and fourth examples – *the doctor and the patient* and *the engineer and the vintage-car owner* – involve negligent omissions rather than negligent positive acts, it is more difficult to answer the question posed by condition (b) in respect of them. However, one *could* describe the doctor's conduct in the third example as: the failure to administer 140–150 ml of the drug in question to B. And one *could* describe the engineer's conduct in the fourth example as: the failure to design a foundation of 3 metres or more. And one *could* then go on to ask: if we were to eliminate the doctor's conduct, so described, and the engineer's conduct, so described, what probably would have happened in each case? We know the answer: B probably would still have been alive in the first case and B's cars probably would not have been damaged in the second. It would therefore seem that with respect to the third and fourth examples, but for A's conduct, B's harm probably would not have occurred.

So, provided the four examples are exceptional in their circumstances, rule (4)** yields the result that A's negligent conduct probably was a factual cause of B's harm in all of them. It does so, in respect of all four examples, because of condition (c). It does so, *certainly* in respect of the two examples involving negligent positive acts and *possibly* in respect of the two examples involving negligent omissions, because of condition (b). These results have two further

consequences. First, in all four examples, the condition in rule (3) is satisfied. Secondly, it follows from the first consequence that the condition in rule (1) is satisfied too. It follows that A, in all four examples, could be held delictually liable for B's harm. At the very least, if A could not be held delictually liable to B in any of the four examples, then it would have to be because delictual liability requires the satisfaction of a desideratum other than factual causation.

However, this result does not yet mean that rules (2)* and (4)** yield contrary liability outcomes to rules (2) and (4) in respect of any of the four hypothetical examples. For this conclusion to follow, it would have to be the case, in addition, that the proviso stated at the beginning of the previous paragraph is met. That is, the circumstances in the four examples would have to be *exceptional*. This condition raises the following question: What, for the purposes of conditions (b) and (c) in rule (4)**, makes circumstances exceptional rather than ordinary? Unfortunately, the attempt to answer this question receives next to no assistance from the *Lee CC* Court. The analysis that follows is thus speculative in nature, and based on a fragment here and a passing remark there.

One possibility is that the circumstances of a case are exceptional if the application of rules (2) and (4) to that case would produce an *injustice*. This possibility is suggested by the following remark in the majority judgment: 'There are cases in which the strict application of the [but-for] rule would result in an injustice, hence a requirement of flexibility.'[24] Elsewhere, the judgment speaks of 'the injustice of an inflexible approach to factual causation'.[25]

These remarks are hardly helpful. The *Lee CC* Court clearly believes that the application of rules (2) and (4) to the facts of the case before them *would* result in an injustice. However, in the absence of some explanation as to the nature of that injustice, no reason exists to think that a similar injustice might not also be occasioned by the application of rules (2) and (4) to the four hypothetical examples.

A second possibility is that the circumstances of a case are exceptional if the application of rules (2) and (4) to the case would have the result that the question of factual causation became 'a mixed question of fact and law', so that '[t]he distinction between factual and legal causation made in our law becomes unnecessarily less clear'.[26] However, this too is unhelpful. The majority judgment's reason for thinking that the application of rules (2) and (4) might have this result is that it necessarily 'involves an evaluation of normative considerations'.[27] Why? Because it requires one to eliminate as much of the conduct as, but no more of the conduct than, *was negligent* (or wrongful); and because the degree to which conduct *was negligent* (or wrongful) is an evaluative or normative question. As is explained in the next section, it is doubtful that this reasoning is valid. But assume for a moment that it is. Clearly, the application of rules (2) and (4) would have this result no matter what case those rules were applied to. Judged by this criterion, therefore, *all* cases are exceptional.

[24] Ibid at 162D–E.
[25] Ibid at 173C–D.
[26] Ibid at 166F–G. See also at 167E–F.
[27] Ibid at 166E.

A third possibility, also suggested (but no more than that) by the *Lee CC* Court is that the factual predicates of a case are to be treated as exceptional if the harm suffered also impugned a constitutionally-protected interest and the conduct was negligent because it posed a risk of harm to that interest.[28] But this twist also does not work. The specified condition is satisfied by the facts of *Lee*. The harm suffered by the plaintiff was to his bodily integrity: a right protected by s 12 of the Constitution.[29] Moreover, the risk to this interest constitutes the defendant's negligence. The difficulty, however, is that the specified condition is also satisfied in every one of the four hypothetical examples. And it would make no difference if the condition were expanded so as to require also that 'the conduct was wrongful because it breached a constitutionally-grounded duty not to pose a risk of harm to that interest'. The right to bodily integrity in s 12 and the right to property in s 25 (as implicated in the example of *the engineer and the vintage-car owner*) undoubtedly justify the imposition of such duties, not only on the state, but also on private parties. In short, these rights may have direct horizontal application.[30]

A final possibility is suggested by the majority judgment's assertion that 'the wrong done to [the plaintiff] is not treated as a *mere* omission', as well as its assertion that '[t]he wrong [the plaintiff] complains of is, in our law, based on *being detained* in [certain] conditions'.[31] Perhaps the majority, in making these assertions, is drawing a line between two kinds of omissions: those constituted by the failure to protect another against a danger posed by a third party or an external force, on the one hand, and those constituted by the failure to protect another against a danger created by one's own prior (or concurrent) positive conduct, on the other. And perhaps it believes that cases involving omissions of the first kind (call them 'pure' omissions) are ordinary – and thus are to be dealt with by applying condition (a) in rule (4)**, whereas cases involving omissions of the second kind (call them 'impure' omissions) are exceptional – and thus are to be handled by applying conditions (b) and (c).[32]

This possibility holds more promise. It is by no means unusual for negligent conduct to take the form of a pure omission. Indeed, the doctor's failure to administer enough of the drug in the example of *the doctor and the patient* and the engineer's failure to design a foundation of sufficient depth in the example of *the engineer and the vintage-car owner* are both 'pure' omissions. It follows that, if the distinction between ordinary and exceptional circumstances were to be drawn in this manner, a large number of cases would be treated as ordinary rather than exceptional. There also would be a large number of cases in which the question

[28] Ibid at 167F–168A, 171A–B.

[29] Constitution of the Republic of South Africa, 1996.

[30] The specified condition could be further tightened up by requiring also that 'there was no other way of holding to account the person who negligently posed the risk to the constitutionally-protected interest'. Alternatively, it could be relaxed by requiring simply that 'the negligent conduct breached a constitutionally-grounded duty, owed to the person suffering the harm in question, either to perform or refrain from performing some act'. *Lee CC* provides some support for both propositions. *Lee CC* (note 1 above) at 167F–G, 170E–F, 171A–B. However, neither alternative would serve to distinguish the facts of *Lee* from the four hypothetical examples.

[31] *Lee CC* (note 1 above) at 163G, 167F–G (emphasis added in both quotes).

[32] This possibility was suggested to me by my colleague Helen Scott.

as to whether the negligent conduct in question was a factual cause of the harm would have to be answered by application of condition (a), rather than conditions (b) and (c), in rule (4)**.

However, further consideration of the distinction between pure and impure omissions shows it to be less useful than it may initially seem. In the first place, it is rarely the case that a person is under a legal duty to not negligently harm another by a pure (as opposed to an impure) omission. Most pure omissions, even if negligent, are therefore not wrongful. It follows that, in respect of most pure omissions, the question as to whether the omission was a factual cause of another's harm is of no practical significance: no prospect of delictual liability obtains. In the second place, impure (as opposed to pure) omissions are by no means uncommon. Indeed, the law reports are full of them. One only has to think of the long line of so-called 'municipality cases' to realise that this is so.[33] But it is not only public bodies which commit impure omissions. Private persons do so too. An obvious example would be my failure to keep a proper look-out when driving my car down the high street. In the third place, the basis for the distinction between pure and impure omissions is the connection that the former have, but the latter lack, with positive conduct. There can be no reason, therefore, to treat impure omissions as exceptional but positive conduct as ordinary. To put this another way, there may be a rational basis for dividing the set consisting of positive acts, impure omissions and pure omissions into positive acts, on the one hand, and impure and pure omissions, on the other, or into positive acts and impure omissions, on the one hand, and pure omissions, on the other. But there can be no rational basis for a division between positive acts and pure omissions, on the one hand, and impure omissions, on the other.

The foregoing has the following two implications. First, even if the line between exceptional and ordinary circumstances were to be drawn using the distinction between impure and pure omissions, there would be no reason not to treat as exceptional the two examples involving positive conduct, that is, the example of *the driver and the pedestrian* and the example of *the nurse and the patient*. Secondly, while there might then be reason to treat as ordinary the third and fourth examples, that is, the examples of *the doctor and the patient* and of *the engineer and the vintage-car owner*, there would be no reason to treat as ordinary the omission in the example below:

> *The municipality and the driver:* Municipality A digs a large hole at the side of a road under its control. A negligently omits either to put up any sign warning drivers about the hole or to fence it off. B is driving down the road on the opposite side to the hole. In order to avoid a collision with a drunk driver, B swerves to his right, and plunges into the hole. The force of the impact kills B. Had A not dug the hole at all, or had A put up a barrier solid enough to deflect a car going at the speed B was going at, B would still have been alive. However, even if A had put up a warning sign of sufficient size, and had erected a barrier sufficiently

[33] See, for example, *Halliwell v Johannesburg Municipal Council* 1912 AD 659; *Cape Town Municipality v Clohessy* 1922 AD 4; *De Villiers v Johannesburg Municipality* 1926 AD 401; *Moulang v Port Elizabeth Municipality* 1958 (2) SA 518 (A); *Pretoria City Council v De Jager* 1997 (2) SA 46 (A); *Cape Metropolitan Council v Graham* 2001 (1) SA 1197 (SCA). See also *SAR&H v Estate Saunders* 1931 AD 276; *Administrator, Cape v Preston* 1961 (3) SA 562 (A); *Silva's Fishing Corporation v Maweza* 1957 (2) SA 256 (A).

> substantial, for it to have acted without negligence, B probably would still have swerved as he had, plunged into the hole, and died of the impact.

It follows that, in respect of this example, as in respect of the examples of *the driver and the pedestrian* and *the nurse and the patient*, rule (4)** would permit, or even require, the question whether the negligent conduct was a factual cause of the harm to be answered by applying conditions (b) or (c) rather than condition (a). For reasons that need not be repeated, application of conditions (b) and (c) would yield the result that A's negligent conduct probably was a factual cause of B's harm, and therefore also the further result that A could be delictually liable to B. Those, of course, are exactly the opposite results to the ones that would be yielded by the application of rules (2) and (4) to this example.

C From the Particular to the General

Of course, not all readers of the cases – and of my analysis – will feel compelled to reach the same conclusions. At least one thoughtful respondent presented the following (possible) objection:

> The stated aim of this section is to show that it is of great practical significance whether factual causation is determined by applying rules (2) and (4) or rather by applying rules (2)* and (4)**. The section was supposed to show this by demonstrating that there are many situations in which different outcomes will be yielded, depending on whether the former or latter pair of rules is applied. So far, the section has proved that the two pairs of rules produce divergent results in respect of the examples of *the driver and the pedestrian*, *the nurse and the patient*, and *the municipality and the driver*, and possibly also in respect of the examples of *the doctor and the patient* and *the engineer and the vintage-car owner*. But it has not yet proved that the two rule-pairs produce divergent results in a large number of cases. For the examples are unusual ones, specially-tailored to make the point.

This objection is, as I have been at pains to show, *not* valid. The five examples have the following three features in common:

(1) It is true – under some description of A's conduct – that A's conduct was negligent.
(2) It is true – under some description of A's conduct – that, but for A's conduct, B's harm would not have occurred.
(3) It is *not* true – under any single description of A's conduct – both that A's conduct was negligent and that, but for A's conduct, B's harm would not have occurred.

In respect of *any* situation possessing these three features, the application of what follows the conjunctions in conditions (b) and (c) in rule (4)** will produce the opposite results to the application of rule (4). So, therefore, may opposite results be produced by the application of rule (4)** and rule (4). That depends only on whether the situation is an exceptional one – which it seems many, even most, situations are. Moreover, the number of situations possessing these three features is not a small one. On the contrary, it is very large indeed.

III The Rules the Common Law Did and Should Contain

This section has two main aims. The one, which is pursued in subsection A, is to show that, at the time of the *Lee*, the South African common law *did* in fact

contain rules (2) and (4), as the Supreme Court of Appeal had assumed, and *not* rules (2)* and (4)**, as the majority of the Constitutional Court maintained. The other, which is pursued in subsections B and C, is to show that there is no reason, of either an analytical or moral nature, to think that the South African common law would be improved if it were to replace rules (2) and (4) with (2)* and (4)**. Indeed, as subsection C shows, the opposite may well be the case. An ancillary concern, which forms the subject matter of subsection D, is to quell possible doubts about the interpretation which this article has placed on the *Lee CC* Court judgment. In particular, the concern is to justify the claim that the majority attributed rules (2)* and (4)** to the common law.

A Evidence that the Common Law Contained Rules (2) and (4)

On the face of it, the South African cases provide overwhelming support for the conclusion that, at the time of *Lee*, the common law contained rules (2) and (4) rather than rules (2)* and (4)**. Rules (2) and (4), to recall are as follows:

> (2) Negligent conduct was a factual cause of harm if and only if, but for the negligent conduct, the harm would not have occurred.
> (4) A plaintiff has proved that a defendant's negligent conduct probably was a factual cause of the plaintiff's harm if and only if the plaintiff has proved that, but for the negligent conduct, the harm probably would not have occurred.

During the four decades leading up to *Lee*, these two rules had been expressly or impliedly endorsed by a number of Appellate Division and Supreme Court of Appeal judgments – often as a part of a judgment's *ratio*. Here are some of those judgments:

> *Minister of Police v Skosana*;[34] *Siman & Co (Pty) Ltd v Barclays National Bank*;[35] *International Shipping Co v Bentley*;[36] *Groenewald v Groenewald*;[37] *Mukheiber v Raath*;[38] *Minister of Safety and Security v Van Duivenboden*;[39] *Minister van Veiligheid en Sekuriteit v Geldenhuys*;[40] *Minister of Safety and Security v Hamilton*;[41] *Minister of Safety and Security v Carmichele*;[42] *Minister of Finance v Gore*;[43] *mCubed International (Pty) Ltd v Singer*;[44] *Delphisure Group Insurance Brokers Cape (Pty) Ltd v Dippenaar*.[45]

On the basis, presumably, of this long list of cases, the *Lee CC* Court acknowledged that 'the [theory on causation] frequently employed by courts in determining

[34] *Minister of Police v Skosana* 1977 (1) SA 31 (A) at 35A–37B, 44C–45F.
[35] *Siman* (note 14 above).
[36] *International Shipping Co v Bentley* 1990 (1) SA 680 (A).
[37] *Groenewald v Groenewald* [1998] ZASCA 17, 1998 (2) SA 1106 (SCA).
[38] *Mukheiber v Raath* 1999 (3) SA 1065 (SCA).
[39] *Minister of Safety and Security v Van Duivenboden* [2002] ZASCA 79, 2002 (6) SA 431 (SCA).
[40] *Minister van Veiligheid en Sekuriteit v Geldenhuys* [2003] ZASCA 90, 2004 (1) SA 515 (SCA).
[41] *Minister of Safety and Security v Hamilton* [2003] ZASCA 98, 2004 (2) SA 216 (SCA).
[42] *Minister of Safety and Security v Carmichele* [2003] ZASCA 117, 2004 (3) SA 305 (SCA).
[43] *Minister of Finance v Gore* [2006] ZASCA 98, 2007 (1) SA 111 (SCA).
[44] *mCubed International (Pty) Ltd v Singer* [2009] ZASCA 6, 2009 (4) SA 471 (SCA).
[45] *Delphisure Group Insurance Brokers Cape (Pty) Ltd v Dippenaar* [2010] ZASCA 85, 2010 (5) SA 499 (SCA).

factual causation is the *conditio sine qua non* theory or but-for test'.[46] In other words, the majority conceded that the South African courts had on many occasions applied rules (2) and (4) in determining factual causation. Nevertheless, said the majority, 'the rule regarding the application of the [but-for] test in positive acts and omission cases is not inflexible'.[47] That is, according to the majority, the South African common law accepted that, at least in some cases, factual causation was *not* to be determined by application or rules (2) and (4).[48] Moreover, the *Lee CC* Court seemed to believe that this more 'flexible' stance had been the common-law position for some time. It went on to assert that '[t]his flexibility has a long history'.[49]

The majority in *Lee* did not think it necessary to present the long line of cases which, in its view, had adopted a 'flexible' approach to rules (2) and (4).[50] That is not altogether surprising: *for these cases do not exist*.

How, then, did the majority seek to justify its contention that the South African common law takes a 'flexible approach' to the application of rules (2) and (4)? It did so, in the first place, by citing four *dicta* to the effect that, when applying the 'but-for' test to determine whether negligent conduct was a factual cause of harm, a court is to apply 'common sense' and cannot always adhere strictly to 'logic'.[51] However, both the wording of these *dicta* and the context in which they appear make it plain that they were never meant to suggest that common sense may justify the *non*-application of rules (2) and (4), but only that common sense is to be exercised *in* the application of rules (2) and (4).The majority tried to support its claim that the common law takes a 'flexible approach' to the application of rules (2) and (4), in the second place, by claiming that the application of these rules was regarded as 'wrong or inappropriate' by the majority in *Siman*.[52] But this claim is false, and patently so. Far from regarding the application of rules (2) and (4) as mistaken, the majority judgment in *Siman* applied those rules in textbook fashion.

One of the questions in the *Siman* case was whether the allegedly negligent and wrongful conduct of a manager in the employ of the defendant had factually caused certain loss suffered by the plaintiff. The majority found that it had not. It reached that conclusion by reasoning as follows:

(1) The manager's conduct had two aspects to it: (a) a refusal (or decision not) to carry out a certain request by the plaintiff; (b) a statement of his reason for refusing – which statement contained a falsehood.[53]

(2) The plaintiff alleged that (b) was negligent and wrongful. But it never alleged that (a) was.[54]

[46] *Lee CC* (note 1 above) at 161F–G.

[47] Ibid at 162D.

[48] Ibid at 162E, 163B–D, 163F–H, 164B–C, 166A–D, 173E–G.

[49] Ibid at 163H.

[50] Ibid at 163G–H.

[51] *Lee CC* (supra note 1) at 163H–166B. The *dicta* are from *Kakamas Bestuursraad v Louw* 1960 (2) SA 202 (A) at 220B–D; *Siman* (supra note 14) at 917H–918A; *Van Duivenboden* (supra note 39) at 449E–F; *Gore* (supra note 43) at 125E–G.

[52] *Lee CC* (note 1 above) at 166B–C.

[53] *Siman* (note 14 above) at 904H.

[54] Ibid at 905B, 907D–E, 908A–B.

(3) It follows that, to determine whether the manager's allegedly negligent and wrongful conduct factually caused the plaintiff's loss, one has to 'eliminate' only (b), and not also (a), and ask whether the loss probably would still have occurred in that event. To put this in another way: one has to ask whether, but for (b), and not but for (a) or but for (a) and (b), the loss probably would still have occurred.[55]

(4) The answer to that question is that it probably would have: for (a) probably was a sufficient condition for the plaintiff to have suffered its loss.[56]

(5) Hence, the manager's allegedly negligent and wrongful conduct was not a factual cause of the plaintiff's loss.[57]

This reasoning perfectly applies rules (2) and (4). It is therefore puzzling that the majority judgment in *Lee CC* could have understood the majority in *Siman* to be rejecting them. The explanation for this confusion on the part of the majority in *Lee CC* seems to be that it misinterpreted the following passage in *Siman:*

> '[I]t is wrong or inappropriate, in applying the "but for" test to [the manager's] misstatement, to eliminate it and substitute another, different statement to the effect that the Bank was willing to procure the forward cover that afternoon … .'[58]

The majority judgment in the *Lee CC* interpreted this passage to mean that 'the majority in *Siman* considered it wrong or inappropriate to apply the substitution exercise of a hypothetical course of lawful conduct for unlawful conduct'.[59] But that is not at all what it means. The majority judgment in *Siman* had no objections to 'apply[ing] the substitution exercise of a hypothetical course of lawful conduct for unlawful conduct'. In fact, the substitution exercise of a hypothetical course of lawful conduct is precisely what it did. What the judgment insisted upon, however, was that the unlawful conduct in question had to be restricted to the misstatement, and could not include the refusal (as the plaintiff had alleged only that the former, and not that the latter, was unlawful). Hence, when substituting lawful conduct for the manager's (allegedly) unlawful conduct, one could not do so in a way which effectively removed the refusal too. But that precisely would have been the effect of substituting, for the manager's (allegedly) unlawful conduct, 'a statement to the effect that the Bank was *willing* to procure the forward cover that afternoon'.[60] As was pointed out by the majority in *Siman*, 'that hypothetical statement would be entirely contrary to or inconsistent with [the manager's] refusal'.[61]

So, neither the four *dicta* emphasising the use of common sense in the application of the 'but-for' test, nor the majority judgment in *Siman*, vindicate the *Lee CC* Court's claim that the South African common law takes a 'flexible approach' to the application of rules (2) and (4). It is possible, however, that the *Lee CC* Court believed its claim to be justified on a third ground: namely the fact that several judgments of the Appellate Division and Supreme Court of Appeal

[55] Ibid at 907D–G.
[56] Ibid at 905C–908A.
[57] Ibid at 907A.
[58] Ibid at 907E.
[59] *Lee CC* (note 1 above) at 166B–C.
[60] *Siman* (note 14 above) at 907E.
[61] Ibid at 907F.

claimed no more than that rules (2) and (4) are *generally* to be applied in order to determine factual causation. That there are such judgments is shown by the following extracts from *Skosana Siman*, and *International Shipping Co v Bentley*:[62]

> [G]*enerally* speaking (there may be exceptions …) no act, condition or omission can be regarded as a cause in fact unless it passes [the but-for] test … .[63]

> The enquiry as to factual causation *generally* results in the application of the so-called "but-for" test … .[64]

> The enquiry as to factual causation is *generally* conducted by applying the so-called "but-for" test … .[65]

However, for several reasons, the fact that these judgments make this claim does not justify the inference, drawn by the majority in *Lee*, that the common law viewed the application of rules (2) and (4) as 'flexible'. In the first place, the claim that the 'but-for' test is *generally* to be applied was, in every judgment in which it appears, *obiter* – for every one of those judgments did in fact go on to apply it. In the second place, there are a couple of judgments which expressly state the opposite. That is, they expressly say that the 'but-for' test is *always* to be applied. So, for example, in *Mukheiber*, it was stated that '[a]s far as *factual causation* is concerned, this Court follows the *conditio sine qua non* – or "but for" – test'.[66] So, too, in *Carmichele*, the Court wrote that, to the question whether harm was factually caused by negligent conduct, 'the answer has to be sought by using the "but-for" test'.[67] And, in *Gore*, it was said that '[i]n our law the time-honoured way of formulating the question [of factual causation] is in the form of the "but for" test'.[68]

There is yet another reason why the claim that the 'but-for' test is *generally* to be applied, as it appears in some Appellate Division and Supreme Court of Appeal judgments, does not justify the inference that the common law viewed the application of rules (2) and (4) as 'flexible'. The origins of this claim can be traced back to two judgments by Corbett JA (as he then was), namely his majority judgment in *Skosana* and his dissenting judgment in *Siman*.[69] In these two judgments, Corbett JA contemplated only two possible exceptions to the 'but-for' test. The one is in cases of so-called 'concurrent' or 'duplicative' causation. The other is in cases of so-called 'supervening' or 'pre-emptive' causation. This is clearly shown by the following passage in the second judgment: 'In a case such as the present one, which is uncomplicated by concurrent or supervening causes emanating from the wrongful conduct of other parties … the but-for or, *causa*

[62] In addition to these three cases, see also *Fourway Haulage SA (Pty) Ltd v SA National Roads Agency Ltd* 2009 (2) SA 150 (SCA) at 163G–H; *mCubed International* (note 44 above) at 479F.

[63] *Skosana* (note 34 above) at 35C–D.

[64] *Siman* (note 14 above) at 914H.

[65] *International Shipping Co* (note 36 above) at 700F. This sentence has been quoted, with approval, in subsequent judgments. See *Groenewald* (note 37 above) at 1113D–E; *Van Duivenboden* (note 39 above) at 449A–B; *Hamilton* (note 41 above) at 239I–240A.

[66] *Mukheiber* (note 38 above) at 1077I–1077J.

[67] *Carmichele* (note 42 above) at 327I–328A.

[68] *Gore* (note 43 above).

[69] *Skosana* (note 34 above) and *Simon* (note 14 above).

sine qua non, test is, in my opinion, an appropriate one for determining factual causation.'[70]

This has the following implication. Even if the claim that rules (2) and (4) are *generally* to be applied could have been attributed to the South African common law, because it had appeared in the *rationes* of certain judgments and had not been directly contradicted by *dicta* in certain others, that would mean no more than that the South African common law had adopted the following qualified versions of those rules:

> *Except in cases of concurrent or supervening causation*, negligent conduct was a factual cause of harm if and only if, but for the negligent conduct, the harm would not have occurred.
>
> *Except in cases of concurrent or supervening causation,* a plaintiff has proved that a defendant's negligent conduct probably was a factual cause of the plaintiff's harm if and only if the plaintiff has proved that, but for the negligent conduct, the harm probably would not have occurred.

However, for the common law to have adopted these qualified versions of rules (2) and (4) would still fall way short of its having adopted the 'flexible approach' which is envisaged by the majority judgment in *Lee*. Why? Because the 'flexible approach' envisaged by the majority judgment is meant to allow for the non-application of rules (2) and (4) also in cases not involving either concurrent or supervening causes. If that were not so, the 'flexible approach' would have been of no relevance or use to the majority judgment – for the simple reason that the judgment concerned a case involving neither concurrent causes nor supervening ones.

B The Lee CC Court's Reasons for Preferring Rules (2)* and (4)** to Rules (2) and (4)

Presumably because of its conviction that rules (2) and (4) had already been rejected by the common law and that rules (2)* and (4)**, if they had not already been adopted by the common law, were at least consistent with it, the *Lee CC* Court did little to explain why it preferred the latter to the former. However, a few passages in the judgment suggest that the majority believed that analytical and doctrinal reasons justified its preference.

One analytical reason provided by the *Lee CC* Court is that rules (2) and (4) have 'the potential to cause confusion between factual causation and negligence'.[71] But this reason does not survive scrutiny. Consider the following questions: 'Did I buy this because it is beautiful?'; 'Did I marry her because I love her?' Could it really be said that if, in order to answer the first question, I were to ask myself: 'But for its beauty, would I have bought it?', I would be confusing a causal question with an aesthetic one? Or that if, in order to answer the second question, I were to ask myself: 'But for my love for her, would I have married her?', I would be running together an enquiry into a causal relationship with an enquiry into an emotional state? Obviously not. That being so, it is puzzling why the majority in

[70] *Simon* (note 14 above) at 915A–B.
[71] *Lee CC* (note 1 above) at 162I.

Lee should have thought that if, in order to answer the question whether harm was factually caused by negligent conduct, one were to ask oneself the questions: 'But for the negligent conduct, would the harm have occurred?' or 'But for the negligent conduct, would the harm probably have occurred?', one would be at risk of muddling an enquiry into factual causation with an enquiry into negligence.

Another analytical reason provided by the *Lee CC* Court has already been highlighted in the previous section. Ostensibly, rules (2) and (4) inevitably result in the question of factual causation becoming 'a mixed question of fact and law', and therefore also in '[t]he distinction between factual and legal causation … becom[ing] unnecessarily less clear'.[72]

But this line of reasoning is nonsense.

It is true that, in order to determine whether negligent conduct was a factual cause of harm, rules (2) and (4) require one to eliminate as much of the conduct as, but no more of the conduct than, *was negligent*. It also is true that the degree to which conduct *was negligent* is an evaluative question. And it is true that, as it is understood by the South African common law, the enquiry into legal causation is in part evaluative.[73] But the evaluative question raised by the enquiry into negligence is *entirely different* from that raised by the enquiry into legal causation. The former evaluates conduct, to see whether the conduct was unreasonable from an *ex ante* perspective. The latter evaluates a causal connection between conduct (which is, or is assumed to be, unreasonable from an *ex ante* perspective) and harm or loss, to see whether – given the nature of that connection – it is reasonable to impose liability on the person who performed the conduct for that harm or loss. So even if it were the case (which, as explained the previous paragraph, it is not) that rules (2) and (4) confused factual causation and negligence, it would not, as a result, make the distinction between factual causation and legal causation any less clear.

Not only is the second of these analytical reasons invalid, but it may also have been self-contradictory for the majority in *Lee* to invoke it. As was explained in the previous section, in order to apply rule (4)** to a set of facts, one has to decide whether those facts are exceptional or ordinary: otherwise one would not know whether to apply condition (a) or rather conditions (b) and (c). As was also explained in the previous section, the majority in *Lee* seems to have supposed that, in order to make this decision, one may or even should consider what liability outcome would be the most *just*. This truly *would* make the question of factual causation a mixed question of fact and value. It really *would* do what the majority purports to be anxious to avoid, namely to 'contaminate the factual part of the causation enquiry with … normative considerations based on social and policy considerations'.[74]

So: the *Lee CC* Court's analytical reasons for preferring rules (2)* and (4)** over rules (2) and (4) are invalid, and possibly self-contradictory. It seems, however, that the majority also had a moral reason to prefer rules (2)* and (4)** over rules

[72] Ibid at 166F–G.

[73] See, for example, *International Shipping Co* (note 36 above) at 700H–I; *Smit v Abrahams* 1994 (4) SA 1 (A) at 15E–F; *Standard Chartered Bank of Canada v Nedperm Bank* 1994 (4) SA 747 (A) at 764I.

[74] *Lee CC* (note 1 above) at 167E–F.

(2) and (4). To understand that reason one has to go back to the judgment of the Supreme Court of Appeal. In particular, one has to take a closer look at how the Supreme Court of Appeal reached its conclusion that, on the evidence before it, the condition stipulated by rule (4) had not been satisfied. The two critical passages from *Lee SCA* follow:

> It is just as likely as not that [the plaintiff] was infected by a prisoner whom the prison authorities could not reasonably have known was contagious.[75]
>
> The difficulty that is faced by [the plaintiff] is that he does not know the source of his infection. Had he known its source, it is possible that he might have established a causal link between his infection and specific negligent conduct on the part of the prison authorities. Instead he has found himself cast back upon systemic omission. But, in the absence of proof that reasonable systemic adequacy would have altogether eliminated the risk of contagion, which would be a hard row to hoe, it cannot be found that but for the systemic omission he probably would not have contracted the disease.[76]

Because it is important that this line of reasoning be properly grasped, it is best to set it out again, but this time in point form:

(1) The plaintiff would have satisfied the condition stipulated by rule (4) had he been able to prove that, but for a *specific* negligent omission on the part of the defendant, he probably would not have contracted tuberculosis.
(2) But he had not been able to prove this, as he had not been able to identify the *source* of his infection – that is, he had not been able to identify the fellow prisoner from whom he had contracted the disease.
(3) Consequently, the plaintiff could only satisfy the condition stipulated by rule (4) by proving that, but for a *systemic* negligent omission, he probably would not have contracted tuberculosis.
(4) However, to prove this the plaintiff would have had to prove that reasonable systemic adequacy would have *altogether eliminated* the risk of contagion – that is, the plaintiff would have had to prove that, but for the systemic omission on the part of the defendant to provide an adequate system for the management of tuberculosis in the prison in which the plaintiff had been incarcerated, the chances of the plaintiff's contracting tuberculosis would have been nil.
(5) This the plaintiff could not do.

The majority in *Lee CC* accepted this line of reasoning as valid.[77] So did the *Lee CC* minority.[78] But neither the majority nor the minority liked the further consequence that followed if – as had been done by the Supreme Court of Appeal – rule (4) was combined with rules (1), (2) and (3). This consequence was that not only the plaintiff in *Lee*, but also others similarly situated, would be without a delictual remedy.[79] The minority expressly described this outcome as contrary to 'con-

[75] *Lee SCA* (note 2 above) at 631D.
[76] Ibid at 631E–F.
[77] *Lee CC* (note 1 above) at 163E–F, 170C–D, 171B.
[78] Ibid at 177E–179B, 181D–E, 182F–G.
[79] Strictly speaking, the plaintiff and others similarly situated would only be without an Aquilian remedy. As to the question of whether they would have a remedy under an *actio iniuriarum*, neither rules (1), (2), (3) and (4), nor the line or reasoning set out above, is of any relevance.

stitutionally tailored justice'.[80] The majority clearly thought the same.[81] Where the majority and minority judgments differed from one another, however, was in their responses to this conclusion. According to the majority, this conclusion could be avoided merely by recognising that the Supreme Court of Appeal had misconstrued the common law. Contrary to what had been assumed by the *Lee SCA* Court, the common law did not accept rule (4), nor rule (2), but rather rules (2)* and (4)**. According the minority, by contrast, the Supreme Court of Appeal had made no mistake about what the common law was.[82] It had erred in assuming that the common law was as it ought to be. It concluded that *Lee* presented the Court with a matter in which the common law required development.[83] In other words, because the outcome produced by rules (1), (2), (3) and (4) was unjust and constitutionally unacceptable in circumstances similar to those presented in *Lee*, one or more of those rules had to be, and indeed should have been, changed.

However, contrary to what was believed not only by the Supreme Court of Appeal, but also by the majority and minority in the Constitutional Court, the line of reasoning set out above is *not* valid. It is *not* valid, because its fourth premise is false. To see why it is false, consider the following two *new* hypotheticals:

> *Hospital A and patient B:* In a case between hospital A and patient B, the court finds the following: 100 patients in hospital A, including patient B, contracted disease *a*; the hospital had taken no steps to protect its patients from the disease, and that failure was negligent; had the hospital acted reasonably (ie, without negligence), it would have adopted disease-prevention scheme *aa*; had the hospital adopted disease-prevention scheme *aa*, it would not have prevented every one of the 100 patients who contracted disease *a* from contracting it, but it would have prevented 80 of them from doing so; however, it is impossible to know which of the 100 patients who contracted disease *a* would have been in that 80 and which not.
>
> *Hospital C and patient D:* In a case between hospital C and patient D, the court finds the following: 100 patients in hospital C, including patient D, contracted disease *c*; the hospital had taken no steps to protect its patients from the disease, and that failure was negligent; had the hospital acted reasonably (ie, without negligence), it would have adopted disease-prevention scheme *cc;* had the hospital adopted disease-prevention scheme *cc*, it would not have prevented every one of the 100 patients who contracted disease *c* from contracting it, but it would have prevented 51 of them from doing so; however, it is impossible to know which of the 100 patients who contracted disease *c* would have been in that 51 and which not.

Imagine that, in these two cases, the court were to apply rules (1), (2), (3) and (4). Would it be compelled to conclude that, because it was impossible for factual causation to be established, there was no possibility of delictual liability? No, it would not. In the first case, an 80 per cent probability exists that, had hospital A adopted disease-prevention scheme *aa*, patient B would not have contracted disease *a*. In second case there is a 51 per cent probability that, had hospital C adopted disease-prevention scheme *cc*, patient D would not have contracted dis-

[80] *Lee CC* (note 1 above) at 181D, 180D, 181B, 182A, 182G–H.
[81] Ibid at 170D–171B.
[82] Ibid at 178C.
[83] Ibid at 180D–E, 182A–B, 184D–E.

ease *c*. In both cases, therefore, the condition stipulated by rule (4) is satisfied: but for the negligent conduct of hospital A, the harm to B probably would not have occurred; but for the negligent conduct of hospital C, the harm to D probably would not have occurred. Since the condition stipulated by rule (4) is satisfied, so are the conditions stipulated by rules (3), (2) and (1).

The hypotheticals have in common two features which, for present purposes, deserve emphasis. The first is that there is no suggestion, in either, that the patient – B in the first case, D in the second – knows, or even could know, the *source* of his infection. The second is that, *ex hypothesi*, even if the hospitals had adopted reasonable disease-prevention schemes, they would still not thereby have *altogether eliminated* the risk that the patients would contract the diseases in question. Even if hospital A had adopted scheme *aa,* there would still have been a 20 per cent probability that B would contract disease *a*. Even if hospital C had adopted scheme *cc*, there would have remained a 49 per cent probability that D would contract disease *c*. However, and this is the critical point, neither of these features has any bearing on the question as to whether the condition stipulated by rule (4) is satisfied. Neither feature therefore precludes the conclusion that the negligent failure to provide a reasonable disease-prevention scheme by the hospital in question *was* a factual cause of the patient in question's harm. Nor therefore, does either feature preclude the possibility that the hospitals *could* be delictually liable to the patients.

The implication of this analysis for the outcome in *Lee* should be clear. The fact that the plaintiff in the case was unable to identify the *source* of his infection did *not* entail that, to prove that the defendant's negligent omission probably was a factual cause of his having contracted tuberculosis, the plaintiff had to prove that, but for that omission, the risk of his contracting the disease would have been *altogether eliminated*. As the discussion of the two hypotheticals makes plain, it would have been sufficient for the plaintiff to have provided evidence that, if the defendant had put in place a reasonable tuberculosis-management system, the general incidence of the disease in the prison during the period of his incarceration would have dropped by 51 per cent. That is, it would have been enough for him to have shown that, out of every 100 prisoners who contracted the disease in that time, 51 would not have contracted it. For, in that case, the plaintiff *would* have proved that the defendant's negligent omission probably was a factual cause of his having contracted tuberculosis. He would, therefore, have satisfied the condition in rule (4). He also, as a consequence, would have satisfied the conditions in rules (3), (2) and (1).[84]

[84] It would appear that, in *Lee*, the plaintiff did not in fact meet this burden. That is, he did not manage to show that, if the defendant had put in place a reasonable tuberculosis-management system, the general incidence of the disease in the prison would have dropped by 51 per cent. As the Supreme Court of Appeal put it: the evidence regarding the defendant's systemic failure was 'scant' (*Lee SCA* (note 2 above) at 630E–F). That finding of scantiness suggests that the Supreme Court of Appeal's conclusion, that the plaintiff had not proved that the defendant's negligence probably caused his harm, was correct, even if not for the reasons expressly stated.

The foregoing analysis eviscerates the *Lee CC* Court's moral objection to rules (2) and (4).[85] As has been explained, the objection rests on two premises. The first is that, as a matter of constitutional justice, the plaintiff in *Lee* and others similarly situated should not be denied a delictual remedy just because: (a) they cannot identify the source of their infection and (b) they cannot prove that, but for the systemic negligent omission on the part of the prison authorities, their risk of infection would have been zero.[86] The second premise is that, as a matter of logic, rule (4) – when combined with rules (1), (2) and (3) – has precisely this effect.[87] It now turns out that the second of these premises is false. Rule (4) does not have this effect because, even where (a) and (b) are the case, the condition in rule (4) can still be satisfied by proving: (c) that, but for the systemic negligent omission on the part of the prison authorities, the general risk of infection would have been reduced by 51 per cent.

C Other Reasons, For and Against

Application of rules (2) and (4) to the two hypotheticals introduced in the previous subsection has another implication, which also deserves mention, as some may find it worrisome. Consider again the first case, that is, the one involving hospital A and patient B. It is not only patient B who would succeed in establishing factual causation, but every one of the 100 patients who contracted disease *a*. For, in respect of every one of those 100 patients, there is an 80 per cent probability that, had hospital A adopted disease-prevention scheme *aa*, he or she would not have contracted disease *a*. Every one of the 100 patients would thus be able to establish that, on a balance of probabilities, the hospital's negligent failure to adopt disease-prevention scheme *aa* was a factual cause of his harm. Hence, also, every one of the 100 patients would be able to hold the hospital liable.

However, the finding upon which the probability of 80 per cent is based is that, if hospital A had adopted disease-prevention scheme *aa*, it would have prevented 80 of the 100 patients from contracting disease *a*, but would not have prevented 20 of them from contracting it in any event. It is therefore implicit in the finding upon which the 80 per cent probability is based that hospital A's negligent conduct was *not* a factual cause of the harm suffered by 20 out of the 100 patients. It follows that the application of rules (1), (2), (3) and (4) to the first hospital case yields a result that is *over-inclusive*: the hospital's failure to act as it should have factually caused harm to only 80 patients, but the application of the four rules results in the hospital's being held liable to all 100 of them. To put it another way, the result is that 20 of the 100 patients get to recover from the hospital even though they were not in fact harmed by its wrongdoing.

Consider now the second case, the one involving hospital C and patient D. Once again, every one of the 100 patients who contracted the disease will succeed

[85] The analysis is equally destructive of the minority's moral objection to the *combination* of rules (1), (2), (3) and (4). It also suggests that the minority's invocation of *Fairchild v Glenhaven Funeral Services* [2003] 1 AC 32 in support of its conclusions may have been spurious.

[86] *Lee CC* (note 1 above) at 170C–171B.

[87] Ibid at 170C–D, 171B.

in holding the hospital liable. In the case of every one of the 100 patients, there is a 51 per cent probability that, but for the hospital's failure to adopt the relevant disease-prevention scheme, he or she would not have contracted the disease in question. However, the basis for this probability is a finding according to which 49 of the 100 patients would have contracted the disease even if the hospital had put the disease-prevention scheme in place. The over-inclusivity in the second case is therefore even greater than that in the first. In the first case, 20 patients got lucky. In the second case, 49 do.

More or less the same outcomes would be true in *Lee*. Imagine that the plaintiff had managed to show that, had the defendant introduced a reasonable tuberculosis-management system, the general incidence of the disease in the prison during the period of his incarceration would have dropped by 51 per cent – but also that he had managed to show no more than that. Assume further that the number of prisoners infected by the disease during the relevant period was 1 000. It would follow, if rules (1), (2), (3) and (4) were applied, that every one of the 1 000 prisoners would have a delictual claim against the defendant. In respect of every one of the 1000, there would be a 51 per cent probability that, but for the defendant's negligence, he would not have contracted the disease. It would also follow, however, that the result yielded by the application of those rules would be *significantly over-inclusive*: 490 prisoners would get damages from the defendant, even though its wrongdoing had in fact made no difference to their contraction of the disease.

Whether the potential over-inclusivity of rules (2) and (4), when combined with rules (1) and (3), should count against them is doubtful. But that is not of present concern. For present purposes, it is enough to point out that the described over-inclusivity could not possibly have served as a reason for the *Lee* majority to object to those rules. For the rules preferred by the *Lee* majority, namely rules (2)* and (4)**, are clearly far *more over-inclusive* in their potential effect. This is shown by the following variation on the hospital-and-patient hypotheticals. Imagine that, if one of the hospitals had adopted the disease-prevention scheme that reasonableness required, 10 (but no more than that) of the 100 patients who contracted the disease in question would not have done so. And, again, suppose that it is not possible to tell who the 10 would have been. In this scenario, rules (2) and (4), when combined with rules (1) and (3), would yield the result that no one recovers. Rules (2)* and (4)**, by contrast, would yield the result that all do. In other words, the effect of rules (2)* and (4)** would be that 90 patients out of the 100 get to recover from the hospital, even though the hospital's wrongdoing had no impact on them.

The discussion so far in this and the previous subsection has aimed only to refute certain reasons *for* preferring rules (2)* and (4)** over rules (2) and (4). But I have yet to offer any argument *against* that preference. Here is one. The argument thus far presented relates solely to rule (2)*. However, with minor modification, it can be made to apply with equal force to rule (4)**. Consider the following three questions, all asked about an atheistic gay man who was recently refused appointment as a Constitutional Court judge:

(1) 'Was his being *a man* a cause of his being refused appointment as a Constitutional Court judge?'
(2) 'Was his being *a gay man* a cause of his being refused appointment as a Constitutional Court judge?'
(3) 'Was his being *an atheistic gay man* a cause of his being refused appointment as a Constitutional Court judge?'

As the word 'cause' is used in ordinary speech and, more importantly perhaps, as the concept of causation is employed in ordinary thought, there is an obvious distinction between these three questions. Moreover, because the three questions are distinct, it is possible that they do not all have the same answer. That is, it is possible for the answers to the three questions to be: 'yes', 'no' and 'no'. And it is possible for the answers to the three questions to be: 'yes', 'yes' and 'no'. Furthermore, these distinctions are not in any way value-dependent. Thus, if a person were to argue that, because that would serve some moral or political goal, a yes-answer to question (1) should be taken to entail yes-answers also to questions (2) and (3), he would be assumed – and rightly so – to have misunderstood the very concept of causation.

The same is true of the following two sets of questions, the first asked about a man who conducted himself violently and the woman who divorced him, the second asked about a man who performed unreasonable-risk-posing conduct and a woman who suffered harm:

(1) 'Was his *conduct* a cause of her divorcing him?'
(2) 'Was his *violent conduct* a cause of her divorcing him?'

(1) 'Was his *conduct* a cause of her harm?'
(2) 'Was his *unreasonable-risk-posing conduct* a cause of her harm?'

As the word 'cause' is used in ordinary speech and as the concept of causation is employed in ordinary thought, the two questions in each set – just like the three questions about the atheistic gay man who was recently refused appointment as a Constitutional Court judge – are distinct questions to which different answers are possible. That is, it is possible that the answers to the questions in the first set are: yes, his conduct was a cause of her divorcing him; but, no, his violent conduct was not a cause of her divorcing him. And it is possible that the answers to the questions in the second set are: yes, his conduct was a cause of her harm; but, no, his unreasonable-risk-posing conduct was not a cause of her harm. Moreover, once again, these distinctions are value-independent.

Now consider again rule (2)*. According to this rule, there are circumstances in which *negligent conduct* will be a factual cause of harm if, but for *the conduct*, the harm would not have occurred. This result is a function of condition (b) in the rule. Moreover, what those circumstances are depends, it seems, at least in part on *justice*. Negligent conduct is, of course, unreasonable-risk-posing conduct. This proposition can be put another way. As rule (2)* defines the concept of causation, there are circumstances in which questions (1) and (2) in the second set above are not in fact distinct, and in which a yes-answer to the former does entail a yes-answer to the latter. More than that, what those circumstances are, and thus also whether these two questions are distinct or not, seem to be partly value-

dependent. Clearly, therefore, the concept of causation in rule (2)* is at odds with the concept of causation in ordinary thought and speech.

It may be said: 'All true, but so what? Why should the fact that the concept of causation in rule (2)* diverges from that used in ordinary language and thought count against the common law's adoption of rule (2)*? After all, the common law is free to define its concepts in ways that do not accord with common usage.' Well, yes, the common law no doubt is free to do so. But it always does so at a cost. And the cost is a moral one. To the extent that the common law defines its concepts in ways that deviate from ordinary usage, it compromises an important moral, political and constitutional value. That value, one held by the Constitutional Court to be at the very heart of our constitutional project on numerous occasions, is mentioned more than once by the *Lee CC* Court: the rule of law.[88] Whatever else the rule of law demands, it demands that the law be clear to those who are subjected to it.[89] And, of course, the more the law defines its concepts in ways deviating from the ordinary understanding of those concepts, the less clear the law will become.

Rules (2)* and (4)** compromise the rule of law, not only because their concept of causation is eccentric, but also for a second reason, namely that the rules are excessively indeterminate. As was explained in the previous section, in respect of a large number of situations, contrary liability outcomes will be yielded by application of, on the one hand, condition (a) in rule (4)** and, on the other, what follows after the conjunctions in conditions (b) and (c) in that rule. As was also explained, whether (a) is to be applied, or rather (b) or (c), depends on whether the situation is ordinary or exceptional. Although this was not previously discussed, the same holds for conditions (a) and (b) in rule (2)*. The line between the ordinary and the exceptional is by its nature a vague one. Moreover, as was shown in the previous section, the majority judgment in *Lee* did almost nothing to draw it more sharply.

There may even be a third respect in which rules (2)* and (4)** compromise the rule of law. As was explained in the previous section, it is possible that the majority in *Lee* meant the line between the exceptional and the ordinary to be drawn by reference to *justice*. In other words, one is to apply condition (a) in rules (2)* and (4)** unless a more just liability outcome would be achieved by applying condition (b) in rule (2)*, or conditions (b) or (c) in rule (4)**, in which case one is to apply them. But this trades in rule-based decision making for palm-tree justice. Rather than liability being determined by application of a set of rules designed to serve justice, it is in fact being determined by applying justice (or at any rate, the decision-maker's understanding thereof) directly. To be clear, it should be noted that the difficulty here is not that rules (2)* and (4)** require causation to be determined by evaluative reasoning – the same, after all, is true of the determination of negligence, wrongfulness, and legal causation. The difficulty is the nature of the evaluative reasoning which their application requires.

[88] Ibid at 172A–B, 172H–I, 173A–B.

[89] See L Fuller *The Morality of Law* 1969, (revised ed) 63–65.

Finally, no student of South African constitutional law needs reminding that the rule of law has been recognised as a foundational value both in the text of the South African Constitution and immense body of jurisprudence built up by the Court itself over its first 20 years of existence.[90] More than that, it is in fact inextricably linked with some of the most important values protected by the South African Bill of Rights.[91] The connection is fully explained by the following three quotations from Joseph Raz's article 'The Rule of Law and Its Virtue':

> But there are more reasons for valuing the rule of law. We value the ability to choose styles and forms of life, to fix long-term goals and effectively direct one's life towards them. One's ability to do so depends on the existence of stable, secure frameworks for one's life and actions. The law can help to secure such fixed points of reference … by a policy of self-restraint designed to make the law itself a stable and safe basis for individual planning. This last aspect is the concern of the rule of law.'[92]
>
> [O]bservance of the rule of law is necessary if the law is to respect human dignity. Respecting human dignity entails treating humans as persons capable of planning and plotting their future. Thus, respecting people's dignity includes respecting their autonomy, their right to control their future.[93]
>
> The violation of the rule of law can take two forms. It may lead to uncertainty or it may lead to frustrated and disappointed expectations … The evils of uncertainty are in … restricting people's ability to plan for their future. The evils of frustrated expectations are greater. Quite apart from the concrete harm they cause they also offend dignity in expressing disrespect for people's autonomy.[94]

D Evidence that the majority in Lee really did endorse rules (2)* and (4)**

The claim that the majority in *Lee* took the South African common law to contain what this article has called 'rule (1)' and 'rule (3)' is unlikely to prove controversial. The same holds for the claim that the majority took the common

[90] See s 1*(c)* of the Constitution of the Republic of South Africa, 1996. See also *Fedsure Life Assurance Ltd v Greater Johannesburg Transitional Metropolitan Council* [1998] ZACC 17, 1999 (1) SA 374 (CC); *Ex Parte President of the Republic of South Africa: In re Pharmaceutical Manufacturers Association of South Africa* [2000] ZACC 1, 2000 (2) SA 674 (CC), 2000 (3) BCLR 241 (CC) at para 33; *Affordable Medicines Trust & Others v Minister of Health & Another* [2005] ZACC 3, 2006 (3) SA 247 (CC), 2005 (6) BCLR 529 (CC). See, on the meaning of the rule of law in s 1*(c)*, J Fowlkes 'Founding Provisions' in S Woolman and M Bishop (eds) *Constitutional Law of South Africa* (2ndEd, RS 5, 2014) Chapter 13; F Michelman 'The Rule of Law, Legality and Constitutional Supremacy' in S Woolman & M Bishop (eds) *Constitutional Law of South Africa* (2nd Ed, OS, 2006) Chapter 11.

[91] For detailed commentary on the rule of law jurisprudence built up by the Court and its relationship to various substantive provisions in the Bill of Rights, see C Hoexter *Administrative Law in South Africa* (2nd edition, 2012) 121; Michael Bishop and Alistair Price have recently developed legality or rule of law jurisprudence in sophisticated ways that take their cue from Professor Michelman's work. See M Bishop 'Rationality is Dead! Long Live Rationality! Saving Rational Basis Review' in S Woolman & D Bilchitz (eds) *Is this Seat Taken? Conversations at the Bar, the Bench and the Academy on the South African Constitution* (2012) 1; Alistair Price 'The Content and Justification of Rationality Review' in S Woolman & D Bilchitz (eds) *Is this Seat Taken? Conversations at the Bar, the Bench and the Academy on the South African Constitution* (2012) 37. See, in addition, H Botha 'The Legitimacy of Legal Orders (3): Rethinking the Rule of Law' (2001) 64 *THRHR* 523.

[92] J Raz *The Authority of Law* (1979) 220.

[93] Ibid at 221.

[94] Ibid at 222.

law *not* to contain what this article has labelled 'rule (2)' and 'rule (4)'. It is possible, however, that someone may object to this article's assertion that the majority in *Lee* accepted that the South African common law contained, or at any rate was not inconsistent with, 'rule 2*' and 'rule 4**', which are set out once more below:

(2)* Negligent conduct was a factual cause of harm if and only if:
  (a) *but for the negligent conduct*, the harm would not have occurred; or
  (b) the circumstances were exceptional and, *but for the conduct*, the harm would not have occurred.

(4)** A plaintiff has proved that a defendant's negligent conduct probably was a factual cause of the plaintiff's harm if and only if:
  (a) the plaintiff has proved that, *but for the negligent conduct, the harm probably would not have occurred*; or
  (b) the circumstances were exceptional and the plaintiff has proved that, *but for the conduct, the harm probably would not have occurred*; or
  (c) the circumstances were exceptional and the plaintiff has proved that, *but for the negligent conduct, the risk of that harm would have been reduced*.

In response to this possible objection, it should be conceded right away that working out exactly what rules the majority in *Lee* took the common law to accept, if not rules (2) and (4), is no easy task. The judgment nowhere provides a clear formulation of those rules. And its reasoning concerning the content and application of the rules is at times opaque. That said, the evidence for the majority's endorsement of condition (c) in rule (4)** is strong, as the following passages from the judgment show:[95]

> What was required, if the substitution exercise was indeed appropriate to determine factual causation, was to determine hypothetically what the responsible authorities ought to have done to prevent potential TB infection, and to ask whether that conduct had a better chance of preventing infection than the conditions which actually existed during [the defendant's] incarceration.[96]
>
> It would be enough … to satisfy probable factual causation where the evidence establishes that the plaintiff found himself in the kind of situation where the risk of contagion would have been reduced by proper systemic measures.[97]
>
> As I understand the logic of the Supreme Court of Appeal's approach, it is not possible to make this kind of inference of likely individual infection from the fact that a non-negligent system of general systemic control would generally reduce the risk of contagion. I do not agree.[98]

Admittedly, there is no equally strong evidence for the majority's endorsement of condition (b), not only in rule (4)**, but also in rule (2)*. However, the judgment clearly does endorse and apply a *second* condition for determining whether negligent conduct was a factual cause of harm, apart from the one featured in

[95] Indirect evidence for this view can be found in the minority judgment in *Lee*. The dissenting justices understood the majority judgment in the very same way. *Lee CC* (note 1 above) at 182I–183A.
[96] Ibid at 168G–H.
[97] Ibid at 169C–D.
[98] Ibid at 170B–C.

the extracts above.[99] The judgment also makes it clear that this other condition, unlike the one featured in the extracts above, does not require one to pose the counterfactual: 'but for the negligent conduct of the defendant'. Indeed, the judgment says this repeatedly:[100]

> '[I]t was not necessary for the substitution of reasonable alternative measures to determine factual causation'[101]
>
> '[O]ur law requires neither the inflexible application of a substitution exercise in the application of the but-for test'.[102]
>
> 'Our existing law does not require … the use of the substitution of notional, hypothetical lawful conduct for unlawful conduct in the application of the but-for test for factual causation.'[103]

At the same time, the judgment did not suggest that the second condition could be applied without posing any counterfactual at all. But what was that counterfactual, if it was not: 'but for the negligent conduct of the defendant'? A simple process of subtraction yields, as an alternative: 'but for *the conduct* of the defendant'. That the majority had this particular counterfactual in mind is borne out by a passage in the judgment in which it is stated that, to determine whether the defendant probably had factually caused the plaintiff to contract tuberculosis by its negligent failure to put in place certain precautions against the disease, it was acceptable simply to ask 'whether the factual conditions of [the defendant's] incarceration were a more probable cause of his tuberculosis, than that which would have been the case had he not been incarcerated in those conditions'.[104]

IV Conclusion

For the reasons discussed in sections II and III of this article, it matters whether the South African common law contains rules (2) and (4), or rather rules (2)* and (4)**. Consequently, it also matters whether the majority judgment in *Lee* actually changed the common law. It may seem that it must have done so. After all, the judgment's rejection of rules (2) and (4) and acceptance of rules (2)* and (4)** were part of its *ratio*. However, as this conclusion briefly explains, there is room for doubt.

As was shown earlier in this article, the majority judgment supposed – erroneously – that the South African common law had already rejected rules (2) and (4). And it supposed – erroneously again – that there was no inconsistency between the South African common law and rules (2)* and (4).** Moreover, on the basis of these two erroneous suppositions, the majority judgment concluded

[99] The fact that it is here called the 'second' condition must not be taken to mean that it was introduced second. It was not. The majority judgment completed its discussion of the second condition, and commenced its discussion of the first condition with the words 'Even if one accepts that the substitution approach is better suited to factual causation …'. Ibid at 168C.

[100] In addition to the extracts below, further evidence of the *Lee CC* Court's view can be found elsewhere. Ibid at 162C–E, 166A–C, 166E.

[101] Ibid at 163B–C.

[102] Ibid at 163F–G.

[103] Ibid at 166B–C.

[104] Ibid at 168B.

that, in so far as it was rejecting rules (2) and (4) and accepting rules (2)* and (4)**, it was not developing the South African common law.[105] This raises interesting questions, the answers to which are by no means self-evident:

1. Can the common law be developed by a Constitutional Court judgment which expressly disavows any intention of – and therefore also provides no constitutional reasons for – doing so?
2. Can the common law be changed by a mistaken statement in a Constitutional Court judgment as to what the common law is, as opposed to a statement (true or false) as to what – for constitutional reasons – the common law ought to be?

One might be inclined to answer to each of these questions in the negative. Some support for this conclusion can be found in FC ss 167(3) and 168(3).[106] It is possible, therefore, that rules (2)* and (4)** will be still-born. The Supreme Court of Appeal and the High Courts may push back (for the very reasons suggested in this article). Were such resistance to occur, the *Lee CC* Court's holding's risk of harm to the common law of causation, may not , in the end, cause that harm.

[105] Ibid at 173C–174A.

[106] Not all mistakes have law-changing effect. Thus, if the Constitutional Court were to hold that the Constitution requires a particular development of the common law, either to give effect to a right in terms of s 8 or to promote the spirit, purport and objects of the Bill of Rights in terms of s 39(2), as part of the *ratio* of a judgment, that holding would be binding, even if it were mistaken. However, imagine that the Constitutional Court were to reach a decision on the basis that a particular statutory provision required it, but that it had gotten the *wording* (and not merely the interpretation) of the provision wrong. Would we regard this mistake as law-changing? More specifically, would we conclude that the Constitutional Court had in fact struck down the original provision and replaced it with its mistaken version? Would lower courts thereafter be obliged to apply the provision as worded (mistakenly) by the Constitutional Court, or rather the provision as worded in the original statute? The answers, surely, are no, no, and no again – at any rate if the Constitutional Court provided no constitutional reasons for the wording it applied. Of course, what is true in respect of legislation is not necessarily true in respect of the common law. So, for example, a mistaken statement by the Supreme Court of Appeal as to what the common law is, if part of a *ratio*, *would* have precedential force (even though its effect would not necessarily be to overrule the existing common law, but perhaps only to create a conflict within the common law). In the case of mistaken statements by the Constitutional Court, at least at the time *Lee* was decided, as to what the common law is, a further complication was added by FC ss 167(3) and 168(3). These provisions seem to have presupposed a division of expertise, and thus also of authority and labour, between the Constitutional Court and the Supreme Court of Appeal. It is in large part that division which raises the two questions posed above.