Must we have a theory of proof?

ANDREW PAIZES*

University of the Witwatersrand

It is common, when writing a legal textbook, to begin with a general introductory chapter in which the theoretical foundations of the subject are expounded and explained. Chapters of this kind usually hold little of interest for those who are concerned primarily with finding practical answers to legal problems. They are, typically, ignored by them as being the province of academics who amuse themselves and each other by posing and solving riddles that stimulate the mind but have little bearing on the often prosaic world of legal practice. The question that forms the title of this chapter would, no doubt, suggest to some, this kind of exercise. ‘Who cares,’ one can imagine some people thinking, ‘about competing “theories of proof?”’

‘Evidence is lawyers’ law. We know what we want to achieve, and we make sure that our evidentiary rules deliver the results we need. There is little to be gained by exposing what we do to penetrating analysis or by trying to construct sophisticated models that rely on metaphysical distinctions, synthetic metaphors or scientific precision, since most of the things we do when it comes to legal proof take place at a visceral level where everything rests on such unscientific and non-academic notions as intuition, experience, “gut-feeling” and observations on human nature. It is of such things that the world of forensic fact-finding, the so-called real world, are made, and any attempt to contain them within a coherent scientific theory, while it may dignify the subject and lend it a certain intellectual veneer, are bound to be artificial, unhelpful and, as a result, tend toward sophistry.’

My purpose in writing this chapter is to show that such views are harmful and wrong. I hope to show that the selection of one or other ‘theory’ of proof is of the greatest practical importance and that fundamental every-day problems of proof cannot adequately be addressed without making such a selection. I propose to indicate, too, what the implications of the chief competing theories are, and, finally, to put forward some ideas on how we might go about determining which of the theories might best serve us in particular cases.

I THE FIRST PROBLEM: IS THE COURTROOM LIKE A CASINO?

The unlikely starting point for this exploration of theories of proof is the casino. Imagine that you wish to play a game of roulette. Imagine a

*BCom LLB PhD (Witwatersrand); Professor of Law, University of the Witwatersrand, Johannesburg.
roulette wheel of 36 numbers, 18 of which are marked in red, the remaining 18 in black; 18 of which are odd, the remaining 18 even. If you were to place a single token on a bet that the spin of the wheel will yield, say, an odd number, it is accepted that your chance of success is one in two, or, to express this in the language of probability theory, 0.5. A bet on a black number would have the same chance of success. A bet that combined these outcomes, one that, in other words, is taken on an odd and black number resulting, would yield a probability of 0.25 (or one in four), a ratio that is the product of the other two ratios.

That it is entirely appropriate to employ this ‘product rule’ in cases of this kind is plain. It is a simple application of the classical or mathematical system of probability that has come to be referred to as ‘Pascalian’, after Blaise Pascal, who developed the early analysis of the probabilities involved in gambling. But what is much more controversial is whether it is appropriate to engage in this kind of reasoning in the legal context.

It should be emphasised that the inquiry is not a synthetic one. One is frequently required, in law, to consider the probability of the truth of an entire series of propositions taken together where none of the prepositions may be considered, in itself, to be certain. To take but one example: X disappears in mysterious circumstances while hunting in a remote area of bushveld. He was in the company of Y who returns with an implausible account of the disappearance. Y had a motive to kill X, who has not re-appeared in the intervening two years, and Y comes back with some of X’s valuable personal possessions giving, again, an implausible account as to how they came into his possession. The area in which X disappeared is, however, notoriously dangerous, with countless natural and other hazards, including wild animals, marshes, rivers, floods, extreme weather conditions, and human occupants hostile to tourists. The propositions that would have to be established by the prosecution in a case of murder against Y would include the following:

1. X is dead;
2. Y killed X;
3. The killing was unlawful (and, therefore, carried out in the absence of defences such as self-defence, compulsion, and necessity); and
4. the killing was intentional.

A probability ratio may be allocated to each of these propositions, none of which may be taken to be certain, each ratio being determined, however, on the assumption that the preceding propositions are certain.

1 1623–1662.
2 This example is taken from a murder trial that took place some 15 years ago in Botswana. It is, to my knowledge, unreported.
Of these propositions the most likely is, of course, (1), and the least likely (2), followed, probably, by (4). If we, for the purpose of argument, allocate numerical values to these ratios (say, 0.9, 0.7, 0.9 and 0.8 respectively, the following question arises: in order to establish the probability of all four propositions being true, is it appropriate to view it as being equivalent to the product of the probability ratios of each of the four propositions (which would be approximately 0.45)?

Before I attempt to answer this question, it should be pointed out that the problem does not arise merely because we have allocated numbers to these statements of probability. It would seem to arise with equal force if we were to use a taxonomy that employed such descriptions as, say close to certain, well beyond a reasonable doubt, only just beyond a reasonable doubt, falling just short of being beyond a reasonable doubt, more probable than not, and so on.

And so to the real problem. Multiplication quickly and significantly expands doubt, as can be seen in the example above, where the final ratio falls below 0.5 and cannot be said to have satisfied even the civil standard of proof, which requires a preponderance of probabilities, let alone the criminal standard of proof beyond a reasonable doubt. Even individual ratios, each of 0.9, fall short of the civil standard if taken in sets of seven or more. To take this observation a little further would mean this: that if a criminal conviction (or even a civil judgment) depended on the truth of each of a great enough number of propositions, each established well beyond a reasonable doubt but none quite certain, no such conviction (or judgment) would be possible.

Can this be so? Law students are brought up on the axiom that the prosecution is required, in a criminal case, to prove every element of liability beyond a reasonable doubt, and this is what our courts seem to require in practice. A court will examine each element separately, test it against the standard of proof beyond a reasonable doubt, and, if satisfied that it meets this standard, move on to the next element where it will adopt the same approach. It is an approach, however, that offends against the tenets of Pascalian reasoning, which would require the probability ratio of the ultimate issue (the accused’s guilt) to be expressed as a product of the probability ratios of each of its constituent elements of liability, so that a sufficient number of elements each established on the basis that they only just exclude a reasonable doubt would be unable to sustain a conviction.

As it happens, there is an alternative system of probability that would seem to describe more accurately what our courts seem to do in practice, one which has come to be called ‘Baconian’ after Francis Bacon.\(^3\)

\(^3\) 1561–1626.
grappled with a more inductive system that has, in recent times, been championed by Jonathan Cohen. In a Baconian system, the product rule does not apply to cases of this kind. Instead, Cohen would say that the chance of two or more events having occurred would be equal to the lowest of the various probabilities put forward. Baconians are averse to the use of numerical values, but it is clear, at least, that they would regard an issue as having been established beyond a reasonable doubt where that issue depended on the truth of a number of assertions, each of which has, in turn, been established beyond a reasonable doubt.

It is plain, however, that our courts have never expressly adopted the Baconian in preference to the Pascalian system. They have, to my knowledge, never even alluded to the fact that such a choice exists. Any defence of the former would, therefore, have to convince one of the fact that the courtroom is, in fact, unlike a casino in the sense that the Pascalian rules of probability, appropriate though they may be in the context of gambling odds, do not describe the way we ought to reason about forensic fact-finding.

A second situation may be considered in order to demonstrate, first, that the problem already identified arises frequently in practical situations, secondly, the lack of attention it has received by the courts and, thirdly, how crucial the choice of theory may be to the outcome of a case. It is the problem of what is sometimes called ‘cascading inferences’, or ‘inferences upon inferences’. Consider this example: X’s body has been found some time after his death. The cause of death is not certain, but some traces of a strong poison, used in a commercial brand of weed-killer have been found in the body and the condition of the body is consistent with that kind of poisoning. W, a salesman, testifies that some ten days before X’s death he sold a large quantity of the weed-killer to someone answering to the description of Y, a person who had a motive to kill X. The evidence, then, is used to establish what is called an ‘intermediate’ fact – that Y bought the poison shortly before X’s death. This is established, inferentially, from W’s evidence. It is used, in turn, to sustain a further inference, that Y somehow applied the poison to X, and, then, the further inference that this poison was the cause of X’s death. The question I wish to consider is this: what level of proof must be achieved in respect of the

---


5 According to Cohen ‘The logic of proof’ (n 4) 93, it can be shown ‘that where, on given evidence, two propositions both fall short of certainty, their conjoint assertion has just the same Baconian probability as that of either proposition separately, if they are equally probable, or as that of the less probable one if their probabilities are unequal’. As he puts it (at 97), ‘in terms of Baconian probabilities, the chain . . . is as weak as, but no weaker than, its weakest link’.  
intermediate fact, the fact, that is, that it was Y who bought the poison from W?  

A Pascalian would answer this question by saying that this would depend on a number of things. One would, first, have to identify the quantum of proof needed to sustain Y’s conviction on a charge of murdering X which is, in our law, proof beyond a reasonable doubt, a level very difficult to quantify but represented, for the purpose of this argument by the symbol ‘Q’. One would then have to quantify, first, the probability that Y applied the poison to X given that he bought the poison from W, and, second, the probability that the poison caused X’s death given that Y applied it to him. If we identify these ratios as, say, 0.8 and 0.9, we can answer the question and find the probability ratio necessary for the proof of the primary fact, which I shall call P, by using the product rule, so that P × 0.8 × 0.9 = Q, and P = Q/0.72, a number considerably larger than Q itself.

A Baconian would take a completely different view of the problem. Under that system a primary fact that is necessary to sustain a conviction would, quite simply, have to be proved beyond a reasonable doubt, since anything less than this would not allow for the proof of the ultimate issue on the same level of proof, and anything more than this would, simply, be unnecessary.

The Pascalian value for P, then, is dependent on the values of the other variables in the given equation, with the result that, while it may never be less than the normal criminal standard, it may well require proof on an even higher level. The Baconian value for P, however, remains constant and is unaffected by the strength of the inferential leaps that are taken after the proof of the primary fact. For a Baconian, in short, P = Q.

Which of these approaches, then, is taken by our courts? Once again, nothing has, to my knowledge, expressly been said by our judges on the question. Even in jurisdictions such as Australia, where there has been much judicial attention paid to the proof of primary and intermediate facts, this particular issue has not been prominent. The focus, there, has been on distinguishing between modes of reasoning akin to the ‘links of a chain’, where the proof of a particular fact is essential to the proof of an ultimate issue, and modes of reasoning akin to the strands of a rope, where this is not so but where the primary or intermediate facts are, in the aggregate, sufficient in number and weight to allow for the proof of the

---

6 It is this that distinguishes the present situation from the previous one relating to the disappearance in the bushveld. Even though the first case involved chains of inferential reasoning, each step related to an element of liability. In the present case, one of the steps, the first, relates to a fact (the purchase of the poison by the accused) that does not constitute an element of liability as such, but goes to proving elements of liability. It is, in other words, truly an ‘intermediate’ fact.
ultimate issue. As regards the latter mode of reasoning, it is correctly accepted that proof at a level lower than that of the criminal standard may well be sufficient, even in a criminal case. As regards the former mode of reasoning, however, while it is correctly said that nothing less then the criminal standard will do, the question that is not asked is whether something more might not, in certain cases, be required.

In the given context, it is my impression that the courts in South Africa incline intuitively towards the Baconian model. But the question has once more to be asked: if this approach only obtains if the rules of the courtroom are not those of the casino when it comes to reasoning about probability, to what, if anything, are we to attribute this differential treatment?

1. Excursus on probability: weather maps, prison yards and gatecrashers

In attempting to distinguish between modes of reasoning that may be employed by gamblers and judges, it may be instructive to begin with a simple form of probability assessment familiar to the layperson. A weather forecaster predicts a 70% chance of rain over a certain area of the country. What is he saying? Is it that

1. rain will fall over 70% of that area;
2. seven out of every 10 towns in that area will have rain;
3. he is confident that the entire area will have rain;
4. he is almost certain that at least the greater portion of that area will have rain; or
5. he believes some part, at least, of that area will certainly have some rain?

One could, of course, add more possible outcomes to the above list, but these are enough to raise what I believe to be an important question. Mathematics is, of course, a language, one that operates within its own context of abstract notions. Once we move away from that context and attempt to apply that language within the world of human experience so that the language of mathematics is used in conjunction with other languages, is there not the danger that what works within the world of abstract ideas either cannot be translated at all or can be translated only with great difficulty and at some cost to the accuracy of description into the language of everyday discourse? After all, we all know what is meant by a 70% probability of the happening of an event; we know it in a pure


8 This is, too, the South African position: see R v Membu 1950 (1) SA 670 (A) at 679–80; R v Manda 1951 (3) SA 158 (A) and R v Sibanda 1965 (4) SA 241 (SRAD).
mathematical sense, and when the event is a non-complex one that can be described in non-mathematical language that is simple, unambiguous and clear, the two languages merge reasonably happily with few complications. Whether a red number or an even number will come up on the spin of a roulette wheel is such an event – it can be taken from our material world to enliven a mathematical idea and to render that idea susceptible of practical application. Whether it will rain over a certain area, however, is less easily accommodated within such a marriage. To convert a confident prediction of rain over a certain area into the mathematical language of probability in a way that is meaningful within both languages is rather more difficult.

When it comes to the courtroom, the task is more difficult still. Consider the well-known hypothetical case of the prison yard: 9 Of a group of 100 prisoners in a prison yard, all but one (who storms off in protest at what the others have decided to do) attack and kill a prison guard, Y. The prisoners are identically dressed and cannot be distinguished by the surviving guard who witnessed the event and testifies to the attack. Assume that one of the 100 prisoners, say X, is chosen at random and charged with the murder of Y. Should he be convicted? Could one not say that since, statistically, the chance that X is the innocent prisoner is only one in a hundred, there is a 99% probability of his guilt – a figure that may be raised to an even higher number if we raise the number to say 1 000 prisoners or even, since this is a hypothetical situation, 1 000 000 – and that this represents a far higher figure than the probability that can ever be generated by, say, eyewitness testimony on the strength of which accused persons are convicted every day by our courts?

To take another example: 10 assume that it is known that, of a crowd of 1 000 spectators that attended a football match, 501 were gatecrashers who gained entry to the ground without paying the admission fee. If the owners of the stadium sued one of the spectators, Z, again selected at random, for the admission fee, would they succeed on the ground that it was more probable than not that he or she was a gatecrasher? Would they be more likely to succeed if the number of people known to be gatecrashers was, say, 800 or, even 999?

What these examples point to, in my opinion, is the danger of slipping from one language to another without realising that subtle differences in linguistic nuance disguise significant contextual differences in meaning. It is one thing to say, in the prison yard case, that, in randomly selecting X,
we have a one in a hundred likelihood of chancing upon an innocent man (since, if we undertake this task 100 times, it is certain that we will have selected one innocent man). But it is quite another thing to say that the probability that X is guilty of murder is 99%. What, after all, do we have by way of evidence to link this particular person with the crime in question? Nothing apart from the cold statistic that he was one of a group of 100, 99 of whom committed that crime. To say that this establishes a 99% probability of his guilt in respect of that crime is to betray a singular ignorance of the foundations upon which our legal system in general, and our procedural jurisprudence in particular, operates. When we conduct criminal trials, we are not playing a game — either for our amusement or our financial enrichment. We are not assessing the odds in order to determine whether we should be placing or accepting a bet, as a gambler or casino operator would do in a game of roulette. Neither are we engaging in a purely abstract exercise where the numbers exist in a world of their own and, for the people who imagine and manipulate them, as an end in itself. We are, rather, participating in a very human activity in which the issue is whether or not to impose on one of our own the highest mark of opprobrium conceived by our society. In order to do so in a manner that will earn the respect and deference of that society, we have designed a body of evidentiary and other procedural rules cast in one or other non-mathematical language that the members of our society are able to understand. In so far as we borrow from the language of mathematics to describe the process of proof, we do so once the evidence has been adduced as a metaphor to describe the extent to which we have been persuaded by the evidence. But I doubt whether it can ever actually be the evidence when there is no other evidence. I do not believe it is appropriate to translate the prison yard situation into the mathematical statement ‘the probability that X is guilty is 99%’. Neither do I believe it to be appropriate to translate the gatecrasher case into the mathematical statement ‘it is more probable than not that Z was a gatecrasher’.

One has, I believe, to ask oneself what the purpose is, within a human context, of relying on an assessment of probability. The simpler the exercise, and the closer it appears to be to the abstract world of numbers and ideas as an end in itself, the more able we are to use the mathematical notion and the more appropriate the translation of that notion becomes to the world of human affairs. The more complex the exercise, and the further it strays from that world into the quite different (and ostensibly much messier and more textured) world of human and social experience, on the other hand, the less useful and the more dangerous the translation becomes.

In order to test this idea, consider a situation that combines the world of the casino with that of the courtroom: A plays a single game of roulette
in a casino where the wheel has 18 red numbers, 18 black numbers and two green numbers (being 0 and 00). Assume that one of the bets open to him is that the winning number is, say ‘not red’, which would include both the black and the green numbers. Assume, too, that this is the bet that A makes, placing a sum of R10 000 on this outcome. And assume, finally, that the state has imposed a tax on all gambling to the extent of 10% of all winnings. Would the receiver of revenue, in a civil action against A, succeed in claiming the sum of R1 000 (being 10% of the winnings, which, let us assume for this purpose, would, if earned, have amounted to R10 000) if no other evidence at all were adduced in support of a contention that A has played, won, and failed to pay his tax?

A Pascalian might maintain that the chance of A’s winning would be 20 in 38, which makes it clearly more probable than not that one would win if one took this bet and, therefore, that A did win in this particular case. It would seem to follow a fortiori, therefore, that this would suffice for a successful action, since A’s civil liability in this regard seems to be wholly dependent on the probability of his having won the bet.

It is here, however, that the difference between the two worlds must assert itself. It is one thing to say that one has a greater than 50% chance of winning if one places the bet in question. A statement such as this is entirely valid in a purely abstract mathematical sense and, even, if applied outside of that realm as a guide to gamblers in their betting activities. But it is an entirely different thing to say that this means that A must be found, in law, to have won the bet in question. A statement such as this has no regard to either the evidentiary and procedural rules with which we endow our legal system or the broader socio-political role played by that system within a human society. Just as it would offend us to convict the prisoner in the prison yard case (to do so would, of course, justify the conviction of all 100 prisoners, creating the certainty of at least one wrong conviction), so would it offend us to find for the state where no evidence, in the sense in which we understand that word, has been adduced of his having won the bet. How bizarre it would be to dismiss the state’s case if the bet was on ‘black’, only to allow it if the bet was on ‘not red’, with the mere existence of the two green numbers being the ultimate clincher. This is not the way our legal world does or should work. The statement of probability is not, therefore, a valid translation from the mathematical (and gambling) world to which it properly belongs into the language of the law.

We have, of course, to be very sure of ourselves on this point, since our approach, if we now embrace Baconian reasoning, is one that leads us to convict in situations where Pascalian logic tells us that we should not. To return to the example of the poison, for instance: if we are satisfied beyond a reasonable doubt, but only just so, that the accused (a) bought
the poison; (b) assuming he did so, applied the poison to the deceased; and (c) assuming (a) and (b) to be true, thereby caused the deceased’s death, our position would be that we are entitled to convict the accused of murder even though Pascalian reasoning would expand the amount of doubt as to the accused’s guilt to such a degree that it would, at the very least, amount to a reasonable doubt. Are we confident enough about our sense of how the law does and should work to get past the powerful presence that is the product rule?

If we examine more closely the two hypothetical situations considered above (the ‘bushveld’ case and the ‘poison’ case), it seems to me that one may take one of at least three different positions. The first is to maintain that there is, in principle, a difference between the two cases, even though both are instances of cascading inferences, in that the first concerns inferences which, at every step, yield conclusions which are themselves elements of liability whereas the second is concerned, first, with an inference that yields an intermediate fact – the fact, that is, that the accused bought the poison in question from the witness – before sustaining other inferences that constitute elements of the crime of murder. There is a respectable line of thought, led by Cohen himself, that considers chains of inference to have in common with physical chains the fact that weaknesses in the links are less damaging to the final result the further down the ‘ladder’ they happen to be. It is a view that maintains that relatively weaker inferences drawn from facts firmly established count for more than relatively stronger inferences drawn from rather flimsier ‘proven’ facts. A purely mathematical perspective, favoured by writers such as Schum, on the other hand, would hold that there is no rational foundation for such a distinction. Lawyers who find Cohen’s view more appealing might, I suppose, defend it by arguing that elements of liability are, in a way, destinations or end-points in a criminal case. It is true that the guilt of the accused rests on the need for the state to prove all of these elements, but once a single element has been proved beyond a reasonable doubt, there is a powerful sense that any rule that would require such a conclusion to be revisited in order for a determination to be made of how much the amount of doubt concerning the proof of this particular element affects the amount of doubt there may be, finally, as to the accused’s guilt, would be so foreign to our practice as to distort what we really mean by proof beyond a reasonable doubt. Intermediate facts, on the other hand, as the name suggests, are not end-points. There may,

11 See Cohen, The Probable and the Provable (n 4) §§ 22–3, 72 and 73. Cohen claims that to ‘indulge in a long chain of inferences is rather like lowering yourself out of an upstairs window by a real chain’, since you ‘have less distance to fall if the links at the lower end of the chain are weak than if those at the upper end are’.

then, be more reason to revisit such a fact, proved beyond a reasonable doubt, and determine that an even higher level of proof is necessary to satisfy, ultimately, the criminal standard of proof beyond a reasonable doubt.

The first position, then, is a Pascalian one in so far as it concerns intermediate facts, but it concedes to the Baconian when it comes to elements of liability. The second position is, on the other hand, uncompromisingly Pascalian. It finds unacceptable the fact that Baconian concessions are made in respect of facts that are, in an analytical sense, no less intermediate than any other facts which adherents to this position would recognise as being such and which are, moreover, the most significant facts of all, since their proof establishes nothing less than an element of liability.

A third position, though, is possible, one that is purely Baconian in its conception and style and which takes the Baconian arguments to their logical conclusion. It attacks the very starting-point of the Pascalian argument and claims that the error in that argument stems from attempting to convert the criminal standard of proof into a notion susceptible of mathematical expression. If it be accepted that this standard does not describe a point on a linear graph, located between 0.5 and 1.0 in the notional realm of proof but describes, instead, a state of mind that is, although impossible to define or describe with complete accuracy, one that lawyers are able to understand and to share, then the foundations of the Pascalian model crumble. Although all would agree that it is not possible to identify a specific point in the range of proof as describing the criminal standard, the fault lies in even making an assumption that any such point is at all capable – even for analytical or hypothetical purposes – of existing. It is, quite simply, not a ‘point’ at all. Neither is it something that can, in any way, be formulated in the kind of language that allows mathematicians to go to work on it for the purpose of devising models such as those relied on by the Pascalians. Seen in this way, it is difficult to translate the criminal standard into any non-legal language at all – let alone a mathematical one. It is a standard which, once attained, is attained for all practical purposes. It serves, rather like the speed of light in theoretical physics, as both a constant and an upper ceiling, and any attempt to go beyond it is both unnecessary and pointless. It carries with it the judgment of legally trained people that it is safe to accept what is sought to be proved for any legal purpose – even for imposing the sternest penalties that our system is capable of imposing. Once a fact, any fact, has been proved at this level, it is safe to accept it without question and carry on with any other outstanding business, since that fact has cleared the greatest obstacle our law chooses to put in its way.
It is the third of these positions toward which our legal practice, in my view, seems to incline. It is a strong Baconian position and it asserts that the difficulty of translating this legal practice into mathematical language is so great that the temptation to do so has to be strenuously resisted. Once the imperfect translation of the criminal standard into mathematical number-lines has been effected, it is too late to stop conclusions being drawn that are utterly inconsistent with our legal practice.

II THE SECOND PROBLEM: IS A CIVIL CASE A TRIAL OF CASE STRENGTH OR IS IT A DIVISION OF CASE MERIT?

The second problem is as interesting as the first. It is also at least as important from a practical point of view. To illustrate it, a return to the casino is necessary. Assume, again, a roulette wheel with numbers of only two colours, black and red. This time, however, assume that the number of red numbers is not equal to the number of black numbers, that there are \( x \) numbers in total, the number of red numbers being \( y \) and the number of black numbers \( z \), where \( y \) is not equal to \( z \). The chance of a spin of the wheel yielding a red number will, therefore, be \( \frac{y}{x} \), while the chance of a black number winning will be \( \frac{z}{x} \). The important thing, for now, however, is that the sum of these two ratios will always amount to 1, since no other outcome is possible, so that there is what we might call a ‘complementarity’ of proof.

The question we have to consider is whether such a complementarity obtains as much in the courtroom as it does in the casino. Imagine a simple civil case in which the plaintiff, P, sues the defendant, D, for damages arising out of an alleged assault. D denies having assaulted P. There are no witnesses apart from the two litigants, both of whom testify and give contradictory accounts on the crucial question of whether the argument which D admits having had with P led to the alleged assault. The court, let us assume, makes an assessment of the credibility of the two witnesses, concluding that it is more inclined to believe, say P than D. The court, however, is not favourably impressed with either of the witnesses, since both accounts are contradictory, unclear, unconvincing and laced with errors. The court would not, of course, in a case of this kind, be inclined to quantify these notions, but if it did, and if it were of the view that there was, say, a probability of \( X \) that P was telling the truth and only a lesser probability of \( Y \) that D was truthful, how should the court decide the matter?

One’s first reaction might be to say that, once it is clear that \( X \) is greater than \( Y \), the matter is easily resolved in terms of the civil onus of proof, which requires proof on a preponderance of the probabilities, with the result that P even if he or she bears the onus, must win. But the problem is rather more difficult than that. It is clear, first, that there would seem to be nothing in principle to compel complementarity between \( X \) and \( Y \). They
need not, in other words, add up to 1, and it will, indeed, be entirely fortuitous if they do. Secondly, it may well be that $X$, even though greater than $Y$, is not greater than 0.5. What do we do if, say, $X = 0.4$ and $Y = 0.2$?

We could do one of two things. We could, on the one hand, take the view that, although 0.4 is greater than 0.2, it is not greater than 0.6, a figure that represents the probability of P’s version not being true, so that P fails on the basis that he has not proved his version to be ‘more probable than not’. Such an approach rests on a requirement of complementarity, since the probability that a proposition is true and the probability that it is false must always add up to 1. It is an approach, therefore, that weighs the probability of the truth of a proposition against its mathematically conceived converse rather than against the probability of the truth of a version that happens to be put forward in competition to it in the real world of civil litigation.

We could, on the other hand, relinquish the notion of complementarity and decide that 0.4, although it does not constitute an abstract ‘probability’, is good enough to defeat 0.2, which happens to be its only adversary, so that the plaintiff wins despite failing to prove that his case is more probable than not. Such an approach would see civil litigation as a trial of case strength rather than a division of case merit. The Pascalian approach, according to which there is a determined or fixed amount of ‘case weight’ which is allocated to the plaintiff and the defendant, is attacked by Baconians as failing to reflect the realities of our legal system. As Cohen puts it, ‘if certain premisses afford zero Baconian probability for Q, this does not mean that they imply the certainty of not-Q, as would zero Pascalian probability. Instead it means either that the premisses do not state any relevant evidence at all or that, on balance, they give at least some measure of support to not-Q’. He concludes then, that,

‘even if the “posterior” Baconian probability of Q on P is greater than zero, the “prior” Baconian probability of Q – the standing of Q irrespective of premisses – may still be zero, because without the premisses there may be no reason at all to prefer Q to not-Q’.

In short, as Cohen observes, the ‘scale of Baconian probability runs downwards from provability to non-provability, not as in the Parcalian scale from provability to disprovability’.

13 This situation could not, of course, arise once $X$ exceeds 0.5.
14 See Schum (n 12)467.
15 ‘The logic of proof’ (n 4) 93.
In South Africa, the leading texts\textsuperscript{16} and some judicial authority\textsuperscript{17} give formal support to and recognition of the Pascalian approach. It was clearly articulated in the following pronouncement of Lord Denning in \textit{Miller v Minister of Pensions}.\textsuperscript{18}

‘If the evidence is such that the tribunal can say “we think it more probable than not”, the burden is discharged, but if the probabilities are equal it is not.’

Eggleston,\textsuperscript{19} after examining the cases in Australia, too, finds ‘plenty of cases’ in which the test of ‘more probable than not’ has been applied,\textsuperscript{20} and none in which a judge has said something like this:

‘The plaintiff’s evidence did not establish the case at the level of more probable than not, but his account was more probable than the defendants’s and I will decide in his favour.’\textsuperscript{21}

For all this, however, it is difficult to shake off a persistent doubt as to whether this is what our courts do in practice. Most practitioners will tell you that the test in civil cases is quite clear in practice, and that everything hinges on one simple question: which of the two versions is more probable? The more subtle approach of weighing a proposition against its notional converse appears, to the practitioner, synthetic and even a little ridiculous, belonging more to the etherised world of abstract inquiry than to the adversarial contest that is civil litigation.

But if the Baconian approach squares more with the practice of our courts than its Pascalian counterpart, would one not expect to find traces of it in the judgments of our courts? One is, after all, dealing here with a crucial choice which must, one would suspect, affect the outcome of countless civil cases. If our courts really do apply the civil onus the way many practitioners might consider they do rather than the way the courts profess they do, there must be some instances where the outcome may be supported by Baconian but not Pascalian reasoning.

I resolved, with this idea in mind, to look for some evidence of this disparity between theory and practice. At first I achieved little success. The really big ideas in law, it seems, receive little formal treatment – the more foundational a notion is (consider, for instance, the question of

\begin{itemize}
\item \textsuperscript{17} See \textit{Ocean Accident and Guarantee Corp Ltd v Koch} 1963 (4) SA 147 (A) 157D. See, further, \textit{S v Siyavule} 1980 (4) SA 257 (B) at 260F–G and \textit{Selamolele v Makhado} 1988 (2) SA 372 (V) 375–6.
\item \textsuperscript{18} [1979] 72 All ER 372 at 374.
\item \textsuperscript{19} (1991) 65 \textit{Australian LJ} 130.
\item \textsuperscript{20} Eggleston (n 19) 135.
\item \textsuperscript{21} \textit{Ibid.} See, for instance, \textit{R v Jenkins: Ex parte Morrison} (1949) 80 CLR 626 at 642–3, where Dixon J clearly favoured the Pascalian over the Baconian approach.
\end{itemize}
unlawfulness in criminal law and delict) the less there seems to be said about its precise meaning and ambit. It is almost as if our powers of analysis take us only as far as a particular point, beyond which we become increasingly and even solely reliant upon those human qualities that resist empirical expression or scientific description – qualities such as judgment, intuition, and human experience. I had reconciled myself, therefore, to the fact that I was unlikely to find any articulated rules or principles that would betray a conflict between Pascalian form and what I suspected was Baconian substance, when the proof for which I was searching came to light in a somewhat unexpected quarter.

It was held in the important case of *R v Blom*22 that there are, when it comes to reasoning by inference, two ‘cardinal rules of logic’ which may never be ignored. The first is that the inference sought to be drawn must be consistent with all the proved facts. The second is that the facts must be such that they exclude all other reasonable inferences save the one sought to be drawn. For some years, the rules in *Blom*, even though they found expression in a criminal case, were applied equally to civil cases.23 The view was then taken that the second rule, which some see as a restatement of the criminal onus of proof,24 could not be applied to civil proceedings, and a variation of that rule emerged so as to adapt it to civil cases. For this purpose, the inference could be drawn only if it was the ‘most natural and plausible inference’.25 The significant thing about this test for our purposes is that an inference, say A, may be more ‘natural and plausible’, then inferences B, C and D without being, at the same time, more likely than ‘not A’. And where the drawing of an inference leads to the proof of a fact, it would appear that it is possible to reach significantly different results, depending, quite arbitrarily, on the route one happens to choose. If A is more natural and plausible than B, which is the inference argued for by the defendant, but is not more probable than ‘not A’, then a court applying the Pascalian civil standard as it is defined by our courts would find that the plaintiff would have failed to discharge the onus. A court that approaches the matter as one of inferential reasoning, however, would

22 1939 AD 188 at 202–3.
23 See, for instance, *Basson v Basson* 1943 OPD 204 at 213; *Friedman v Harrismith Municipality* 1945 OPD 212 at 226 and *Kanel v Fiddel* 1948 (4) SA 466 (C) at 472.
24 The views for and against this proposition are considered by his Lordship H C Nicholas in ‘The two cardinal rules of logic in *Rex v Blom*’ in E Kahn (ed) *Fiat Justitia: Essays in Memory of Oliver Denys Scheinrein* (1983) 312 at 319.
25 See, for instance, *Ocean Accident Guarantee Corporation* (n 17) at 159B–D; *Govan v Skidmore* 1952 (1) SA 732 (N) at 734; *AA Onderlinge Assurance Assosiasie Bpk v De Beer* 1982 (2) SA 603 (A) at 614–5; *Taylor NO v National Mutual Life Association of Australia Ltd* 1988 (4) SA 341 (E) at 346H–J; *Asumnestaal CC v South African Eagle Insurance Co Ltd* 1992 (2) SA 662 (C) at 668D–E; *Parenci Committee of Namibia v Nujoma* 1990 (1) SA 873 (SWA) at 887C–E; *Hans v Minister van Wet en Orde* 1995 (12) BCLR 1693 (C) at 1699A–C; *Macleod v Rens* 1997 (3) SA 1039 (E) at 1047A–C and *Santam Bpk v Poiglieter* 1997 (3) SA 415 (O) at 422–3.
find that the inference is one that may safely be drawn and would find the ensuing 'fact' relied on by the plaintiff to have been properly established. An examination of those cases in which this version of the second leg of the test has been applied reveals that the courts, when they follow the second route, do not allude at all to the first and, what is more, show no awareness of the fact that the two routes may, in many situations, lead to different results. It is highly probable, therefore, that results consistent with the Baconian approach and inconsistent with the Pascalian approach have frequently (and, probably, unwittingly) been reached by our courts. There may, in other words, be a strong Baconian overlay to a test that, on the face of it, embodies Pascalian reasoning, with the result that one might find it difficult to predict which of these two theories will prevail in a given case.

And so to the next question: given that the Pascalian and Baconian approaches must, in situations that will in practice be neither uncommon nor artificial, lead to opposite outcomes within the same trial, which approaches should the courts choose? The choice we face may, perhaps, be better understood if we consider the practical context within which it has to be made. Consider, then, the hypothetical example I raised earlier concerning the lost tourist in the bushveld. Imagine, this time, a civil action in which the dependants of the missing man bring an action against the tour operators in which they allege that the guide in question either intentionally or negligently caused the death of their breadwinner. The account given by the guide of the man's disappearance is, once again, considered by the court to be entirely implausible: he told the court that the man came unexpectedly upon some friends, decided to go off with them to explore other parts of the area, and decided to reward his guide for his services by making him a gift of all the valuables he had with him at the time (including a vehicle, gold watch, travellers' cheques and clothing). Far more probable, in the court's assessment, is the case contended for by the plaintiffs, that the guide killed the tourist, either intentionally or negligently, and stole his personal possessions. But slightly more probable than even this, in the court's view, is that the tourist died accidentally in the hazardous terrain, after which the guide decided to make off with those items and, when confronted by the authorities, lied in order to exculpate himself in respect of the lesser crime that he had almost certainly committed.

A court faced with a situation such as this might either (a) find for the plaintiffs on the basis that the version put forward by them is far more probable than that advanced by the defendant, or (b) rule that the plaintiffs have failed to discharge the onus of proving that the deceased's death was caused intentionally or negligently by the guide. Which of these approaches a court should take will depend on whether we wish to see a
MUST WE HAVE A THEORY OF PROOF?

civil trial as a contest of case strength or a division of case merit. It is a question that penetrates to the very core of our structure of civil procedure and evidence. The dilemma faced by a court in such a case is this: that whereas the facts contended for by the plaintiffs are more probable than those relied upon by the defendant, the issues relied upon by the plaintiffs have not, in themselves, been proved. The plaintiffs, in such a case, would have had to allege, in their particulars of claim, that (a) the deceased was, in fact, dead; (b) the guide had caused his death; (c) he had done so wrongfully; and (d) he had done so either intentionally or negligently. These are the four essential pillars upon which the plaintiffs’ case rests – they could be described as the ‘issues’, the criteria according to which the assorted facts relied on by either side are or are not considered to be relevant. To find for the plaintiffs in circumstances where one of these issues has not been found to be ‘more probable than not’ would tend to strike at the very centre of our procedural edifice. What is the sense in requiring a plaintiff to allege something, say X, in circumstances in which ‘not X’ leaves him without a cause of action, without requiring him to prove X? To take this further, and to find for the plaintiff where, as in this case, as many as three of the four issues are less probable than not, would seem to require us to redesign our entire practice. Pleadings would have to be seen merely as describing the broad and general ambit of an intended contest, and a cause of action as a variable notion over which the actions of one’s adversary may have an influential and even dominant effect, with the result that the very structure and shape of civil actions become not the solid, absolute notions we are accustomed to but, rather, notions that assume their shape only in relation to what facts happen to have been proved by the other side.

This is not a world that the average practising lawyer would recognise. Nor, in the absence of compelling reasons, should it replace the current practice. It is true, as I have pointed out in a recent article, that our current practice concerning the meaning, function and incidence of the onus of proof is confused, contradictory and even, at times, incomprehensible. It is even true that we have no certainty as to whether the onus applies in our law, to ‘facts’ or ‘issues’. That this is so may be easily shown


27 I should point out that what I have argued in this chapter does not, I believe, in any material respect, affect the chief submissions made in my previous article. In that article I submitted that the onus of proof appeared to be an unsatisfactory and deficient mechanism for resolving who should win and who should lose. It was called upon to play this role, however, only where there was an equipose – where, that is, the probability of the truth of the proposition sought to be proved by the onus-bearer was equal to the probability of the untruth of that same proposition. In every other case, there would be certainty as to who should win or lose, since the onus-bearer would win if the proposition was more likely than not true, and lose if it was not. That this is so, is due entirely, I argued, to the quantum of proof required in civil cases, and not to the rules that determine the incidence of the onus. If we change the rule
by examining the problem as it concerns one area of the law of property – vindication. An owner of property (say A) who wishes to vindicate his property from one in possession (say B) has, in our law, to allege and prove (a) that he, A, is the owner of that property and (b) that B is in possession of it. If B wishes to rely on any right to possession, it is up to B to allege and prove it. The ‘issues’, then, appear to be quite plain: A plaintiff must prove his ownership and the possession of his property by the defendant, whereas a defendant has to prove any relevant right to possession in order to defeat the plaintiff’s claim. What happens, then, if the plaintiff happens, in his particulars of claim, to aver that the defendant holds the property pursuant to a contract of lease which has come to an end, with the defendant alleging that the lease still has another year to run? One might expect the onus to be on the defendant to prove his version of the lease agreement, since its only relevance is on the issue of whether the defendant has a right to possess the property. Our law, however, places the onus on the plaintiff, on the basis that, in view of his acceptance of the contract of lease, he is left without a cause of action unless he alleges (and, hence, proves) that the contract has come to an end by the effluxion of time.28

The fact that our approach is, as I have submitted in my article, sometimes fact-oriented and sometimes issue-oriented when it comes to determining the incidence of the onus of proof is, however, far removed from a practice that would allow a plaintiff to win where most of the issues which go to defining a particular civil claim – in the sense that a failure to aver any one of them would leave him without a cause of action – are not shown by him to be more probable than not. Such a practice would be entirely alien to our experience and would, moreover, I submit, lead to very difficult questions in those situations where there was only a relatively small probability that the onus-bearing party’s version of the facts was true but where the probability of the truth of the account given by the other party was even smaller. Could one realistically expect a court to hold for the plaintiffs in the above case if the probability that the guide intentionally or negligently killed the deceased was, in the court’s view, say 0.1 (with the result that the probability that he did not do so would be 0.9) merely because the probability of the truth of the version relied on by the defendant was even smaller (say 0.01)?

relating to the quantum of proof so as to reflect a Baconian rather than a Pascalian approach, and so as to give victory to the party with the stronger case, we are still left with a point of equipoise – this time where the cases are of equal strength – where the position must be resolved by having rules relating to the incidence of the onus. My arguments concerning the weakness of such rules would apply, I believe, with equal cogency, to such a position.

28 See Henning v Petra Meubels Bpk 1947 (2) SA 407 (T), discussed in Paizes (n 26) 549 and 555.
It is my view, therefore, that it is the Baconian approach this time, that does not and should not reflect the way our legal system deals with problems of this kind. It may be, however, that my view is coloured by my belief—not altogether supported by all the cases—that the onus should not be employed in a mechanical, unthinking way as an accessory to the ‘he who alleges must prove’ policy which has, I believe, had a malign effect on our law. Such an approach sees the onus relating to ‘facts’ only, and not to ‘issues’, which, as I have described them, are abstract categories into which facts of a particular type are placed. In a fact-oriented model, in which no attempt is made to look for or give effect to issues, there is a convergence between the Baconian and Pascalian approaches, since every case will be made up of a series of factual averments and every contest can be broken down into a finite number of single battles between a fact, ‘F’, and its converse, ‘not F’. Liberated from having to consider the effect of those facts on issues, a court has, simply to weigh the probability of the truth of that fact (P (F)), against the probability of the truth of its falseness (P (not F)). One side of this equation will necessarily represent what is contended for by the plaintiff, and the other the defendant, with the winner being the party with the higher probability ratio. I have attempted to demonstrate, however, that this fact-oriented picture is insufficiently nuanced or textured to give expression to the full range of ideas that our law is capable of using in the operation of a mechanism such as the onus of proof. A far richer and more appealing model is one that sees the onus as attaching to issues. And as I have pointed out, the Pascalian argument serves this model far better than does the Baconian within the broader context of our existing procedural and evidentiary framework.

Consider the situation that faced the High Court of Australia in Anchor Products Ltd v Hedges, described thus by Owen J:

‘A sues B for damages claiming to have been injured by B’s negligence. A leads evidence of an act done by B which, if the evidence is accepted, would justify a finding that B had been guilty of negligence and thus caused A’s injuries. But the tribunal of fact is not satisfied that B did conduct himself in the way which A has described. The happening as a result of which A was injured, was however, one which more probably than not would not have occurred if B had taken reasonable care for A’s safety. Is it open to the tribunal, in these circumstances, to infer that B was negligent although unable to determine what was the particular act of negligence C and find in favour of A?’

The court, correctly, I believe, decided that it was. The fact that the facts relied on by A were no more (and may even have been less) probable
than those relied on by B were rightly not allowed to shift the court’s attention from the fact that the issue that had to be established by A – that of negligence on the part of B – had been shown to be more probable than not.

A corollary of my conclusion in this section of the chapter is this: that the test used by the courts in civil cases to determine when an inference may properly be drawn from certain ‘proved’ facts is, in so far as it allows for the proof of any issue where the probability of the existence of that issue has not been shown by the onus-bearing party to be more probable than not, flawed. The only approach that may be supported is one that ‘works backwards’, in the sense that one applies to each step in the chain of inferences (each yielding, in effect, successive intermediate facts) such a standard of proof as will allow, ultimately, for the proof of the ultimate issue or issues as more probable than not. Whether this is done in a Pascalian fashion (with the corrosive effect of the product rule) or according to the Baconian model put forward by Cohen (according to which a final issue is regarded as more probable than not if no intermediate fact is established at any lower level of proof) is a question that takes us back, of course, to the first of the two problems considered in this article.

III CONCLUSION

It was fashionable, once, to believe that legal notions could be expressed as functions of mathematical concepts. In fact, until 1977, when the views of Cohen began to make an impact, the conventional wisdom would have been that there was little reason to doubt that evidentiary principles, in so far as it was at all possible or appropriate to think of them in an empirical fashion, would rest upon the accepted canons of probability theory. There was even, from time to time, express judicial recognition of this belief. It is best reflected in what was said by Lord Simon of Glaisdale in Davies v Taylor:32

‘Beneath the legal concept of probability lies the mathematical theory of probability. Only occasionally does this break surface – apart from the concept of proof on a balance of probabilities, which can be restated as the burden of showing odds of at least 51 to 49 that such-and-such has taken place or will do so . . . Perhaps forensic science experts could use the mathematical theory more frequently – for example, in combining items of circumstantial evidence (say, similarities of boot-prints, dust and cigarette ash); the respective odds do not combine simply by addition.’

32 [1974] AC 207 (HL) at 219f–H.
It is typified, too in the approach taken by Dixon J in the decision of the Australian High Court in *R v Jenkins: Ex parte Morrison*:\(^{33}\)

‘With any chain of circumstantial evidence the chances of error in the conclusion arise first from the chances of error in each fact or consideration forming the steps and second from the chance of error in reasoning to the conclusion from the whole of those facts and considerations. It is therefore wrong to take each fact or consideration separately, to assess the possibilities of error in finding it is established and then if you think it should be found afterwards to treat it as a certainty and pass to the next fact or consideration and so on to the conclusion. The possibilities of error at all points must be combined and assessed together.’

An obvious drawback to the use of mathematics in legal contexts is that lawyers are not, by and large, mathematically trained, and it is deceptively easy to take the wrong path. Three examples, one each from the United States, Australia and South Africa, will demonstrate my point.

In the Californian case of *People v Collins*\(^{34}\) the defendant had been convicted of second-degree robbery on the strength of the following evidence: The complainant had been robbed of her purse by a white woman, weighing about 145 pounds, wearing dark clothing and wearing her medium blond hair in a ponytail. Shortly after the robbery, she ran out of an alley and was picked up by a black male, with a moustache and beard, and driving a yellow automobile. The description of the woman matched that of one J, who, her employer testified, was picked up at about the time of the robbery on the day in question by someone driving a yellow automobile. The defendant, who was identified as the driver of the yellow automobile seen to have picked up J after the robbery, was a black male who, at the time, was found to be driving a yellow Lincoln. In an attempt to bolster the identifications made by the various witnesses, the prosecutor called a college mathematics instructor to seek to establish that, assuming the robbery was committed by a white woman with a blond ponytail who left the scene in a yellow automobile in the company of a black man with a moustache and beard, there was an overwhelming probability that the crime was committed by any couple answering such distinctive characteristics. The prosecutor then attached probability ratios to each of the six chief characteristics:

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partly yellow automobile</td>
<td>1/10</td>
</tr>
<tr>
<td>Man with moustache</td>
<td>1/4</td>
</tr>
<tr>
<td>Girl with ponytail</td>
<td>1/10</td>
</tr>
<tr>
<td>Girl with blond hair</td>
<td>1/3</td>
</tr>
</tbody>
</table>

\(^{33}\) (1949) 80 CLR 626 at 644.

\(^{34}\) 438 P 2d 33.
Applying the product rule to these factors, the prosecution argued that there was only one chance in 12 million that any couple possessed these distinctive characteristics, so that there was only one chance in 12 million that the two were innocent.

The Supreme Court of California, however, correctly held that this reasoning was gravely defective. Not only was there inadequate proof of the statistical independence of the six factors (and ‘black men with beards’, to begin with, clearly overlapped with the category ‘men with moustaches’), so that the product rule would ‘inevitably yield a wholly erroneous and exaggerated result even if all of the individual components had been determined with precision’, but there was also no evidence to support any of the individual probability factors ascribed to the six characteristics. Moreover, the court added, the entire enterprise was flawed and misguided. Even if only three couples in the entire city matched the distinctive characteristics in question, the prosecution’s approach could furnish no guidance at all on the crucial remaining question: of the admittedly few such couples, which one, if any, was guilty of committing this robbery?

The Australian case of *TNT Management Pty Ltd v Brooks* involved an action arising out of a collision in which it was necessary for the plaintiff to prove some negligence on the part of the driver of a pantechnicon. Murphy J decided the case in the following way. First, he concluded that there were only four possible explanations: one driver was solely negligent; the other driver was solely negligent; both drivers were contributorily negligent; or neither driver was negligent. Secondly, he dismissed the fourth possibility as being only a theoretical possibility—a fact conceded by defendant’s counsel. Thirdly, he considered that each of the remaining three possibilities was equally probable. And fourthly, he concluded that the probability that the driver of the pantechnicon was in some way negligent amounted to ‘two out of three’, so that it was more likely than not that he was, in fact, negligent. The majority of the court would have no part of Murphy J’s mathematical analysis. It rests, of course, on the assumption of equi-possibility (also called the principle of indifference) which, as Ligertwood points out, can constitute a good reason for decision only where there are no reasons for distinguishing between members of a class. In this case, however, there was evidence that rendered such an assumption inappropriate, such as the conditions at

---

35 Collins (n 34) 39.
36 (1979) 23 ALR 345.
37 At 22.
In Ligertwood’s view, was not sufficiently considered by Murphy J.

The third case, *S v Mukunga*, is South African. The question that confronted the court here was how to deal with a situation where six firearms had been found in a hut occupied by seven people in circumstances where their possession of these guns would be unlawful in terms of a statutory provision and where it was clearly probable from the evidence of the dispersion and location of the guns, that the guns were not collectively possessed, whether by any one person alone or by any number of persons jointly. Since one of the accused must, on the strength of this evidence, inevitably have lacked a firearm, and because that accused was not identifiable and might equally have been any one of the seven, it was argued that all seven were entitled to the benefit of the doubt. The court held that while this reasoning would have been faultless if it had been incumbent on the state to prove the actual possession of a firearm by each accused, it had to be rejected on the ground that the statutory provision placed the onus on the accused in circumstances of this kind to prove that they were not in possession of the firearms. That the onus could not be discharged by the accused merely by the above statistical evidence could, said Didcott J, be demonstrated by a ‘simple example’:

> Ten men and ten knapsacks are found together in a room. The actual possessor of each individual knapsack is unknown. One knapsack contains a revolver. Each man, when subsequently prosecuted for its unlawful possession, argues that, because the odds against his having been its actual possessor are ten to one, the preponderance of probability favours the conclusion that he was not that particular person. The argument is however fallacious. The true position in the circumstances postulated is that each of the men is inherently as likely as any one of the others to have been the actual possessor. When each of them is called upon to prove that he was not that person, more is required of him than a mere reference to those same circumstances. His evidence that he did not in fact possess the revolver will of course suffice, provided that the denial is acceptable. In this case the denials of the accused cannot however be accepted.

Even if one were to agree with the learned judge’s finding on the facts before him – and one has reservations even here on the basis of what has been said above about the meaning and assessment of probability in situations such as these, particularly where one is faced with the certainty

---

38 1976 (3) SA 193 (O).
39 The presumption that effected this reversal of the onus has since been held to be unconstitutional in *S v Mbatha; S v Prinsloo* 1996 (1) SACR 371 (CC).
40 *S v Mukunga* (n 38) 198.
of a wrong conviction – the ‘simple example’ does not support his conclusion. As Schmidt points out, the fact that ‘each of the men is inherently as likely as any one of the others to have been the actual possessor’ is not the point. The real question is whether it is more probable that each, taken in isolation, is not the possessor than that he is the possessor (since the onus is on the accused), a question that must, even if one judges Didcott J’s argument by its own lights, be answered in the affirmative, since the ‘probability’ of his innocence would be 0.9 and not 0.5 as the judge would have it. On such a determination, the onus resting on each accused must be found to have been discharged.

Cases such as Collins, Brooks and Mukunga demonstrate that courts may apply mathematical principles erroneously in order to reach conclusions of law that may appear seductively convincing to those not sufficiently well versed in the use of those principles. The deeper question, however, remains to be answered. Even if we do not make mathematical errors, to what extent is it appropriate to be using mathematical principles to solve legal problems in the first place?

It is this question, so summarily answered in Davies v Taylor, that is so pivotal in our quest to find a theory of proof. The last two decades have, on the whole, seen a waning enthusiasm for the use of mathematics to solve legal problems. In 1975, Laurence Tribe of Harvard University published a highly influential article, ‘Trial by mathematics: Precision and ritual in the legal process’, in which he argued persuasively against the use of mathematics. His reasons are not easily summarised, but they seem, in essence to derive in part from the tendency of more readily quantifiable variables to dwarf those that are harder to measure, in part from the uneasy partnership of mathematical precision and certain important values, in part from the possible incompatibility of mathematics with open-ended and deliberately ill-defined formulations, and in part from the intrinsic difficulty of applying techniques of maximization to the rich fabric of ritual and to the selection of ends as opposed to the specification of means.

No one, of course, would argue that all evidentiary principles could be reduced to mathematical algorithms. Apart from the fact that any claim of this kind would be untenable even within a purely mathematical argument since, as Gödel’s famous incompleteness theorem demonstrates, not even mathematical principles rest on notions that are entirely computable in this sense, any practising lawyer will tell you that our legal system just does not work in this way. To exclude mathematical ideas entirely from legal proof, on the other hand, would seem to be

41 Schmidt and Rademeyer (n 16) 89.
42 (1971) 84 Harvard LR 1329.
43 Ibid 1393.
 unacceptable. As the discussion relating to the civil onus and the quantum of proof demonstrates, there is clearly a place for such ideas, and the inductive reasoning of the Baconians would seem, on occasion, to lead to error. The deployment of the tenets of probability theory have, moreover, led to the development of celebrated models such as Bayes' theorem, a mathematical mechanism that purports to allow one, with precision, to assess the probative impact of a new item of evidence on a legal argument. Even though it is seldom possible, in practice, to make a precise calculation, since values are rarely available in empirical quantities, writers such as Eggleston claim that a proper understanding of the mathematical construction of these models enables lawyers to detect fallacy and to prevent the commission of serious error. I believe Eggleston is right. It is dangerous to jettison mathematical aids. Some writers have great confidence in the descriptive power of Pascalian methods, even in so far as they may be applied to basic legal concepts. Schum and Du Charme, for instance, have derived certain formulations for evaluating the inferential value of testimonial evidence from witnesses whose credibility is less than perfect. They claim, rather remarkably, that these formulations not only show that the inferential value of such testimony depends upon the relative rareness of the event being described or reported and the credibility of the witness reporting it, but also prescribe the exact nature of the interaction between these two factors. They have devised, too, other Pascalian-styled techniques for evaluating an item of evidence in the light of previously given items, showing, in their view, that the process of evaluation is akin to the scientific techniques employed in the field of contrast processes such as colour and brightness in sensory perception.

If one bears in mind the fact that most judges, if asked whether mathematical probability underpins legal probability, would probably take the view – even today – that there would seem to be no reason why it should not, it would seem to be rash to set one's mind against all Pascalian models. A better approach would be to accept that there may well be areas where such a model sits well with our legal judgment and serves legal ends without the need for adjustment or replacement. Where this is not the case, it is most instructive to understand precisely why the Pascalian model will not do. It becomes, easier, then, to find a new model, one based on Baconian or any other type of reasoning (such as,

---

44 Put forward by Reverend Thomas Bayes in 'An essay toward solving a problem in the doctrine of chance' Philosophical Transactions of the Royal Society of London (1763).
45 Op cit (n 4) 139.
46 See Schum and Du Charme 'Comments on the relationship between the impact and the reliability of evidence' (1971) 6 Organizational Behaviour and Human Performance 111. See, too, Schum (n 14) 472–3.
say, deductive reasoning). It was the Baconian model, for instance, that seemed to produce the best results when it came to cascading inferences and the proof of complex or multiple issues.

The point seems to be this: that there is no reason to suspect that one particular theory will apply to all situations. Neither, I believe, can one predict which theory will best solve a particular problem. The solution will, however, always be a legal one, in that it will always rest on a judgment made from norms and principles that do not belong to mathematics, probability theory, inductive reasoning or any other non-legal system or body of rules. All we can do is test the various candidates and make a legal judgment as to which best serves the policies of our legal system. We reject the Pascalian model when it comes to cascading inferences because, I believe, it leads to results which are, when tested by the lights of a working practical system, alien and absurd. It is too exacting in setting standards and does not reflect the human (if not mathematical) experience of accepting that once a fact is proved to our satisfaction, what little doubt there is as to its existence should not go into a ‘pot’ with other small ‘quantities’ of doubt attending the proof of other facts so proved until the mix becomes big enough to constitute what we may call a ‘reasonable’ doubt. The Baconian model gives expression to this experience, so we use it to describe what we do in that context.

IV IN SHORT

Yes. We do need a theory of proof. Not one that fits all sizes, but one for each size. We must, in every evidentiary context, ask which, if any, of the various theories works and, equally important, why it works. I believe our courts have done this at a deeply intuitive and subconscious level. It is time to do so openly and expressly. It will not always be an easy task. I believe, therefore, that our courts should tread warily, that they must be alive to the complexity and relative newness of these ideas, yet mindful of their crucial practical and theoretical importance. For, difficult as it may be for lawyers to make choices based on areas of learning peripheral to the legal experience, it is undeniable that such choices are unavoidable. And even if a choice is made on an incomplete or even inadequate grasp of some of these ideas, such a choice must, I believe, be preferable to the only alternative, which would be to make what amounts to a choice without engaging those ideas at all – either in ignorance of their existence or out of a timid or obdurate refusal to accept a difficult challenge.

47 See Eggleston (n 4) 131.