

Choice and Responsibility*

JIM EVANS**

Introduction

The first of the books reviewed in this essay is a collection of essays on responsibility written by Tony Honoré. With one exception they were written after 1985, the date of the second edition of Hart and Honoré's, *Causation in the Law*. The exception is an essay entitled 'Can and Can't', published in *Mind* in 1964, that is relevant to the issue of free will and determinism. Because the basic point in it is used in some of the later essays it is re-published here as an appendix. All the essays, apart from the first, have been previously published. The first essay is a useful introduction, and contains a very interesting account of the evolution of the ideas expressed in the first edition of *Causation in the Law* in 1959. The second book is a collection of eight essays deriving from two conferences on the first book: the first in Canberra and the second in New York. The book finishes with a helpful response by Honoré to each of the essays.

The manner in which the essays in the second book engage with Honoré's ideas in the first book varies. Two philosophers, Michael Smith and Phillip Pettit, connect only with the idea expressed in 'Can and Can't', in developing their own contributions to aspects of the free will problem. William Lucy, in a wide-ranging essay, attempts only to place Honoré's work within a theory he develops about types of modern legal scholarship. The other five essayists each undertake a close critique of aspects of Honoré's views. Among these, Jane Stapleton considers Honoré's views about causation, while the other four, Peter Cane, John Gardner, Stephen Perry, and Arthur Ripstein, focus primarily on his views about responsibility.

By the time I had finished reading both books I had concluded that I disagree with Honoré's basic thesis. Although I can find occasional support for the views I hold among the essayists in the second book, I felt that none of them identify the flaws in Honoré's position with sufficient clarity.

* A review essay on Tony Honoré, *Responsibility and Fault* (Hart Publishing, Oxford, 1999) hereafter 'RF'; and Peter Cane and John Gardner, (eds), *Relating to Responsibility* (Hart Publishing, Oxford, 2001) hereafter 'RR'.

** Professor of Law, University of Auckland.

Indeed, I often found myself disagreeing with the essayists for the same reasons that I disagree with Honoré. It seemed appropriate, therefore, to write an essay setting out my disagreement with Honoré in which I would bring in points in the second book when these were relevant. I shall start by explaining, in general terms, the character of my disagreement with Honoré.

Honoré argues that the most basic form of responsibility in human affairs is what he calls ‘outcome responsibility’. His account of outcome responsibility is in two parts: one dealing with acting and the other with not acting. In the case of acting, we are outcome responsible for all the outcomes we cause by our voluntary actions.¹ So long as the actions themselves were voluntary, it does not matter if the outcomes were unforeseen or unintended.² It does not even matter if the outcomes were not foreseeable.³ Thus, outcome responsibility for adverse outcomes does not depend upon fault, and may depend on sheer bad luck.⁴ In the case of not acting, we are not outcome responsible for everything we do not do, but only for ‘omissions’. Honoré defines these as violations of a norm placing us under a specific responsibility to act.⁵ My concern here is with responsibility for actions. Although I shall occasionally talk about praise, my principal concern is fault.

Honoré’s position about responsibility strikes me as both odd and wrong. I disagree with many details of his argument, but two complaints are central.

The first is that he fails to distinguish carefully between different senses of ‘responsible’ and its cognate ‘responsibility’. These terms can be used in many different ways, but for the present discussion it is important to distinguish at least the following three uses:

1. To say someone is responsible for a certain outcome can mean simply that he caused it. If an action was involved, it isn’t even necessary that it be voluntary. To the question, ‘Who was responsible for the broken window?’ the answer might be ‘John, while he was sleep-walking.’ I shall call such responsibility ‘causal responsibility’. Honoré’s notion of ‘outcome responsibility’ is a species of causal responsibility. It is confined to outcomes that are caused by voluntary actions, and it includes as forms of causing ‘near-causal relationships’, such as inducing, advising, or providing opportunities

¹ RF 7, 76-8, 129 fn 22; RR 223, 227.

² RF 129: ‘what we do includes its unintended aspects’; see also RR 227.

³ Ibid.

⁴ RF 9, 14-15.

⁵ RF 77, and 46-54.

for, others to act.⁶ In this sense of ‘responsibility’ it is, of course, tautological that we are responsible for the outcomes we cause.

2. To be distinguished from this sense is a sense of ‘responsibility’ that has become common in philosophical literature, which I shall call ‘agency responsibility’. Let me clarify it. We can usefully distinguish between two types of conditions under which a person can fairly be blamed or praised for conduct. The first is that the conduct is appropriately connected to the agency of the person; the second is that it possesses features – other, of course, than such connection – that make it worthy of praise or blame. A person is ‘agency responsible’ for conduct when she satisfies conditions of the first type in relation to it.
3. The third relevant sense of ‘responsibility’ means, simply, a legal, moral, or similar, obligation that is understood as imposing a burden. We use the term in this sense when we say, ‘It is my (his, her) responsibility to do X’. We don’t always say that a person who has an obligation imposing a burden has in the relevant respect a responsibility, but we tend to when the obligation is undertaken, arises from the obligor’s fault, or arises from a status or role the obligor has. A responsibility, in this sense, may vest on the happening of some untoward event. When it does, the untoward event need not have been the fault of the person who comes under the responsibility; indeed, he need not even have caused it. For example, in the case of strict liability in the law of tort, we may say that a person who satisfies the conditions of such liability has a responsibility to take the remedial action the law requires. That means only that he has an obligation to do so: it does not imply that he was agency responsible (ie blamable) for the event that vested the liability.⁷ It doesn’t even imply that he caused this event. Often a person who is strictly liable will have caused the vesting event; but an owner of property may be liable to abate a nuisance originating from his property, even though he did not in any sense cause the nuisance. I shall call this form of responsibility simply ‘responsibility meaning obligation’.

‘Outcome responsibility’ is, as I have said, a species of causal responsibility, but Honoré often attributes to it consequences that are appropriate to agency responsibility. For example, although outcome responsibility does not justify *moral* praise or blame, it does justify ‘credit’ and ‘discredit’ among members of one’s community. These look awfully like forms of praise or blame, especially when we learn that credit goes

⁶ RF 1-7; RR 227. Cf Cane RR 89.

⁷ Cf Cane RR 100.

together with 'esteem' and discredit with 'resentment'.⁸ Admittedly, Honoré mounts a number of arguments intended to show that outcome responsibility justifies these responses. I think these arguments fail, but I also think that his position often gains apparent plausibility from confusion between causal responsibility and agency responsibility.

Let us, however, turn briefly to the merits, being careful not to confuse the two ideas. In my view esteem and resentment are not justified unless an outcome is more closely linked to the agent's choice than the limited link required for it to be caused by the agent's voluntary conduct. That there is an important issue here follows from the fact that there are many outcomes that are caused by our voluntary actions that we did not choose, and could not reasonably have been expected to avoid. To take a simple case, which derives from Donald Davidson:⁹ coming home in the evening, I enter my flat and turn on the light in the usual way. This causes a burglar who is on the premises to leap from a window and suffer severe injuries. My turning on the light was voluntary, and it caused the burglar to leap. But I did not choose this outcome and could not reasonably have chosen to avoid it. Alternatively, in case anyone thinks that the burglar's jumping breaks the chain of causation, let it be that turning on the light starts a fire because of an electrical fault that I had no reason to suspect. Again, my voluntary action causes an outcome that I did not choose and could not reasonably have chosen to avoid.

In my view, nothing regarding the merits of Honoré's position turns on any difference between 'esteem' and 'praise', or 'resentment' and 'blame'. The real question is what are the conditions for agency responsibility. Since I believe that Honoré also tries to co-opt for outcome responsibility other consequences that properly belong only to agency responsibility, such as being essential to our identity as persons, the conditions of agency responsibility are central to my argument. I shall argue that an agent can only be blamed for producing an outcome if the outcome satisfies one of three conditions: (i) it was wanted by the agent, (ii) it was foreseen and accepted by the agent, or (iii) it was a consequence of the agent's negligence. The first two of these are clearly linked to choice. Knowingly creating a risky situation will reduce to either (i) or (ii): I shall not consider it further. However, it may seem that inadvertent negligence is not linked to choice. I shall argue that it is. Why these particular conditions? I can give here an answer that is only a place-holder, and will need to be elaborated later: they are important because each allows us to say that the agent could, through choice, have made things otherwise. I shall not explore closely the conditions for praise, but they plainly follow a similar structure.

⁸ RF 26.

⁹ 'Agency' in *Essays on Actions and Events* (Clarendon Press, Oxford, 1980) 43, 53.

Honoré is also capable of confusing responsibility, meaning obligation, with agency responsibility. For example, as he points out,¹⁰ a relative can become morally responsible for bringing up an orphaned child, although the death of the child's parents was from the relative's point of view an unexpected piece of bad luck. This means only that the relative has an obligation to care for the child. But Honoré takes it as rebutting the position that moral blame should never depend on luck. It doesn't: the responsibility that attaches to the relative under these circumstances doesn't imply that the relative was morally to blame (ie agency responsible) for these circumstances, as Honoré acknowledges.

The second general complaint that I have against Honoré applies also to the essayists in the second book (other than Lucy, to whose essay it is irrelevant). Neither Honoré, nor the essayists, pay sufficient attention to the various relationships in human life between reasons for action, choice and conduct. Only when we understand that, can we see how fault, and hence blame, are related to choice.

I shall start with this topic, and then turn to the issues discussed by Honoré and the essayists. In the preface to the second book, the editors usefully organise the issues discussed into categories. I shall follow their order, discussing 'capacity', 'responsibility', and 'causation' in successive sections. The first of these sections, on capacity, is concerned with the free-will problem: that is how anybody can be responsible for anything. The second considers Honoré's arguments for outcome responsibility. The third, which is very brief, considers the role of causation in the law of negligence. It applies to that question the account of negligence given in the next section.

Choice and Fault

We can begin with the question 'What are reasons for action?'. The answer, I think, depends on whether we are concerned to identify that which motivates an agent, to explain an agent's intentions or action, or to consider what ought to motivate an agent.¹¹ I shall start with the first, which I think is fundamental. I shall also start with a straightforward case in which an agent considers reasons for and against an act, decides to perform it, and then does so. What motivates an agent in such a case, I contend, are propositions believed true about a type of act. Let me defend this position.

¹⁰ Cf RF 131.

¹¹ Cf the Introduction to Joseph Raz, *Practical Reasoning* (Oxford University Press, Oxford, 1978) 2-4, which initiated some of the reflections reported here.

Clearly, it is beliefs, rather than truths, that motivate an agent, since an agent cannot know when her beliefs are false. It is not, however, the fact that an agent has a belief that motivates, but the content of the belief. Of course, the autobiographical proposition that I hold a certain belief may form part of a more complex belief that motivates me to act; but the thing that then motivates will be the content of this further belief, not any of: the initial autobiographical fact, the proposition that expresses that fact, or the new autobiographical fact that I hold the more complex belief. Let us take a simple case to illustrate these points. Suppose I have been arguing on theoretical grounds that the stock market has reached its zenith. Later I reflect that I really do believe that, and that if I am to be consistent with this belief I ought to sell my shares now. Here I may have two distinct reasons for acting: that I want to be self-consistent, and that I want to maximize the return on my investments. The initial autobiographical proposition, that I believe that the stock market is at its zenith, forms part of the first reason. More fully stated that reason is that I am committed to maximizing the return on my investment, believe that the stock market is at its zenith, that a self-consistent person who had this commitment and belief would sell now, and that, therefore, if I am to be such a person I must sell now. It is the content of that belief that motivates, not (1) the initial autobiographical fact that I believe the stock-market is at its zenith, (2) the proposition that records this (on its own), or (3) the new autobiographical fact that I believe that if I am to be self-consistent I must sell now. As to the second reason (that I want to maximize my return) the initial autobiographical proposition forms no part of it. Only the content of what is believed – that the share market is at its zenith – forms part of this reason. For that proposition, along with the obvious proposition that selling at the zenith maximizes returns, yields the conclusion that selling now will maximize my return, which provides a reason for selling.

It may surprise some readers, but similar points hold about propositions that the agent wants something that the act will conduce towards (ie produce, tend to produce, or constitute). (For simplicity, I shall treat all pro-attitudes towards things that an agent believes an act will conduce towards as ‘wants’.) Such propositions are certainly needed when we turn to explain an action, but they are not needed as part of the content of the beliefs that motivate. Were they to be needed, we would not be able to act for a given reason until we had taken account of the autobiographical proposition that we want some particular thing that the reason shows the act will conduce towards. But if we really do want that thing, the autobiographical proposition that we want it is redundant. That the act will conduce towards the thing that is wanted is on its own perfectly ample as a reason. It could only be necessary to take account of the autobiographical proposition if we were condemned to acting only when we had not merely a want, but also a higher order want to act on this want. For example, I would

not be able to sell just because selling would maximize my return, but only if I also wanted to sell because (ie on account of the fact that) I would be acting for this reason. That, surely, is not how things are; but in case anyone is inclined to think the contrary, it is worth recording that the problem iterates. We could only act on a second order want to act on a first order want if we had a third order want to act on the second order want, and so on, endlessly. Since at some stage we must act on a want without also acting on a higher order want, there seems no reason why this should not be possible at the first level. Indeed, it patently is: mere recognition that selling will maximize my return is ample to motivate, without it being necessary to add that if I so act I will be acting on my want to maximize the return.

As I have said, it is different when we seek to explain an agent's actions. Here what we call reasons are the beliefs and wants that led the agent to perform the action. Again, it is different when we consider what ought to motivate an agent. There are many points of view from which we may do this: moral, prudential, taking an agent's commitments or wants for granted, and so on. Here it is normally facts rather than beliefs that count as reasons, and not every fact that would motivate an agent if known by him, will count as a good reason.

I said at the outset that the things that motivate us are propositions believed true of *types* of acts. It must be types of acts, because in thinking ahead we cannot plan for every detail of the concrete action that will occur if we act. No matter how carefully we plan there will be things true of the particular action that we did not plan. I plan, let us suppose, to execute a particular dance movement with exquisite precision. But I do not plan just exactly how much breath is in my lungs at the time, which set of micro-organisms I will stand on, or how many windows are open in the building at the time. An infinite multitude of such details will be settled by my habitual mode of acting and the way the world happens to be. Thus, reasons are always general to this limited extent: they apply to an open class of actions. It is, however, too simple to think that what we want when we have weighed the reasons for an act and decided to perform it is to perform *any* action of the relevant type. What we want is to perform an action of that type that has no adverse features other than those we have anticipated and accepted. I want, let us suppose, to turn on the light. There are multitudes of ways of doing it. Among these, I want one that does not break any furniture, knock any people over, or cause similar havoc, along the way. There are always side-constraints that shape our actions.

The points made above allow for an elegantly simple account of practical deliberation. Practical deliberation is deliberation carried on with a view to acting in which we seek to identify the propositions true of a type of act that provide reasons for or against performing an act of that type. We may, of course, start with a goal or a commitment, rather than a

contemplated act, and then seek to find an act (type) that would serve the goal, or fulfil the commitment. But the end-state sought will still be the identification of an act that there is, as a consequence, and given other propositions about it that provide reasons for or against performing it, reason to perform. Practical deliberation no doubt requires special abilities in a mind: the mind must be capable of bringing forward relevant concerns about a type of action for contemplation, and good practical deliberation will require a refined skill in this. However, it does not require any special logic:¹² the only logic needed is the ordinary logic for deriving propositions that are true if their premises are true.

So much for careful deliberation: what about spontaneous actions? Seeing a car bearing down on me, I leap: I do not stop to contemplate that leaping gives me the best prospect of safety. Asked a simple question, I respond immediately: I do not calculate that my answer has the qualities being truthful and helpful. Having risen, dressed, and breakfasted, I travel to work in the usual way: I do not reflect that these actions will equip me to fulfil my employment obligations and earn my salary. Life is full of such actions. At least three things show this type of case to be a close cousin of the deliberative type of case considered above. Firstly, if the action is to be intentional, and not just a mechanical reflex such as starting when we get a fright, some mechanism in the agent has selected behaviour that is appropriate to one or more goals or commitments that the agent has. Secondly, if the agent was asked why he acted in the relevant way, he could truthfully point to the contribution of the action to these standing goals or commitments. Thirdly, the agent does not normally have to work out, by theory construction, as an observer would, what were the goals, commitments, and beliefs, that led to the action, he can re-construct this from memory. So, although an agent does not always deliberate consciously in the way sketched above when choosing an act, the account is an adequate basis for understanding the way in which acts are linked by reasons to goals or commitments of the agent, when action is intentional.

A further point about this account is important. It enables us to see why practical conclusions are defeasible. The critical point is that relevant to some set of truths, say A-D, that constitute reasons for or against an act, there may be reason to perform it, but relevant to a wider set of truths, say A-E, there may not be.¹³ There is reason, Aquinas said, to return a thing borrowed (indeed the natural law requires it); but not if the thing is a sword and the lender a madman in a rage.¹⁴

¹² Pace Anthony Kenny in *Will, Freedom and Power* (Blackwell, Oxford, 1975) ch 5.

¹³ Cf Raz, above n 11, 11-14.

¹⁴ *Summa Theologica* 2a2ae, Q120 (Blackfriars edition, London and New York, 1963-1980, Vol 41) 277; cf 1a2ae, Q94, A4; Q 96, A6 (Vol 28)

It will be useful at this point to formalize a distinction I have already been observing between 'acts' and 'actions'.¹⁵ The term 'act' I shall use in three related ways: when talking about the contemplation of conduct ahead of time, it will refer either to a type of act or to some undifferentiated instance of a type of act; when talking about conduct that has occurred, it will refer to the pair of a concrete performance and a description of that performance capable of marking out a type of act that could be contemplated ahead of time. I shall use the term 'action' to refer to a concrete performance not linked to any description, but capable, of course, of being described by an array of different descriptions.

We can now note that the logically primary use of the term 'intention' is to refer to acts rather than actions. This is because when we form intentions we necessarily contemplate types of act. Even spontaneous conduct is undertaken because it is of a *type* that links it to commitments and goals. Actions cannot be intentional in this way, because we cannot contemplate ahead of time the concrete event that occurs with all its detail. Actions can be intentional only in a derivative sense that they constituted one or more intentional acts. An alternative way of putting this point is illuminating: 'intention', when referring to conduct that has occurred, is primarily a three-place predicate: it serves to link an agent to conduct and a description of that conduct. However, it can be used derivatively as a two-place predicate, merely linking an agent to conduct, when it is believed (or assumed) that at least one proposition using it as a three-place predicate is true. Normally, we use 'intention' in the three-place sense, even when speaking of actions. In this sense, an action may be intentional under one description but not under another. To illustrate: a person in a queue at the theatre buys two seats. They turn out to be the last two seats available, so her action causes disappointment to the person immediately behind. Her buying of the seats was intentional, but her causing the disappointment was not.

When Honoré says that outcome responsibility requires that the action that produced an outcome was voluntary he seems to mean that it must have been intentional under some description.¹⁶ (If he does, that is not quite right, for aimless action, like strutting backwards and forwards across

87-91; 137-141. For further discussion of the causes of defeasibility see Jim Evans, 'Aristotle's Theory of Equity' in W Krawietz, N MacCormick and G H von Wright (eds), *Prescriptive Formality and Normative Rationality in Modern Legal Systems: Festschrift for Robert S Summers* (Duncker & Humblot, Berlin, 1994) 225.

¹⁵ Cf G H Von Wright, *Norm and Action* (Routledge, London, 1963) chs 3 and 4: A Kenny, 'Intention and Purpose in Law' in R Summers (ed), *Essays in Legal Philosophy* (Blackwell, Oxford, 1968) 146, 150.

¹⁶ RF 129, 132.

a room can be voluntary. Perhaps what is important in such a case is that the agent could have chosen to act otherwise if he had wanted not to act as he was doing.) When Honoré says that the outcome itself can be unintended or unforeseen, we can understand him as saying that the act of producing the outcome can be unintended or unforeseen.

Let us now consider how these various points relate to the traditional categories of legal fault: intention, recklessness, and negligence.

An act will be intentional in a full sense if it was anticipated, wanted, and the agent acted for the reasons that led him to want it. An agent who performs an instance of a proscribed act will be at fault if his act was intentional in this full sense, unless there was a circumstance present that justified the act notwithstanding its adverse features.¹⁷ Whether an agent was justified, given the beliefs he had, is, of course, judged from an evaluative point of view.

An agent may know she will, or may, perform an act, without wanting to do so. If the theatre-goer described above knew she was buying the last two seats and this would disappoint the person next in the queue, but did not want to disappoint that person, or wanted to do so but did not act for this reason, then that act was not intentional in the full sense described above. Bentham coined the useful term 'obliquely intended' to refer such acts.¹⁸ In ordinary discourse whether we call such an act intentional depends not only on how likely it was that the agent would perform it (inevitability standing at one extreme), but also, when fault is in question, whether there was good reason to act for the reasons that motivated the agent, notwithstanding that she knew that she might perform this act in doing so. Reckless acts are a proper sub-set of acts that are obliquely intended. I will define them as those *adverse* acts (such as causing an unwanted consequence) that the agent foresaw to a level less than inevitability, the prospect of which did not motivate her, but which she was willing to run an *unacceptable* risk of committing, in the pursuit of the things that did motivate her. Because of the particular terms of my definition, reckless acts in this sense will always involve fault. It is worth noting, because it matters to the later argument, that an agent who gets away with a reckless risk is equally at fault. If an act that was obliquely intended was recognised as inevitable, we are more likely to say that it was intentional rather than reckless, so I have framed my account of 'recklessness' to take account of

¹⁷ Or unless, as can happen, the adverse features are not present. For more detailed discussion see Jim Evans, 'Aristotle's Theory of Equity', above n 14; and section 1 of 'A Brief History of Equitable Interpretation in the Common Law System' in Jeffrey Goldsworthy and Tom Campbell (eds), *Legal Interpretation in Democratic States* (Ashgate, Aldershot, 2002) 67.

¹⁸ Jeremy Bentham, *An Introduction to the Principles of Morals and Legislation*, J H Burns and H L A Hart (eds), (Methuen, London, 1982) 86.

that. Obliquely intended acts of this kind will normally involve fault in the same cases as fully intentional acts, although that an agent did not want to commit the act, may mitigate the fault.

Reckless acts may also be called negligent; but negligence is a wider category than recklessness, since an agent who is negligent need not have adverted to the unacceptable risk. Inadvertent negligence poses a problem for a theory of responsibility because it seems not to involve choice. As H L A Hart noted, the puzzle appears in the following type of exchange:¹⁹ ‘I didn’t mean to do it: I just didn’t think.’ ‘But you should have thought.’ Hart remarks: ‘Such an exchange, perhaps over the fragments of a broken vase destroyed by some careless action, is not uncommon.’ In such cases we appear to accept that no choice was made to run the unacceptable risk, but reject that this precludes fault. I don’t think Hart solves this problem. At times he appears to suggest the fault arises because the careless person has chosen not to think of the risk.²⁰ But while this can happen, it is not the common case. At other times he suggests that the careless person is at fault if he had the capacity to think of the risk, but failed to exercise it on this occasion.²¹ However, if no choice was made whether to exercise the capacity, we need to know more about it, and why it is relevant, before we have an answer to the puzzle. If all that is being said is that nine times out of ten the agent would have thought of the risk in situations of this type, it is not obvious why we should blame him on the tenth occasion.

A valuable clue to solving the puzzle is provided by John Mackie, who points out that behaviour may be caused by the *lack* of any sufficiently strong desire for contrary behaviour.²² We can envisage the presence of such a desire making a difference in two possible ways. Firstly, we may believe that such a desire, if present, would have modified the careless person’s behaviour in the way that standing goals and commitments lead to spontaneous action. For in the case of spontaneous action, the agent’s conduct follows the pattern of his standing goals and commitments without contemporaneous reflective choice. Secondly, we may believe that if the agent had had a sufficiently strong desire, he would have recognised the risk and had an opportunity to choose whether or not to run it.²³ As we have already noted, the mind of a person capable of practical deliberation must

¹⁹ ‘Negligence, *Mens Rea* and Criminal Responsibility’ in *Punishment and Responsibility* (Clarendon Press, Oxford, 1968) 136.

²⁰ *Ibid* 148.

²¹ *Ibid* 149-52. Cf Perry RR 77-9.

²² ‘The Grounds of Responsibility’ in J Raz and P Hacker (eds), *Law, Morality, and Society: Essays in Honour of H L A Hart* (Clarendon Press, Oxford, 1977) 175, 180, 184.

²³ Cf H L A Hart, *Punishment and Responsibility* (Clarendon Press, Oxford, 1968) 133-134.

be capable of bringing to that person's attention concerns that are relevant to a contemplated act. It seems to be just a fact about human beings that we tend to think about concerns that matter to us both when we contemplate acts to which they are relevant, and, unless we have to decide with such haste as to preclude contemplation, in the course of acting. So lack of thought can indicate lack of adequate concern.

Is it, however, possible for an agent to choose the things she cares about so that they will shape her spontaneous action and bring to her attention concerns that are relevant to her contemplated acts? I do not think the choice can be made out of nothing. There has to be a natural ability to feel the force of the relevant reasons for action or an ability to respond to appropriate learning. Further, there are some concerns, such as being free of pain, that we cannot just choose not to have. But it is open to an agent to structure her concerns, to set priorities, both generally, and for common types of case, to train herself in habits, automatic responses, and spontaneous behaviour, that reflect these choices, and to train herself not to forget things she cares about when they matter. We do this in many ways, but it is worth noting that a person of any integrity does it, *inter alia*, when she evaluates conduct, not only her own but that of others, for she cannot apply standards to others that she is not willing to act on herself. Thus, in a broad way, inadvertent negligence can be seen as the result of choice.

Capacity

Honoré's worry about capacity is whether, if physical determinism is true, we can ever say that anyone is responsible or at fault.²⁴ His response is to distinguish between 'can' (particular) and 'can' (general).²⁵ Can (particular) requires that if a person tries on a particular occasion they will do the relevant thing; it is inconsistent with 'I tried but I couldn't do it'.²⁶ He gives more than one description of can (general), but the most common is that it requires that a person normally succeeds in doing an act of the type in question when he tries.²⁷ He uses the example of a golfer and a six-foot put. If the golfer can normally sink a put of this sort if he tries, then he can (general) sink such a put; but if he tries on an occasion and misses, then he could not (particular) sink the put on that occasion. Honoré believes that if physical determinism is true it must be false that a person could (particular) have acted other than he did on an occasion: 'given all the factors, external and internal, that were present.'²⁸ But physical determinism is not

²⁴ RF 11.

²⁵ *Idem*, also RF 37-38, 139, 158-160.

²⁶ RF 143-144 ('Can and Can't').

²⁷ *Ibid* 145.

²⁸ RF 15, 38.

incompatible with saying that a person could (general) have acted otherwise than he did.²⁹ All that is needed for could (general) is that the person would normally have succeeded if he had tried to act in a different way, which is compatible with physical determinism. He believes that could (general) is sufficient to ground responsibility and fault.

If physical determinism is true, there is, in principle, a possible physical description of an agent's action under which, given the antecedent conditions that then existed in the universe, including the state of the agent's physiology, the action could not have been otherwise. Even if physical determinism is not true, whatever probabilistic patterns operate at the level of sub-atomic particles are probably sufficient to yield a description under which, to all intents and purposes, the action could not be otherwise. So it seems we are faced with either denying that the causal laws revealed by science (deterministic or probabilistic) fully operate within the human body or accepting that there is a possible physical description of any action under which the agent could not have done otherwise than he did. As Pettit points out,³⁰ an account of the capacity to have done otherwise that purports to ground responsibility will not be plausible if it denies the understanding of the world provided by science, but if it accepts that understanding it then has the problem of showing why it is justifiable to praise or blame agents for their acts.

I do not think the solution Honoré suggests works. It is compatible with the scientific view of the world, but it fails to show why it makes sense to praise or blame agents. For the mere fact that an agent can normally perform an act does not imply that she acted freely and can fairly be blamed, when, on an occasion, she fails. Suppose a champion darts player can hit the triple-twenty spot 19 times out of 20. I think we would have to say she can normally hit the triple-twenty. But that does not imply that when she tries and misses on the 5 per cent of occasions on which she does, she acts freely and can fairly be blamed.

At times Honoré uses a different account of can (general) under which an agent can (general) do an act only if he can certainly do it unless there is something unusual present that prevents him.³¹ This account doesn't work any better. For whether such an agent is responsible for a failure depends upon what it is that prevents him succeeding. In *Harding v Price*³² the trailer of an articulated truck driven by the defendant clipped a parked car, but owing to the noise of the truck he did not realize the accident had happened. He was acquitted of failing to report the accident to the police.

²⁹ *Idem.*

³⁰ RR 21-25.

³¹ RF 147.

³² [1948] 1 KB 695.

The defendant wasn't to blame for failing to report the accident, although no doubt he was able to notice a motor vehicle accident when he had one, unless something unusual was present. But perhaps this ability is not specific enough. Suppose, then, that on most occasions on which the defendant had been involved in a similar accident in this truck he would have been able to hear the accident, but by chance the truck hit a pothole at the relevant time, causing a noise that masked the sound of the accident. Would it then have been fair to blame him? If a defender of Honoré here insists on even more specificity, he had better indicate where he is going to stop, for otherwise he will make it impossible for any circumstance to count as unusual. One possible answer might be that all external circumstances should be rejected, and only internal conditions of the agent should count as unusual circumstances. However, this doesn't help. Suppose, if you like, that the noise of the truck had brought on temporary deafness, or had put the defendant into a catatonic state in which he couldn't adequately distinguish stimuli. Unless the defendant had some reason to know of his condition this surely makes no difference. The problem with Honoré's account is that it relies on a general capacity the presence of which doesn't establish that the agent could have acted differently in the particular case.

Alternative solutions to the problem appear in the essays by Smith and Pettit, but in my view neither solution succeeds. Smith's essay is mostly about weakness of the will and how we can understand the capacity for rational self-control that is present when an agent is guilty of weakness of the will. (His example is John, who cannot resist buying chocolate, although he knows it is bad for his health.) Having given an account of the capacity, Smith turns to consider the difference between an agent who has the capacity, but fails to exercise it on an occasion, and an agent who lacks the capacity.³³ He suggests it lies in the fact that the possible world in which the first agent succeeds is less different from the possible world in which he fails than the possible world in which the second agent succeeds is from the possible world in which he (the second agent) fails.

Smith's approach suffers from a defect similar to Honoré's. It doesn't matter how close is the world in which an agent fails to exercise self-control to a possible world in which he succeeds, this is still compatible with the agent not having the capacity to exercise self-control on the occasion. Consider again the defendant in *Harding v Price*. Suppose him to be a conscientious person who would certainly have reported the accident if he had been aware of it. The world in which his truck hit a pothole causing a noise that obscured the sound of the accident is surely very similar to the world in which it did not and he consequently reported the accident. But this doesn't imply that he had the capacity to exercise self-control, and thus make a choice to report the accident, when his truck did hit the pothole.

³³

RR 14.

Pettit argues for what he calls ‘an agent-centred approach’, rather than ‘an act-centred approach’, to explaining the required sense of ‘could have done otherwise’. An act-centred approach, he says, focuses on the way the act was generated; while an agent-centred approach ‘bears on the sort of agent that X more generally is’.³⁴ Early on, his account looks very like that version of Honoré’s account of could (general) in which it means that the agent could have done the thing if something unusual had not intervened. He says that when we say of an agent who has failed to observe a standard that she could have done otherwise, ‘we are saying that this failure was not typical’.³⁵ He adds that there need not be any point in the process leading to the action ‘where we can see the presence of something we might describe as volition’.³⁶ Among features of the circumstances that may lead to atypical behaviour he includes ‘a glitch in the way the agent’s memory worked’,³⁷ which, of course, may be due to a cause entirely outside the agent’s control, such as a blow on the head or an impending illness. If these remarks genuinely reflect Pettit’s position, it is open to the same criticism as Honoré’s: an agent may have the required general capacity yet not have been able to act otherwise on the occasion. This is so whatever may be the standards that Pettit is envisaging the agent must (in general) be capable of satisfying.

Later, however, Pettit says that the required capacity must be present at the moment of acting.³⁸ If that only means that it must be true of the agent at that moment that he has the required general capacity – as it was no doubt true of the defendant in *Harding v Price* at the moment of the accident that he was usually able to notice a motor vehicle accident when he had one – nothing has changed. If, however, the idea is that the act itself must result from the relevant capacity, it is hard to see how the ‘agent-centred approach’ differs from an ‘act-centred approach’: in both cases we are trying to explain how the agent’s condition at the time led to the particular act.

In my view the sense of ‘could have done otherwise’ required to justify blame is unavoidably ‘could have done otherwise on the occasion’. Further, since an agent might always have done otherwise by chance, the force of the remark is not fully carried by its literal meaning, but lies also in the suggestion that a choice, or choices, of the agent could have made things different. The issue now is whether this is reconcilable with the scientific understanding of the world.

³⁴ RR 26.

³⁵ *Idem.*

³⁶ *Idem.*

³⁷ RR 27.

³⁸ RR 33.

As many people have pointed out, the strongest evidence in favour of freedom of choice is personal experience: we know that we have a free will. We might sensibly ask, then, what constitutes the experience of free will. I will begin with cases of deliberated choice. What constitutes free choice, in a case of this sort, is deliberating about the options, considering their implications, feeling the force of the conflicting reasons, deciding, when we want to, on a favoured course of action, and then carrying it out for the reasons favoured. The key to the experience of freedom in such a case seems to lie in two things. The first is that no proposition believed true of an act precipitates that act automatically. We decide which of those things we are inclined to want that such propositions draw to our attention we will favour. When we act, we act knowing that we are favouring some things we want against others, and knowing we can at any moment act differently if we favour differently. The second is that such contemplation is not automatically cut off at any point. We decide when we will decide, knowing, again, that we are favouring doing so over continuing to deliberate and that if we favour differently we can proceed differently.

There is another aspect to freedom of choice that is not part of the experience of freedom of choice on an occasion, but is rather part of the experience of such freedom across a life. We know that our ongoing contemplation about how we should act, and the goals and commitments we adopt in the light of it, will shape both our actions and our practical deliberations. For our standing goals and commitments shape our spontaneous actions and the manner in which we carry out deliberated acts, and they also influence the range of considerations that will come to our attention when we deliberate about prospective acts. Further, the potential range of such ongoing contemplation is not constrained. The point here is not just that we are free to follow up different lines of thought as we choose, it is also that we have access to layers of thought about reasons. We are capable of considering the merits of acting on any particular reason, and of taking our view about that as itself a reason for conduct. Since this ability iterates, there is no logical limit to the range of such contemplation that we can undertake if we choose.

It is by the elaboration of such points that we can hope to show that, whenever blame is fairly levelled, choice that lay within the power of the agent could have made a difference. Is there anything in all of this that is contrary to the scientific view of the world? I do not think so. All we need to believe is that organisms that operate according to causal (or probabilistic) laws can be capable of such complex structures of thought and consequent action. That a human being deciding and acting is a physical structure operating according to causal laws need not trouble us, so long as the physical structure is sufficiently complex to constitute a human being *freely* deciding and acting in the manner described above. We need,

of course, an account of how such physical creatures can be capable of thought and feeling. But I think there are prospects of that. If we can develop such an account, I see no reason why the two perspectives here – that of the experienced internal life, and that of the scientific description of the physical world – should not prove to be compatible.

Responsibility

I will here consider the arguments that Honoré advances that might support the claim that ‘outcome responsibility’ should be allowed at least some of the consequences more traditionally preserved for agency responsibility. The relevant consequences are: (a) it justifies our receiving credit and discredit from others in our society, (b) it is essential to our status as a person, (c) it justifies obligations to do something to set things right.

Outcome responsibility is fair

This argument takes different forms in different places. The first is that a system of outcome responsibility will be fair if all those affected participate as equals and if each stands the prospect of benefiting from the system more than she suffers.³⁹ The idea here is that the system should only apply to those with a minimum capacity to calculate the potential outcomes of their conduct, and that such people are likely to receive more credit from the system than discredit. Honoré places some stress on the fairness of people being responsible for outcomes they could have anticipated, so it can seem that the point being made is just that it is fair to hold people responsible when the risks they chose to run come home to roost. But that is not the point. The system undoubtedly includes unmerited credit and discredit.⁴⁰ Since credit and merit could, in principle, be restricted to acts (not actions) believed to depend on choice, the idea has to be that for persons of minimum capacity the system is likely to produce unmerited credit that will amply compensate for any unmerited discredit. This, surely, is a strange idea. For anyone of self-respect, to receive unmerited esteem can hardly be a compensation for anything, let alone for the wrong done when we receive unmerited resentment. The crucial fact here is simply that praise or blame (for that is really what we are talking about) that is not related to a choice that either was made or should have been made is arbitrary. Why should we esteem someone for an outcome she did not pursue? And how could we fairly resent someone for an outcome she could not reasonably have been expected to avoid? Praise and blame are not commodities that we can

³⁹ RF 24-8.

⁴⁰ See eg RF 28.

shuffle into convenient configurations for social ends: they are intrinsically related to choice.⁴¹

I am aware that we may praise the weight-lifter who lifts the heaviest weight, even although we are not confident that he tried harder than the person who came second. But the truth is we never really know what is the cause in cases like this: whether, for example, the training was better conceived, more conscientiously carried out, or involved greater endurance. So in such cases praise tends to get mixed up with something that is certainly different: the admiration of a body that is, in some respect, well formed. This does little harm, so long as the praise and esteem are kept within sensible bounds.

A second argument Honoré puts is that it is fair to make a person who will benefit from an uncertain situation over which she has some control bear the losses that may flow from that situation. The idea is that the outcomes of any action are uncertain, and since a person has control over whether, and how, she acts, it is fair to make her responsible for all the outcomes of her voluntary actions.⁴² We need to question what 'responsible' means here. If it means agency responsible, then, as we have just noted, it is not fair to hold a person agency responsible (ie blamable) for all the adverse outcomes she may cause by her voluntary actions. The fact that she stands to benefit from these actions makes no difference: there are outcomes we may cause that we could not reasonably guard against, and to require us not to act at all because of this general possibility would be absurd. If, however, 'responsible' means legally liable, then, although the general principle Honoré cites is not sound, perhaps sometimes a person who engages in an unusually hazardous activity should be liable for all the harm of certain types caused by this. Whether such strict liability is ever justified in civil law, and if so when, requires careful thought. But whatever we say about this, it has nothing to do with blame. As we have already seen, the whole point about strict liability is that the person who engages in the hazardous activity is not blamable for doing so, and is not necessarily blamable for an adverse consequence that results. She may be blamable if she then fails to compensate, but that is a different matter.

A third argument is that it is fair to place on every member of the community the burden of bearing the risk his conduct may turn out to be harmful to others, in return for the benefit he will receive if his conduct turns out as he plans.⁴³ The problem here is that society does not confer the compensating benefit: it is simply a consequence of being able to act on one's goals. At most society can protect this ability, or enhance it by

⁴¹ Cf Perry RR 64.

⁴² RF 76-80.

⁴³ RF 80.

making possible goals that would not otherwise exist. Whenever it does so it must create 'burdens' in order to provide the protection or enhancement. If it acts even-handedly it will then be able to claim that for the burdens of its scheme each person receives a greater advantage. But it will be the benefits and burdens incident to the scheme that need to be considered in assessing this claim: one cannot bring in extraneous benefits, such as the benefit every normal person has of being able to act to secure ends.

We may need to respond to outcomes we produce even if we are not at fault

Honoré relies extensively on the fact that, by moral or social understanding, we may often be expected to respond when we cause harm even if we were not at fault in doing so. 'If purely by your fault in darting out into the road I run over you', Honoré says, 'I must stop, send for the ambulance and give you what help I can in the meantime.'⁴⁴ Again: 'We can be responsible for what we do by sheer accident, like unavoidably tripping someone. .. An apology is called for, and the person who has been tripped must be helped up and if necessary taken for treatment.'⁴⁵

These are certainly interesting phenomena that require an explanation. But in my view the explanations are particular, and do not establish that in general outcome responsibility for a harm yields an obligation to do something to remedy the harm. The reason we should apologise when we unavoidably trip someone has to do, I think, with the fact that, in real life, occurrences are never as clear-cut as they are in philosopher's examples. An apology in such a case serves to prevent potential hostility: it indicates that no intent was involved, and that if fault of some sort was a cause the agent is sorry for it. It also indicates a willingness to accept authorship, so that if questions of fault arise they can be addressed. Honoré does not say that the driver should apologise to the pedestrian who has darted out in front of him. It is clearer that the driver was not at fault and an apology might carry an inappropriate implication that he was. Still, there is an obligation to give help. The source, I think, is that it is clearly valuable to all of us if those involved in an accident accept such obligations, regardless of fault. Fault can be sorted out later; what is immediately needed is help. The driver is on the spot and able to help. Let all who are causally involved in an accident causing personal injury, and on the spot able to help, accept an obligation to help without fussing about fault, and we all gain.

⁴⁴ RF 31.

⁴⁵ RF 127.

Detailed reasons of social utility such as these perhaps often create an obligation 'to do something to put it right'⁴⁶ when one causes a harm without fault, but I don't think there is always such an obligation. This is not because a *prima facie* obligation to do something to put it right can be outweighed, but because there is no *prima facie* obligation. Unless there are detailed reasons for a particular obligation there simply is no obligation. If, driving innocently along a road I cause a subsidence that is not dangerous but will need to be repaired sometime by the local authority, I have no obligation to do anything.

It is different when someone is at fault for causing harm. When harm to someone else is an agent's fault, there is, on that account, an obligation falling on the agent to take remedial action or to compensate. The broad thought is that since the agent could, through choice, have prevented what has happened, the agent should set it right. Exactly what should be done will depend on the type of harm, and, perhaps the obligation may be overridden; but *prima facie*, at least, the agent should remedy the wrong. No such thought is appropriate for outcome responsibility, because mere causation does not imply that the agent could, through choice, have done otherwise.

But it is not just that there is in general a *prima facie* obligation to remedy a wrong that distinguishes the cases. The obligation to remedy a wrong is a genuine *remedial* obligation in that it arises out of the breach of a prior obligation – an obligation not to cause the harm. In contrast, the obligations that sometimes exist to do something to put it right that arise when one causes harm without fault are not genuinely remedial. They do not arise out of the breach of a prior obligation, but only because there is reason for a primary obligation to come into being in these circumstances.

Outcome responsibility is essential to our status as persons

Outcome responsibility, Honoré contends, enables us to 'own' certain occurrences in the world. It thus gives us a personal history.⁴⁷ He contends that if human beings did not share understandings that enabled them to identify some of the states of affairs produced by human bodies as being caused by them as agents, rather than just as bodies, people could have no continuing history or identity as persons. The requisite understandings, he says, are those that identify all the outcomes caused by voluntary actions, in his wide sense of 'caused', as the product of human agents.

Let me admit some points. No doubt if we could not identify some actual outcomes as our product we would lack much of a personal history –

⁴⁶ RF 133.

⁴⁷ RF 10, 29, 131-3; RF 223.

although, in principle, we could still identify what we had tried to do. Again, if others didn't identify those acts for which people are agency responsible in much the same way that we did, that would certainly be disconcerting. Again, to recognise outcomes as our product we no doubt need, in at least most cases, to use some notions of causation. But here the admissions can stop and the demurrer begin. We can have a perfectly adequate personal history without all the outcomes we cause by our voluntary actions being ascribed to our agency. It is sufficient that we are deemed agents of the outcomes we choose, although a richer, and better, conception of agency will also include those outcomes we could reasonably have avoided. In contrast, widespread use of a conception of agency that attributed to people's agency outcomes they did not choose and could not reasonably have avoided would, it seems to me, be likely to damage both our sense of self-worth as individuals and the quality of our understanding of others.

Gardner adds an additional argument to Honoré's claim that outcome responsibility is essential to our status as persons. If I have understood it correctly, it can be stated in the following propositions:

1. Reasons to succeed are logically prior to reasons to try, because a person cannot have a reason to try to do something unless he has a reason to succeed.
2. Consequently, if success is, across the board, insignificant in the account of our lives as rational agents, trying is also insignificant.
3. Therefore, to deny that success can have independent significance, beyond that of mere attempts, in the account of our lives as rational agents, is to leave us without any significant account at all of our lives as rational agents.

As it happens, it is not quite true to say that a person cannot have a reason to try to do something unless he has a reason to succeed. Consider John who has jumped 1.3 metres, and will get into the inter-school sports, a prospect he abhors, if he can jump 1.5. John's father wants him at least to try, and offers a monetary reward if he sees him trying on a particular afternoon. John knows his father is very good at spotting pretence, and won't pay up if he does. John wants the reward and reluctantly sets about trying as hard as he can, although at the same time he hopes mightily that he won't succeed. John does not want to succeed for any reason at all, but he does want to try. This point doesn't matter greatly, for the case is an odd one: remove the point of success in most cases and you remove the point of trying. Gardner's significant mistake is not here, but in imagining that point 1 establishes that the difference between succeeding and trying is necessarily important in the account of our lives as rational agents, if that means important in assessing the merit of those lives. Mostly if we want to do anything worthwhile we must act for reasons that are reasons to succeed.

But if, despite the best of efforts, we fail through sheer bad luck, that doesn't change one wit the merit of our actions. No doubt we often do judge success as more important than attempts, but that is because often when we succeed we have had to overcome setbacks and try again. We all understand the lesson of Robert the Bruce. That doesn't help Gardner, for his point is that the bit that success adds to attempts, even when it is due purely to luck, adds intrinsic merit to what we are – not what we do – and that unless this were so, attempts themselves would lack merit. I can't think that is right.

Similar points apply to the difference between attempts and success in wrongful acts. The person who shoots to kill and misses is just as morally culpable as the person who, with an equal prospect of success, shoots to kill and succeeds. It is true, as Honoré points out,⁴⁸ and Gardner notes,⁴⁹ that we judge murder more severely than attempted murder, and causing death by dangerous driving more severely than dangerous driving. But, there are reasons for this distinct from the moral quality of the acts. Firstly, as Anthony Kenny has pointed out, we need to ensure that the person who has shot and missed is not left without any motive, so far as punishment goes, not to try again.⁵⁰ Secondly, for better or worse, the criminal law is not only concerned to fit the punishment to the moral quality of the crime, but also to temper the feelings of resentment and hostility that crime engenders. These are likely to be greater for successful crimes than attempts and greater when the defendant's conduct has caused actual harm not just danger.

Gardner's argument is set within a broader argument in which he tries to establish that there can be obligations that are, quite strictly, obligations to succeed. I don't want to discuss the details of Gardner's argument; but I do want to comment that duties to succeed, no matter what, seem to me to make no sense, except as a form of shorthand. Legal systems often do, of course, set out legal duties as duties just to do things – that is, to succeed. That is convenient enough, so long as the duties are ones that normal people can comply with under normal circumstances. When a person who could not have complied, through choice, with such a requirement breaches it, we may if we want call this a 'breach of duty': but we can't, without irrationality, conceive of it as a case in which someone who has breached a duty could, through choice, have complied with it. We may treat it in the same way, and there can be comprehensible reasons for doing so;⁵¹ but without irrationality we can't treat it in that way *because it is such a case*. That is why, as von Wright pointed out nearly 40 years ago, so long as 'punish' is used in its normal sense, it is impossible to punish

⁴⁸ RF 31.

⁴⁹ RR 134.

⁵⁰ Kenny, above n 15, 158-60.

⁵¹ See H L A Hart, *Punishment and Responsibility* (Clarendon Press, Oxford, 1968) 18-20.

someone for breach of a requirement if he could not have complied with it.⁵² We can inflict a suffering on such a person, but it cannot count as punishment. This is because to punish someone is not just to inflict a suffering on him, but to do so for breach of a standard it is believed he could have chosen to comply with. So the criminal defences based on lack of rational capacity, as well as those based on lack of any opportunity to choose to comply (as in *Harding v Price*), should not be conceived as a limitation imposed on a system of punishment; they are intrinsic to such a system, so long as our concern is only to punish for breaches of the law. For that reason it is often useful to limit the idea of a breach of the law to cases in which these defences do not apply. The wider usage noted above need not be proscribed; but it is important that it doesn't lead us into confusion.

I should add a comment about strict liability in tort law. It is, I think, a mistake to treat the compensation awarded in such a case as awarded for breach of a prior duty, as Gardner attempts to do. We think most clearly if we treat the only duty existing here as being the duty to compensate if the event occurs for which there is strict liability.

Causation

I want to make only one point about causation.⁵³ It draws on the account I gave in the section 'Choice and Fault' of the reason negligence is a form of fault. The key ideas were that conduct can be caused by the lack of a concern that if present would have prevented it, and that the lack of a concern can be caused by the way an agent has shaped herself. So when negligence occurs we can say that an agent could have understood a risk of her conduct and acted differently as a consequence. The understanding might have been manifested in spontaneous conduct undertaken without measured contemplation, or it might have been manifested in such contemplation and subsequent choice. Either way, the understanding that could have modified the agent's conduct would need to be of an outcome that the conduct stood a risk of producing. Recall that reasons that can shape conduct ahead of time are propositions believed true of types of acts. The harmful outcome must also be a *type* of outcome, since it is not possible for an agent to anticipate ahead of time every detail of the concrete outcome that will exist if the risk of harm is realized. So now we have a clear idea of the sort of proposition that, when negligence has occurred, we

⁵² Von Wright, *Norm and Action*, above n 15, 114-15.

⁵³ Anyone wanting a clear introduction to the dispute between Honoré and Stapleton about causation might usefully read: (1) Honoré's account in RF 1-7 of the evolution of the ideas in Hart and Honoré, *Causation and Law*; (2) Stapleton's very clear essay in RR 145-85; and (3) Honoré's reply in RR 232-37.

think should have modified the defendant's conduct, but did not. It is a proposition that recognises the potential of a *type* of act to produce a *type* of harmful event. Given that negligence will often consist in the failure to be influenced by side-constraints, the type of act that carries the danger may be complex. It may be a member of one type (the type of act wanted) that is not also within a sub-type of that type that avoids an unacceptable risk of the harm. Still, the basic pattern will always be that a type of act has the potential to produce a type of harm.

These reflections enable us to see the proper role of causation in the understanding of negligence. For if we ask why an agent should be capable of understanding that a type of act may produce a type of harm, the most common answer will be that the agent should understand that acts of that type cause, or tend to cause, harms of that type. But that is something that the agent can understand only if the proposition expresses, or is implied by, a causal law that the agent understands. The causal law need not be a law of exact science: it may be a simple common sense understanding. That human bodies tend to be hurt when hit by heavy, hard objects will do. It may be a generalisation that expresses a chance not a certainty, indeed it normally will be. But it must have the pattern of a causal law.

One important distinction that can be made within the use of causal language is between its use looking ahead and its use looking back. When we look back we may be able to see that in the circumstances that occurred some event played a part in the occurrence of some other event. We may call it a cause. But it doesn't follow that it could have been seen as a cause looking forwards. I hail a neighbour as he walks down the street and he stops to talk. Ten seconds later he is hit by lightning. Looking back, we can say my hailing him caused him to be struck by lightning, but we could not say ahead of time that hailing people causes, or tends to cause, them to be struck by lightning. A generalisation is needed even for backward-looking statements: here it is that people who are standing in a place where lightning strikes will be hit by it. An explanatory generalisation is needed, because not everything that is believed to have played a part in producing an event did so. That my neighbour's young son had angry thoughts about his father the night before played no role in my neighbour's misfortune, even if the son believes it did. But the generalisation used here to support the backward-looking causal statement, although true at an earlier time, could not have helped prediction – except for somebody who knew where lightning would strike. It is causal laws that people can utilize ahead of time that are relevant to negligence.

A consequence that follows from these remarks is that causation and foreseeability should not be seen as two independent elements of negligence, but as parts of a single element. The harm must have been foreseeably caused by the defendant's conduct: that is, caused in the

foreseeable way. When an agent could, and should, have foreseen a harm, and consequently modified her conduct, but does not, she creates a risk of a type of harm that she should not create. If the risk comes to fruition, that is, if harm of that type occurs in the way that she should have foreseen as possible, she has wronged the person harmed, and should compensate that person. The duty, however, only extends to harm of the type that could have been foreseen which is caused in the foreseeable way. For only such harm is within the contemplation that could have led her to act differently.

Conclusion

The difficulty with Honoré's notion of outcome responsibility as an account of agency responsibility is that it connects the conduct of an agent to the agent's choice in only a limited way. For an agent to be responsible for an outcome Honoré requires only that the outcome be caused by an action (a concrete performance of some sort) that was intentional under some description. The outcome itself need not have been intentional, foreseen, or even foreseeable. This is too slight a connection to choice to justify agency responsibility.

I have argued that for an agent to be blamable for producing an outcome it must have been possible for him to avoid the outcome through choice that responded to the possibility that he might produce such an outcome. The words 'such an outcome' are important here, for the choice we think the agent could have made must utilize the same description of a type of act as that under which his conduct is considered wrong. The term 'through choice', used here, is deliberately broad. It is intended to cover not only deliberate choice, but also spontaneous action and inadvertent negligence. The latter two can be blamable because they can in various complex ways result from a choice or choices made by the agent.

Is this account of the conditions under which we consider it fair to hold an agent blamable compatible with a scientific view of the world, given that such an account may imply that as a physical entity the agent could not have acted otherwise, or could only have acted otherwise through accidents having nothing to do with her choice? My argument on this point was that there is no inherent inconsistency between such an account and it being true that the agent was at the relevant time, and had been through a lifetime, a human being enjoying the full freedom of choice that a normal human being enjoys. That may seem to leave a lurking question that I have not tried to address: even if we allow the type of connection to choice that I have argued for, is the whole process of blaming agents for acts justifiable, given that it may be true that in some sense they could not have been otherwise than they are? I don't have a tidy answer to that question, but let me note that there is another sense in which normal human agents guilty of

blamable conduct could certainly have been otherwise than they are. Maybe when we can fully understand this second sense, we will understand what we are doing when we blame agents for acts relevantly related to their choice, why we do so, and what sort of question it is to ask whether this is rational.