GUILTY MIND OR GUILTY BRAIN?
CRIMINAL RESPONSIBILITY IN THE AGE OF NEUROSCIENCE

(Article originally published in 2000 *The Australian Law Journal 74*, 661-80)

DAVID HODGSON*

Current developments in the sciences of the brain and mind sometimes seem to suggest that criminal conduct is a symptom of brain disorder or illness that should be treated rather than punished. This paper argues that the insights of these sciences should be taken very seriously by lawyers, but not to the detriment of common-sense ideas of responsibility or of their incorporation into the legal categories used in the criminal law.

.....

Our system of criminal justice is based in various ways on common-sense ideas of free will and responsibility for conduct, according to which it is fair and therefore just that offenders should be punished to an extent that is in some sense *proportionate* to their guilt. These ideas are called into question by ongoing developments in neuroscience, a word I will use in this paper to refer generally to the various sciences of the brain and mind - neurophysiology, cognitive science, artificial intelligence, psychology, psychiatry, and so on - developments which in a general way tend to suggest that criminal conduct is a symptom of a brain disorder or illness that should be *treated*, rather than a wrongdoing that should be *punished*.

In this paper, I will argue that common-sense ideas of free will and responsibility remain valid and respectable despite assertions to the contrary by some philosophers and neuroscientists; although I will also suggest that the insights of neuroscience should be taken very seriously by lawyers concerned with criminal responsibility, and should be used to refine and improve common-sense ideas and legal categories, and to inform decisions as to how best to deal with offenders. I do in fact accept that there is a role for the *therapeutic* approach to crime, in relation to both crime prevention and punishment; for example, in early identification of risk factors and taking steps to minimise them, and in treatment and rehabilitation of offenders. But I contend that this should not be to the exclusion of common-sense notions of responsibility, or of their incorporation into the legal categories used in the criminal law, or of the moral considerations that should underpin the law.

It is perhaps presumptuous for someone who is neither a practitioner in criminal law nor a scientist of any kind to write about how we as lawyers should respond, in our approach to questions concerning criminal responsibility, to current developments in the sciences of the brain and mind. However, for reasons I will give, this is certainly a matter lawyers should be thinking about; and my interest in the philosophy of mind has led me to pursue this particular interface between law and science.[1]

1 Personal Responsibility and Criminal Justice

I will start by noting some respects in which our system of criminal justice is based on ideas of personal responsibility for conduct. In particular, this can be seen (1) in the general doctrine of *mens rea* or guilty mind, (2) in the mental element involved in the very definition of many offences, (3) in various defences based on factors affecting *mens rea*, and (4) in the principles applied in determining sentences to be imposed.

1.1 *Mens rea*

A key principle of our criminal law is the principle of *mens rea*: generally, no person can be convicted of a crime unless the prosecution proves not only a guilty act, but also a guilty mind. For criminal responsibility, the action in breach of the law must be *conscious* and *voluntary*; and there may be additional requirements, such as that there be *intention* of a particular consequence, or *recklessness* as to this consequence. In the case of murder, for example, generally the act causing the death of the victim must be done with intention to kill or inflict really serious bodily injury, or with reckless indifference to human life.

Because of this requirement of a guilty mind, a person will generally not be guilty of any crime if the act in breach of the law occurs *independently of the will* or *by accident*; or if the person mistakenly *believes* the facts to be such that the act would not have been in breach of the law.

1.2 *Mental elements of crimes*

The requirement of a guilty mind is further illustrated by the mental elements involved in the definition of particular offences, particularly matters of belief and intention.

In cases of theft and other offences involving misappropriation of property, it has to be proved that the accused *did not believe* that he or she had a legal right to take the property, and *did intend* to deprive the owner of it. In cases of fraud, it has to be proved that the accused *did intend* to induce someone to believe something which was untrue and which the accused *did not believe* to be true, and thereby to induce that person to part with property.

Perjury involves not merely giving untrue evidence on oath, but also that the accused *did not believe* it to be true. Bribery involves not merely a payment or offer of payment, but also that it was done with the *intention* of inducing the recipient to act contrary to duty. Blackmail involves not merely making demands with menaces or force, but also that this be with the *intention* of obtaining property. Rape requires proof that sexual intercourse took place *without the consent* of the woman, and that the accused *did not believe* there was consent.

1.3 *Defences*

There are defences which, while not negating the voluntariness of the act in question, are regarded as negating or mitigating the guilty mind or criminal responsibility.

Duress (in cases other than homicide) can negate mens rea if the act was done *under the influence of* a threat of death or really serious injury. Self-defence can negate mens rea if the act was done *in self-defence*, in the *belief* held on reasonable grounds that what was done was necessary having regard to the threat. Necessity can negate mens rea if the act was done *to avoid irreparable harm*, in the *belief* held on reasonable grounds that what was done was necessary having regard to the danger.

Then there are the defences of insanity and automatism. Because these defences are so pertinent to a consideration of *mens rea*, I will look at them in a little more detail.

A useful statement of the relevant legal principles by the High Court of Australia is found in *The Queen v Falconer*,[2] in the joint judgement of Mason CJ, Brennan and McHugh JJ. They note the principle from *Woolmington v DPP*[3] that 'the prosecution bears the ultimate onus of proving beyond reasonable doubt that an act which is an element of an offence charged was a willed act or ... was done voluntarily'; but they observe that it is generally not necessary for the prosecution to lead evidence going specifically to mens rea, because of two presumptions. First, there is a presumption of fact, which they state as follows:

> It is presumed ... that an act done by a person who is apparently conscious is willed or done voluntarily. That presumption accords with, and gives expression to, common experience. Because we assume that a person who is apparently conscious has the capacity to control his actions, we draw an inference that the act is done by choice. ... The prosecution may rely on [this] inference ...

> to discharge the onus unless there are grounds for believing that the accused was unable to control [the] act.

Second, there is the presumption of law stated in *M'Naghten's Case*[4] that 'every man is presumed to be sane, and to possess a sufficient degree of reason to be responsible for his crimes, until the contrary be proved'.

What is meant by the presumption of sanity being a presumption of *law* is that it will prevail unless insanity is affirmatively proved on the balance of probabilities; whereas the presumption of voluntariness will, as a presumption of *fact*, be displaced if the evidence merely goes far enough to raise a reasonable doubt whether the action was voluntary. As stated in *Falconer*:

> An accused bears no ultimate onus of proving that his act was unwilled, but he does bear an evidential onus of rebutting the presumption that he had the capacity to control his actions and, if he chooses to discharge that onus by showing that he was not of sound mind, he must prove that proposition on the balance of probabilities.

The test for insanity was laid down in *M'Naghten's Case* in the following terms:

> To establish a defence of insanity, it must be clearly proved that, at the time of committing the act, the party accused was labouring under such a defect of reason, from disease of the mind, as not to know the nature and quality of the act he was doing; or, if he did know it, that he did not know he was doing what was wrong.

This test does not in terms extend to a lack of capacity to control one's actions; but it may do so in practical effect because it is 'extremely difficult to conceive of a state of things in which a person acting automatically and without any exercise of the will would nevertheless know what he was doing and that it was wrong'.[5]

As is well known, the result of a successful defence of insanity is that the accused is found not guilty by reason of insanity, and is then detained in custody for an indefinite period at the discretion of the government, for the safety of the public and for treatment. On the other hand, the result of a successful defence which merely denies that the act was conscious and voluntary (the defence of sane automatism) is an outright acquittal; and as we have seen, there is no legal onus of proof on the accused in relation to this defence.

However, as suggested in the second passage quoted from *Falconer*, to counter the presumption of voluntariness referred to above, there needs to be *evidence* which suggests that the accused was unable to control his or her acts; and for this to raise an issue to be left to the jury, that evidence generally needs to provide an *explanation* as to how this came about. Evidence from the accused alone that he or she blacked out and cannot remember the incident is not sufficient.[6] If the only explanation suggested by the evidence involves 'a defect of reason from disease of the mind' (or, as it is sometimes put, unsoundness of mind), then only insanity would be left to the jury. That is, in order that the question of sane automatism be left to the jury, there must be evidence of unconsciousness or inability to control one's actions *which is explained by something other than unsoundness of mind*.

This is strikingly illustrated in the English case of *Reg v Sullivan*.[7] In that case, the accused kicked a man violently on the head and body while suffering a seizure due to psycho-motor epilepsy. He pleaded not guilty to causing grievous bodily harm. He gave evidence, which was not disputed, that he had no recollection of the incident; and two medical experts, whose evidence was also uncontested, testified that it was strongly probable that the attack took place during the third, or post-ictal, stage of the seizure, when the appellant would make automatic movements of which he was not conscious. The trial judge ruled, in the absence of the jury, that they should be directed that if they accepted this evidence it would not be open to them to bring in a verdict of not guilty, but they would be bound to return a special verdict of not guilty by reason of insanity. An appeal went to the House of Lords, and the House of Lords agreed with the trial

judge that the evidence could support a defence of insanity, but not a defence of sane automatism. In the course of his reasons in this case, Lord Diplock said this:

> I agree ... that 'mind' in the *M'Naghten* Rules is used in the ordinary sense of the mental faculties of reason, memory and understanding. If the effect of a disease is to impair these faculties so severely as to have either of the consequences referred to in the latter part of the Rules, it matters not whether the aetiology of the impairment is organic, as in epilepsy, or functional, or whether the impairment is itself permanent or transient and intermittent, provided that it subsisted at the time commission of the act.

That is, for a defence of sane automatism to succeed, it must be that the explanation for the automatism does not involve any unsoundness of mind of the accused person; and epilepsy was held to involve unsoundness of mind.

It may come as a surprise to some (as it did to me) that epilepsy was thus treated by English law as insanity: this is not in accordance with the common-sense understanding of insanity. The moral justification suggested for treating epilepsy as insanity was that, if violent injury has been done involuntarily by reason of a condition which is such that something similar could happen again, the State should be given the means to ensure at least that the person gets the treatment appropriate to minimise the chance that it will happen again.

An explanation for automatism not involving unsoundness of mind may perhaps be given by concussion or anaesthesia or possibly an extreme state of intoxication.[8] In *Falconer*, it was held that a dissociative state arising from psychological stress could be such an explanation, provided the psychological stress was sufficient to so operate on a sound mind and did so operate in the particular case: if the psychological stress could have produced the dissociative state only in combination with some unsoundness of mind, then only the insanity defence would have been available. It was suggested in *Falconer* that automatism arising from hypoglycaemia in a diabetic could be such an explanation. *Query* whether hypnosis could be; or is it the case that a person could involuntarily commit a crime under hypnosis only if that person also had some unsoundness of mind?[9] In an early case, it was stated that sleep-walking may give such an explanation; but in *Reg v Burgess*[10] the Court of Appeal noted that the evidence in that case indicated 'that sleep-walking, and particularly violence in sleep, is not normal'. The Court accepted evidence that the sleep-associated automatism alleged in that case involved a disease of the mind; and accordingly, could give rise to a defence of insanity but not sane automatism.

Finally on defences, I should mention two defences which are treated as *lessening* the guilt of the mind and thus as justifying reduction of murder to manslaughter. One is the defence of provocation, which is available where a killing occurs due to a temporary loss of self-control by the accused person, resulting from a provocation by the deceased sufficient to cause such a loss in an ordinary person. The other is the more recently introduced defence of diminished responsibility, which is available where an underlying condition of the brain or mind of the accused has substantially impaired his or her mental responsibility for the killing.

1.4 *Principles of punishment*

In relation to sentencing of offenders also, notions of *mens rea* and personal responsibility for conduct loom large, in at least two ways.

First, there is the overriding principle that punishment for an offence should not be more than what is considered *proportionate* to the offence itself. In particular, a term of imprisonment for an offence should not be increased beyond what is proportionate, in order to extend the period of protection of society from the risk of further crimes by the offender.[11] The extent of guilt or 'criminality' is an important element in determining proportionality. *Veen v the Queen [No 2]*[12] established that in deciding what is proportionate,

the need to protect society and the antecedent criminal history of the offender may be taken into account along with the gravity of the offence, but not so as to result in the imposition of a penalty disproportionate to the gravity of the offence. In no case can the maximum penalty prescribed by law for the offence be exceeded, and this maximum should be reserved for cases falling within the worst category for that offence.

Secondly, particular factors mitigating *mens rea* or personal responsibility may justify the imposition of a penalty less than that considered proportionate to the gravity of the offence. Thus, genetic and congenital disadvantages, early environmental difficulties such as child abuse, and more immediate matters such as pressing social, psychological, or financial problems, may be taken into account so as to justify a reduced sentence.

Admittedly, some of these factors may not work unequivocally in reducing the penalty. For example, if genetic or environmental factors have resulted in a mental abnormality which makes the offender more prone to violence, this may diminish moral culpability and thus point towards a shorter sentence of imprisonment; but at the same time point towards a longer sentence because of considerations of deterrence (greater penalties may be thought necessary to deter such persons from offending) and protection of society from risk.[13] However, this does not deny the general relevance of diminished moral culpability as a mitigating factor.

2 Responsibility and Free Will

2.1 *Free will presupposed*

Now it seems obvious that the notion of personal responsibility for conduct in turn presupposes that we have free will - that we are not caused to do what we do by matters outside our control, but rather in any situation have real choices between alternative courses of action which are truly open to us, in the sense that, given the situation and even given our own natures and characters, it is truly possible for us to take any one these alternatives. Otherwise the legal doctrine of *mens rea* would make little sense.

If we were wholeheartedly to reject free will and accept determinism, the implications would be far-reaching indeed. As Isaiah Berlin put it:[14]

> [Determinism] may, indeed, be a true doctrine. But if it is true, and if we begin to take it seriously, then, indeed, the changes in the whole of our language, our moral terminology, our attitudes towards one another, our views of history, of society, and of everything else will be too profound to be even adumbrated. The concepts of praise and blame, innocence and guilt and individual responsibility from which we started are but a small element in the structure, which would collapse or disappear. Our words - our modes of speech and thought - would be transformed in literally unimaginable ways; the notions of choice, of responsibility, of freedom, are so deeply embedded in our outlook that our new life, as creatures in a world genuinely lacking in these concepts, can, I should maintain, be conceived by us only with the greatest difficulty.

2.2 *Free will questioned*

Philosophers have long questioned whether we have this sort of free will. In the eighteenth century, David Hume argued that everything that happens must be causally determined by its antecedents or else be a matter of chance; and that a person could not conceivably be responsible for events which happen by chance.

The progress of science since Hume's time has in a number of ways seemed to support this approach. Our brains are apparently physical systems, so it seems that all causation in our brains must be causation of the same kind as operates in the physical world generally. This is apparently confirmed by Darwinian explanations of our motivation, by the large role that unconscious processes play in our choices and actions, and by the understanding of the physical operation of the brain and its functional organisation

which is provided by neuroscience. And so, despite the fact that science has not yet any explanation for subjective experience or consciousness, and despite the instinctive belief in free will which we all have and live by, the scientific orthodoxy is (overwhelmingly) that any apparent exercise of free will is no more than a working out in our brains of the ordinary causation we see in the physical world, and thus is itself the inevitable outcome of its causal antecedents.

2.3 *Compatibilism*

However, at least until quite recently, philosophical and scientific scepticism about free will has not had any significant impact on the law's approach to criminal responsibility, or even been considered as calling for any reassessment of this approach. Acceptance in practice of common-sense views about free will and responsibility, and application of them in our everyday dealings with other people and in the practice of the law, has coexisted quite comfortably with this scepticism about free will.

One reason for this has been a quite widespread philosophical view that determinism is *compatible* with a kind of free will, and is compatible and perhaps even necessary for personal responsibility for action. The basic idea of compatibilism is that human beings have freedom and responsibility just because they are free to act in accordance with their own choices and to do whatever it is they most want to do; and that it does not matter that their choices and their wants may themselves be determined by prior circumstances and impersonal laws of nature.

In the eighteenth century, Hume argued that his version of determinism was not inconsistent with personal liberty (by which he meant what we would call free will), because liberty means our power of acting according to the determinations of our will, that is, as we choose - even though our choice may itself be determined by causes. According to Hume, it is constraint, not determination by causes, which can deprive us of this liberty; and if our actions were not determined by causes, then they could only be a matter of *chance*, so that they could not be an exercise of personal liberty.

Hume went on to argue that his approach was not inimical to our common-sense ideas of responsibility, but on the contrary was necessary to make sense of responsibility. This, he said, was for two main reasons.

(1) A person is not responsible for actions unless they proceed from a cause in the character or disposition of the person. The more premeditated an action is, the more we regard it as caused by the person's character and the more responsible we consider the person to be. On the other hand, we consider people less responsible for actions performed hastily.

(2) The imposition of rewards and punishments for actions for which we suppose people are responsible makes sense only if we believe that rewards and punishments have consistent and regular effects on their behaviour; that is, if we believe that the prospect of rewards for certain behaviour and the prospect of punishment for other behaviour will consistently influence people to engage in the former and to refrain from the latter.

The compatibilist tradition continued strongly in the nineteenth and twentieth centuries. There were two significant elaborations of it in the 1960s.

First, philosopher Peter Strawson argued in a famous essay 'Freedom and resentment'[15] that we do in fact, as subjects and participants in the world, regard ourselves as responsible for our own conduct, and we do in fact generally take the same attitude towards other persons with whom we interact; and that this is an attitude we *cannot help* adopting and maintaining, even if determinism is true, and indeed even if we *recognise* that determinism is true.

Second, there was extensive development of Hume's second point, notably by jurist Herbert Hart in essays such as 'Punishment and responsibility',[16] to the effect that the law's concerns about *mens rea* and responsibility can be given justification which does not require appeals to any notion of free will.

Punishment, it was said, must be justified by its beneficial effects, including its deterrent effect. There should be parsimony in threatening and imposing punishment, and since for the most part it is only voluntary actions which are susceptible to deterrence by threat of punishment, it makes sense to threaten and to apply coercion only in respect of voluntary actions. This, it was claimed, fully explains and justifies the law's concerns about *mens rea* and responsibility.

3 Current Problems

However, the time is coming, if it is not here already, when we won't any longer be able to take this modus vivendi for granted, and when it will be necessary to consider very carefully whether the law's approach to criminal responsibility can be maintained in the light of what science tells us about the brain, or whether it needs to be modified or even radically changed. A number of factors are at work.

3.1 *Scientific attacks on mens rea*

As more and more is discovered about the workings of the brain, scientists are increasingly prepared to claim that common-sense ideas about responsibility cannot be maintained, and that the law should recognise this. Scientific reductionists are becoming more confident that they have the complete answer to all problems concerning our minds and our behaviour, and more strident in their assertions that all other views should be abandoned.

One plain example of this trend is provided by the TV series and book *The Mind Machine* by Colin Blakemore, professor of physiology at Oxford University. In the final chapter of that book, entitled 'The Violent Mind', Blakemore commences with three sentences which he says encapsulate the central thesis of his book:[17]

> The human brain is a machine, which alone accounts for all our actions, our most private thoughts, our beliefs. It creates the state of consciousness and the sense of self. It makes the mind.

Blakemore then sets out a number of cases in which criminal acts were committed under various mental conditions: hypoglycaemia, a sub-lethal dose of chlorpyrifos from handling insecticide, pre-menstrual tension, epilepsy, manic depression, and deprivation and abuse in early life. He then comments as follows:

> Surely no one could believe that violent thoughts or actions committed by otherwise model citizens while their blood glucose is very low or their amygdala is exploding with epileptic discharges were *intended* by the person involved. There would clearly be more debate about a killing or near-killing committed after handling insecticide, or just before a menstrual period, or during a manic, drunken high. And I suspect that there will have to be more clearer scientific evidence before the public or the courts will accept that a violent psychopath who has no overt brain injury and whose behaviour has been unremittingly criminal is merely the victim of genetic predisposition and a difficult childhood.

And his main conclusions are stated in two further passages:

> All our actions are products of the activity of our brains. It seems to me to make no sense (in scientific terms) to try to distinguish sharply between acts that result from conscious intention and those that are pure reflexes or that are caused by disease or damage to the brain. We *feel* ourselves, usually, to be in control of our actions, but that feeling is itself a product of the brain, whose machinery has been designed, on the basis of its functional utility, by means of natural selection.
>
> ... The *sense* of will is an invention of the brain. Like so much of what the brain does, the feeling of choice is a mental model - a plausible account of how we act, which tells us no more about how decisions are really taken in the brain than our perception of the world tells us about the computations involved in deriving it. To choose a spouse, a job, a religious creed - or even to choose

> to rob a bank - is the peak of a causal chain that runs back to the origin of life and down to the nature of atoms and molecules.

Somewhat similar views can be found in the work of philosophers such as Daniel Dennett, and Patricia and Paul Churchland; and in the work of other neuroscientists, such as Francis Crick.[18]

A related expression of the current 'scientific' view of the brain and mind is the thesis of *biological determinism*, popularised in books such as Robert Wright's *The Moral Animal*:[19] the thesis that all our mental characteristics and thus all our actions arise wholly from the biological properties of our brains, which in turn are wholly caused by our genes and the environment in which we have been placed since conception - that is, by nature and nurture. Even though this thesis allows a significant role for environment (nurture), the label 'biological' is apt, because all causation of actions is said to be *mediated* by biology; and the label 'determinism' may also be considered apt, even if some randomness is allowed into this picture, so long as any randomness is regarded as making no significant contribution to causation or explanation of actions. This thesis lays claim to explaining not just our characters and actions, but also all aspects of our morality; and it is very hard to reconcile it with common-sense views about responsibility.

Specific mechanisms have been identified whereby a person may be caused to have a tendency towards crime and violence. For example, according to Peter Fenwick, a British forensic psychiatrist, there is evidence that

> abnormal brain development during the fifth month of pregnancy can result in the absence of conscience in the child and hence an inability to empathise with others. High testosterone and low five-hydroxytryptamine [seretonin] levels in the brain result in a tendency to violence. Hence violent behaviour is much more common in men than in women and especially amongst men with low 5-HT levels.[20]

One may think also of the damage caused to infants by abuse of alcohol or other drugs by the mother during pregnancy. And it appears that lack of 'good-enough' carer- infant interaction in the first two years of life causes deficiencies that can result in criminal conduct.[21] The evidence for links between the brain's physiology and criminal behaviour is explored at length in Moir and Jessell's book *A Mind to Crime*.[22]

One particular area of scientific research that is sometimes seen as refuting free will is that undertaken by Benjamin Libet and his colleagues.[23] These experiments indicate that consciousness comes too late to initiate certain actions that the agent considers to be voluntary. In particular, where subjects were asked to make a movement at any time they decided to, neurological preparation for the movement was detected about half a second *before* the time identified by the subject as the time of the decision to move, suggesting that the subject's feeling that the decision was freely chosen must be an illusion.

3.2 *The inadequacy of compatibilism*

Despite these scientific attacks, the compatibilist tradition has continued strongly up to the present. Hart's approach has been continued and developed by writers such as Braithwaite and Pettit, Posner and others.[24]

Michael Moore's *Law and Psychiatry*, for example, presents an extended defence of compatibilism. He argues[25] that persons need not stand outside the causal order of the world in order to perform human actions: the circumstance that the choice to act and the ensuing action are determined by factors such as the chemistry of the brain or early environment does not mean that the action is not a conscious willed action for which the person is responsible. To contend otherwise is to assimilate participation in the causal order with compulsion by external forces.

However, as pointed out by Alan Norrie,[26] this does not meet the deterministic challenge to responsibility: if determinism is true then, although human beings would still have a capacity for conscious

choice and action, this capacity and its exercise would themselves ultimately be determined by factors outside the person's control and thus would be not be adequate to ground notions of individual responsibility, so that retributivist conceptions of punishment could not be justified.

This general line of argument has been forcefully advanced in recent times by Peter Strawson's son, the philosopher Galen Strawson. In a *Times Literary Supplement* article, 'Luck swallows everything',[27] Strawson argues that when a person first starts making choices in life, the person must be as he or she then is as a result of heredity and early experience, which are things for which the person cannot possibly be responsible. And when the person makes later choices, the person must then be as he or she then is as a result of heredity, previous experience, and previous choices, with those earlier choices being the result of heredity and early experience; so that even the later choices are the result of heredity and experience, for which the person cannot be responsible. And if the person tries to change the way he or she is, this attempt, like any other choice, is itself the result of heredity and experience. Strawson accepts that this may not be the whole story, for some changes to the way one is could be the result of indeterministic or random factors; but he claims that indeterministic or random factors, for which one cannot be responsible, cannot contribute to one being truly responsible for the way one is, or for choices made as a result of the way one is.

So compatibilism cannot leave us responsible in any substantial way for our actions because, according to determinism, our actions are the product of our characters and dispositions, and we cannot be responsible in any substantial way for our own characters and dispositions: it must all be a matter of moral luck.

3.3 *Calls for a therapeutic approach*

Along with more forceful and strident attacks on the related notions of free will and responsibility, we are now seeing also stronger calls for abandonment of any idea of retribution in punishment. It is argued that we should adopt approaches which are purely consequentialist and/or therapeutic: that is, that any punishment must be justified solely by its consequences in terms of deterrence, prevention, and rehabilitation; and that crime should be treated as an illness to be treated rather than a wrongdoing to be punished.

One example of this is provided by philosopher Ted Honderich in his book *How Free Are You?* As Honderich points out, the law's current approach combines consequentialist ideas with retribution: punishment is seen as justified partly by its good consequences and partly by notions of desert and retribution. Honderich calls this a package-theory; and claims that, if determinism is true, the part of such a theory which embraces retribution should be abandoned.[28] He contends that if a person's actions are the inevitable result of causes occurring before a person's birth or otherwise outside the person's control, then retribution becomes indistinguishable from primitive vengeance.

As well as calls for the abandonment of retribution, there are also now increasing calls for a therapeutic approach to crime. In their book *A Mind to Crime*, Anne Moir and David Jessel write:[29]

> So if great swathes of crime are a biological disorder, how should it be treated if not by killing or incarceration? Perhaps we can make things easier for ourselves by not seeking to acquit or pardon such individuals, but by confining them for the protection of society. With the introduction of PET scans we should be able more easily to identify those with, for instance, dysfunctional control mechanisms. Identifying the cause, however, loads us with the responsibility of doing something about it - treating the offender. It is easier simply to condemn and lock away - although of course more expensive when the offender re-emerges from prison and inevitably offends again.
>
> ...

> So what's the alternative? We need to know what we do - as electors, lawmakers, jurors, doctors, citizens - with the growing knowledge that crime is as much a function of biology as anything else. Evil may be something no more sinister than a matter of loose connections. The devil may be the term for an accumulation of cerebral wounds. Perhaps the theologians, too, need to think again. Is it practically possible to discard the traditional concept of justice based on guilt and punishment and replace it with a 'medical model' based on prevention, diagnosis and treatment? Do we have the medical tools at hand to treat what has been diagnosed?
>
> ...
>
> What stands out from literally hundreds of papers and studies of the various types of criminal is widespread and cogent evidence of disordered minds resulting from dysfunctional brains. The real crime of many such people is the incapacity to comprehend the nature of guilt. If they had a more recognized learning disability, we would rush to help them with therapy and extra resources. But we do not recognize; we merely condemn. Incarceration is an expensive and wasteful reaction, which does very little good apart from providing a secure hotel facility for a limited amount of time.

So according to these authors, not only is retribution not justified, it doesn't work: what is required is appropriate quasi-medical treatment.

3.4 *Mismatch of concepts*

In discussing *mens rea* in Section 1 of this paper, I mentioned many categories which the law uses in defining crimes and defences, and about which decisions have to be made in criminal trials: categories of belief, intention, consent, unsoundness of mind, knowing what one is doing, knowing that what one is doing is wrong, willed or voluntary action, ability to control one's actions, loss of self-control, impairment of mental responsibility, and so on.

It is important to realise that these categories are all pre-scientific, folk-psychological concepts: the language which describes them is non-scientific, and, except in the clearest cases, the existence or otherwise of the psychological state which they describe is not susceptible to scientific proof.

There is in fact a widening gulf between the categories used by neuroscience and the non-scientific categories used by the law. In general terms, the latter presuppose an active conscious agent, with beliefs and intentions and the ability to make free choices; whereas the former presuppose the regular operation of laws of nature in physical cause and effect. Neuroscience focuses on questions of brain function, and seeks to identify particular causes or explanations of aspects of the functioning (or malfunctioning) of the brain. Apart from some areas of psychology and psychiatry which continue to use folk-psychological concepts, but which tend increasingly to be regarded as unscientific, neuroscience has no place for the concept of an holistic entity, the mind, which according to the law is supposed to be in control of a person's voluntary conduct. There are in effect two languages describing human conduct: the language of neuroscience, dealing in scientific cause and effect, and the language of folk psychology, dealing in the beliefs, intentions, choices, and actions of persons. There is not at present any accepted overall theoretical framework that can make sense together of these two languages and thereby facilitate translation from one to the other.

When scientific evidence is given in criminal cases, particularly those involving defences of insanity, automatism, and diminished responsibility, it is increasingly coming to be expressed in terms of scientific concepts remote from the folk-psychological concepts the law uses in its definitions. Because there is not at present any overall theoretical framework linking these concepts to the fundamental folk-psychological categories used by the criminal law, psychiatrists and other experts can provide little, if any, expert assistance in relating the evidence they give within their areas of expertise to these categories.

The meaning of these categories, their interpretation and explanation, are questions of law for the judge; and their application in particular cases is for the tribunal of fact, generally a jury. Psychiatrist experts do sometimes express an opinion on the meaning and application of the categories, but this opinion can carry relatively little weight: much greater weight is given to the more detailed matters in respect of which the expert witness has scientific or medical expertise, and it is left to the judge and jury to work out what these matters mean for the ultimate folk-psychological questions. The interplay of all these considerations is illustrated by Australian cases such as *Chayna* and *R v Trotter*.[30]

In the former case, for example, an appeal from a conviction for murder was allowed because the trial judge, in his direction to the jury on diminished responsibility, referred to an opinion of a psychiatric witness that the accused was not suffering from an 'abnormality of mind' (a folk-psychological category which was included in the requirements for the defence at that time), without also referring to more detailed evidence from the same witness which the jury could have taken as supporting the existence of such an abnormality. The appeal court remarked that the psychiatrist's opinion as to 'abnormality of mind' was 'the least important and least helpful aspect' of her evidence, and that some of the observations she made about the appellant and some of her opinions on psychiatric matters were more important.[31] Since these observations and opinions were not wholly unfavourable to the defence case, and were not referred to in the summing-up, the appellant was entitled to a new trial.

Along with this problem of translation, there is also the point that, in so far as science can make sense of responsibility, it suggests that responsibility is not all-or-nothing, as the law may seem to assume, but rather a matter of degree. Susan Greenfield, professor of pharmacology at Oxford University, argues that consciousness is not like an on-off light switch but more like a dimmer switch, with continuous gradations of intensity.[32] It seems reasonable to believe that much the same is true of responsibility - that there is a spectrum of responsibility, beginning with acts done without consciousness, through acts done with various kinds of impaired consciousness, acts done without deliberation, acts done under pressure, to acts done after full deliberation.

Problems like the above have led some commentators to suggest that the decisions reached by juries on the folk-psychological legal categories are really not decisions on questions of fact, but are no more than subjective value judgments concerning the degree of guilt of the accused; and that the law would do better to acknowledge this, to cease trying to fit cases into the elaborate but ultimately meaningless categories associated with the defences of insanity and automatism and diminished responsibility, and to focus directly on the real practical question of what should be done about the particular offender.

These points are elegantly made, for example, in an article by psychiatrist John Ellard,[33] prompted by the case of *Falconer* referred to earlier in this paper. In that case, a woman who had shot and killed her husband was found guilty of murder. In support of her defence of automatism, it had been claimed that at the time she was in a state of dissociation brought about by prolonged emotional stress caused by the deceased. Her conviction was set aside by the High Court of Australia on the ground that certain psychiatric evidence rejected by the trial judge should have been admitted, and a new trial was ordered. After considering the problems of translation between legal language and psychiatric language, Dr. Ellard concludes his article with the following 'plea for pragmatism':

> Is it not time for pragmatism? Once the facts are determined sentencing remains. When the death penalty was mandatory then there had to be ways in which convicted persons who were manifestly mad, or who had some other disability, could be extricated from the system if it seemed wise to do that. If judges have complete discretion then the need for all these elaborate fictions is removed and common sense and wisdom can enter the field. In the case in point there was no doubt about who killed whom and why. The only real issue was what to do about it. When the High Court sent the case back down again the accused pleaded guilty to manslaughter, the prosecutor accepted the

plea, the sentencing judge said that the crime was worth five years and then did some elaborate calculations which allowed the convicted person to walk out of the court free. Wisdom got there in the end.

This amounts to a claim that the very concepts used by the law in relation to *mens rea*, particularly concerning questions of insanity, automatism, and diminished responsibility, should be abandoned and replaced by a pragmatic therapeutic approach.

3.5 *Problems of balance*

There appears to be a strong public perception at present that the law is soft on crime, and also much sympathy for the view that perpetrators of crime should be made to suffer no less than their victims and their victims' families suffered from the crime. One often hears it said by victims or relatives of victims that they have been given a 'life sentence' by the crime, so it is unfair that the offender be released after 2 or 5 or 10 or 20 years.

There seems to be no common ground between people holding these views, and those who advocate recognition of the biological and social roots of crime, and application of scientific knowledge to its prevention and to the rehabilitation of offenders: such opposing views seem to be becoming increasingly polarised.

In those circumstances, there is a real need for a theory of retributive justice, which gives a rationally defensible account of responsibility and desert, and which also can take into account and make use of scientific advances in order to refine the theory, without jeopardising the very idea of responsibility.

4 How Should the Law Respond?

In the light of all these factors, there is a need to rethink questions of responsibility. The compatibilism of persons like Peter Strawson, Herbert Hart and Michael Moore may no longer be adequate as a justification of the law's concern with *mens rea*, and plainly it does not provide an adequate theoretical framework to enable the law to take advantage of advances in neuroscience, while appropriately maintaining ideas of responsibility.

What then should be the law's response to these developments?

I suggest that we as lawyers should be open to whatever insights can be provided by neuroscience; and those with lawmaking roles should be prepared to use these insights both in moulding the law's general approach to crime, and in refining the particular categories that the law uses. However, I also suggest that we should be sceptical when neuroscience appears to conflict with common-sense ideas of free will and responsibility, and should strongly oppose any abandonment or significant watering-down of these ideas. We should remain firmly aware of the merits of principles of responsibility and retribution, and of the deficiencies in purported philosophical and scientific refutations of those notions. And we should seek to uphold the moral underpinning of the law.

4.1 *The merits of retribution*

One very important reason for maintaining notions of responsibility and retribution is that such notions underpin deeply-held principles of justice and human rights, which are regarded as essential pre-requisites for civilised societies, and indeed are being increasingly recognised, and where possible promoted, by international law.

These principles give great weight to the autonomy of people and require respect for that autonomy. According to them, a citizen is generally entitled to freedom from interference from the coercive processes of the State *unless he or she voluntarily breaches a fair rule of law*, publicly promulgated by the State. That

is, a citizen should have a choice as to whether to be liable to coercion or not; and in this regard, folk-psychological categories such as belief and intention and voluntariness of action are of central importance.

This is not to say that criminals should be punished purely as retribution, because that is what they deserve: on the contrary, what I am saying is entirely consistent with the view that the coercive *system* which is the criminal law must be justified at least in part by its *utility*, because of the need to protect the majority of citizens from dangerous and anti-social activities of others. But, given that society needs such a system, the question is how should this system identify those persons to whom the coercion is to be applied, and how should it determine what coercion is to be applied to these people.

The solution which has been adopted by civilised societies with respect for justice and human rights is generally along the following lines, which I will call *the human rights qualification* to the power of the State to coerce its citizens: as a general principle, the system should allow coercion to be applied only to people who have been proved, by due process of law, to have voluntarily acted in breach of a public law; and then, no more coercion is to be applied than is 'proportional' to the gravity of the offence. There are qualifications to this principle, the most important for present purposes being one associated with the defence of insanity: where a person involuntarily does harm because of a mental abnormality which may cause that person to do serious harm again, he or she is regarded as less than fully responsible and is liable to coercion by way of restraint and/or treatment to the extent required for appropriate protection of the public.

It is of course argued that there are good utilitarian or consequentialist reasons for this human rights qualification, which do not require acknowledgement of the validity of any idea of criminal responsibility which legitimises punishment, or any basic distinction between acts which involve a guilty mind and those which do not. For example, there is the consequentialist argument mentioned earlier, advanced by Hart and others, that there should be parsimony in threatening and applying coercion; and since it is generally only voluntary actions which are susceptible to deterrence by threat of punishment, it makes sense to threaten and to apply coercion only in respect of voluntary actions. But even this argument presupposes a fundamental distinction between actions which are voluntary and actions which are not,[34] which is difficult to sustain if one rejects free will and responsibility.

And in any event, the argument does not do away with the need to be able to rely on appeals to justice and human rights, which attribute intrinsic significance to the distinction between acts which are voluntary and those which are not - particularly because consequentialist arguments are irredeemably indecisive. Those governments which do not recognise the human rights qualification to the application of the coercive processes of the State generally assert that the good of society overrides other considerations; so that, for example, the detention without trial of political opponents is justified. The strengthening consensus of reasonable opinion throughout the world is that that kind of approach can be justified only in cases of real emergency, and then only as a temporary measure pending resolution of the emergency and as a part of the process of establishing or returning to a normality in which the human rights qualification does apply. However, if one is limited to arguments about consequences, it is impossible to make out a case which would convince anyone who wished not to be convinced: it is impossible to prove what all the consequences of the alternatives would be, let alone to prove which of the totalities of consequences would be 'better'.

Imagine trying to prove to a previous Soviet government (or to those Russians who now hanker for the good old days) that it would have had better consequences not to confine dissidents in mental asylums, or to the present Chinese government that the recognition of human rights qualification to the use of the coercive powers of the State would have better consequences than not recognising it.

But acceptance of the independent force of the considerations of fairness which justify the human rights qualification would mean that there would need to be positive justification for overriding it - in terms of

real emergency, clear and present danger, etc. If justice and human rights are given independent weight, then a heavy onus can be placed on governments seeking to deny human rights to their citizens.

These views about human rights are being increasingly recognised and promoted in international law, and in the domestic law of many countries. As just one example, the Canadian Charter of Rights and Freedom requires that the criminal law respect 'the principles of fundamental justice'. The Canadian Supreme Court, in 1990, decided that those principles required that a person should not be convicted of murder unless he or she actually *foresaw* that death would result from his or her conduct; and the Court held to be invalid a law to the contrary.[35] I am not saying the Court was necessarily correct in its views of fundamental justice - rather that this is an illustration of how the importance of human rights and justice is being recognised.

So I contend that ideas of human rights, which are widely and increasingly accepted, depend crucially upon discriminating between persons whose conduct makes it *fair* and *permissible* that their freedom be curtailed, and persons whose conduct does not make it fair and permissible that their freedom be curtailed. And in order that conduct make curtailment of freedom fair and permissible, it must involve *voluntary action* in breach of a public law, for which the person is responsible.

If the notions of free will and responsibility are discredited, human rights are prejudiced: there will appear to be no rational basis for saying that it is fair, and thus permissible, to curtail the freedom of a person (who has had the bad luck to be caused by genes and environment to become a person) who acts in breach of the law, yet unfair and thus impermissible to curtail the freedom of a person (who, without breaching any law, has had the bad luck to be) regarded by the government as a danger to society.

In relation to principles of punishment, this general approach was given a persuasive exposition by author C. S. Lewis, in an article published in the journal of the Law Students' Society of Victoria *Res Judicatae*,[36] which was referred to and adopted by the leading judgment of the High Court of Australia in *Veen* v *the Queen [No. 2]*.

4.2 *Free will not refuted*

One possible response to such considerations is to argue that, even though neuroscience refutes free will, we should act as if this had not happened: this approach is taken by Steven Pinker.[37] That is to say, we should hold that, although free will is a dead duck, we must pretend it is alive.

However, I contend that a balanced consideration of the evidence and arguments suggests that free will may in fact be alive and well; and that, because belief that it is alive is valuable, there would need to be a powerful case for its demise before it could be rational either to accept that it is in fact dead or to live our lives as if it were. In fact, despite the confident assertions of scientists such as Blakemore, and philosophers such as Galen Strawson, the existence of free will and responsibility is far from being disproved: good sense can be made of these ideas consistently with everything that science reliably tells us about the world.

The basic point about most purported refutations of free will from Hume onwards is that they have proceeded on an unstated assumption which is very close to what is claimed to be proved - the assumption that whatever occurs in the world must either be the *only* occurrence consistent with laws of nature, or else be random. This mechanistic assumption immediately excludes the possibility that consciously-made choices are on the one hand not pre-determined by laws of nature, nor on the other hand merely random occurrences. In circumstances where our choices *seem* to us to be neither pre-determined nor random, one would expect to find some justification of the assumption - and it is not provided.

In common-sense thinking, our capacity to choose depends upon our ability to consciously appraise alternatives and the reasons or motives for and against each of them. I do not suggest that we are conscious of *all* factors that motivate us - but we certainly are conscious of some of them, and it seems

reasonable to believe that our consciousness is closely linked with, indeed necessary for, our ability to choose. I do not think science can properly claim to explain our capacity to choose unless and until it can explain our consciousness; and it is nowhere near doing that. Thus, for example, scientists have not the faintest idea what would be required to construct a machine that could feel pain, much less specify precisely what objective features would distinguish such a machine from one that could not feel pain. Unless and until they can do this, it goes far beyond the competence of science to assert that consciously-made choices must be pre-determined by laws of nature or else be random.

Looking at Strawson's argument outlined towards the end of Section 3.2, it will be seen that there is the unstated assumption that (subject to possible randomness) a conscious choice is pre-determined by the way a person is immediately prior to the choice. Strawson ignores the possibility that the way a person is immediately prior to the choice may only pre-determine the alternatives available and the (inconclusive) reasons or motives the person has for and against each of them. The person, using his or her capacity to make a conscious choice (a capacity which all normal adults have), may then resolve the competing inconclusive reasons - such resolution being neither random nor pre-determined by the way the person is immediately prior to the decision.

In terms of responsibility, one may accept that nothing a person does or can do *at the time of any choice or action* can make the person responsible for what alternatives are then available, for the way those alternatives then appeal, or indeed for having the capacity to choose between them - but one may nevertheless claim that, leaving aside any question of responsibility for *these* matters, the person can still be responsible for *the way the person exercises the capacity to choose*. On this view, the way the person is, in respect of character and motivation, does not predetermine what the person does: it only predetermines what the alternatives are and how they appeal. The person is not responsible for having the capacity to choose between these alternatives, but is not predetermined in how that capacity is exercised, *even by the way the person then is in respect of character and motivation*; and so the person (and no-one and nothing else) is responsible for the way that capacity is exercised. That responsibility may be greater or less, by reason of the nature of the choice posed by the way person now is - the harder it is, by reason of the person's inclinations, for the person to make the 'right' choice, the less blameworthy will be the 'wrong' choice - but on this view, normal adult human beings always have some responsibility for their choices and voluntary actions.

And this means in turn that the person may have some responsibility *through prior choices* for the way the person now is in respect of character and motivation, and thus for presently-operating reasons and the way they appeal.

I would also note that the Libet experiments are far from refuting free will. Even in relation to the simple actions they dealt with, they do not, as Libet himself notes, exclude a conscious *veto* exercisable right up to the time the action is performed. They do not exclude the conscious *shaping* of actions, the details of which are habitual (as in musical performances). And they say nothing about the fully deliberated actions that are the paradigm case of free will.[38]

4.3 *Refining categories and principles*

It follows from what I have said that I believe that the general approach of our system of law to questions of criminal responsibility is along the right lines. In general terms, the human rights qualification to the power of the State to coerce its citizens should be maintained: that is, as mentioned earlier, as a general principle the legal system should allow coercion to be applied only to people who have been proved, by due process of law, to have voluntarily acted in breach of a public law; and then, no more coercion should be applied than is 'proportional' to the gravity of the offence. This is the general principle, and exceptions may be recognised, such as the following:

(1) where any lack of voluntariness or intention is due to self-inflicted intoxication, it may be reasonable to waive the requirement of proof of voluntariness or intention;
(2) in some cases, where carelessness can cause great harm, punishment may be allowed for carelessness (for example, negligent driving);
(3) in some cases, where great harm can be caused in circumstances where intention or carelessness is difficult to prove and where harm can and should be prevented by particular people taking appropriate precautions, strict liability may be imposed (for example, in matters relating to public health);
(4) where a person involuntarily does harm because of some mental abnormality which may cause that person to do harm again, coercion may be applied to the extent necessary to protect the public by restraining the person and/or treating the mental abnormality (this is the rationale of the defence of insanity); and
(5) where prosecution authorities believe on reasonable grounds that a person has acted in breach of a public law and that it is necessary and appropriate to arrest and detain the person so that a trial can take place, then coercion can be applied to that extent, so long as an early opportunity is afforded to the person to apply to a court for bail.

If one is to follow principles of this kind, it is necessary for the law to draw hard and fast lines, despite the fact that there are continuous gradations of responsibility. The lines can take some account of different levels of responsibility, but finer gradations can be given effect to only in sentencing and the determination of treatment after sentencing.

In applying the principles to cases of homicide, I believe it is appropriate to reserve one category for the most serious cases, in our law the category of murder, and to have a category of less serious homicide where responsibility is less, whether because of provocation or because of some kind of diminished mental capacity to make reasonable choices. So despite the persuasiveness of Dr. Ellard's pragmatic approach in the article referred to earlier, I think it should be rejected: the law needs to have categories of the type he criticises, despite the problems of stating them clearly and of applying them.

However, this is not to say that the precise definition of the categories should be immutable. On the contrary, advances in neuroscience may enable better definitions to be drawn up; better in being either fairer, or more readily translated into scientific language, or both. Perhaps also some adjustments of principle could be justified.

For example, it could be found appropriate to have more than just the one category of irresponsibility that we have at present, namely insanity - which is such that on the one hand the stigma and consequences of a successful defence mean that it is relied on only in cases of homicide, while on the other hand a verdict of *not guilty* in some cases of homicide, albeit on the ground of insanity, can seem an affront to the families of the victims and to the public. There could be advantages in having two insanity-like verdicts available, one called (say) *non-responsibility* and the other called (say) *criminal insanity*. In both cases, the verdict would exclude the existence, at the time of the alleged offence, of a genuine capacity to make a choice, and thus would require a verdict of not guilty of the particular offence charged. However, the requirements and incidents of the two verdicts could otherwise be quite different.

The former verdict could be available in those cases where the lack of real ability to make a choice was due to some temporary and/or treatable condition such as epilepsy or hormonal imbalance; and the result of such a verdict would be the minimum intervention required to protect the public, and having a fixed upper limit to ensure that this intervention is less onerous than the consequences of a guilty verdict. The onus of proof for this result would lie on the accused.

The latter verdict would be available in those cases where the lack of real ability to make a choice was due to some more serious and difficult-to-treat mental state, which would cause the accused to be a

serious danger to the community for the indefinite future; and the result of such a verdict would include the possibility of lifetime confinement. The onus of proof for this result would lie on the prosecution.

In applying this approach, it would not be necessary to ask whether the offender is 'mad or bad'. The question would be simply whether or not the accused had a real ability to make a choice, and if not, whether this was due to some serious and difficult-to-treat condition which would make the accused a danger for an indefinite period. It would not matter whether this lack of ability was due to inability to empathise with others, or inability to know what one was doing, or some other deficiency.

Advances in neuroscience could also enable the law to deal more effectively with other cases where responsibility is questionable, such as drug-related offences and offences by children. In the former case, where (say) non-violent theft is committed to satisfy a craving, the law might treat it as a case of non-responsibility *if* the drug-dependence is made out to be temporary and treatable by an appropriate commitment by the accused to undergo rehabilitation. In the latter, it may be possible to formulate a better definition of what amounts to responsibility in children.

In general terms, then, I believe that we should be looking to articulate an overall framework within which good sense can be made both of neuroscience and of common-sense ideas of free will and responsibility; and seeking to improve the categories used by the criminal law, so that they fit as well as possible with neuroscience as understood within this framework.

4.4 *Punishment, treatment, and education*

As I noted at the beginning of this paper, I accept that there is a role for a therapeutic approach to crime, for example in the early identification and minimisation of risk factors and in the treatment and rehabilitation of offenders; and it is in these areas that neuroscience has the greatest contribution to make.

As scientific research shows what are risk factors for criminal behaviour, how to identify such factors in particular individuals, and how to minimise such factors, we should be prepared to use and to apply this knowledge. And similarly, as scientific research produces insights into how best to go about rehabilitating offenders, we should be prepared to accept and apply these insights. However, all this must be done within a framework that recognises the principles of human rights and responsibility which I have discussed, and also recognises the need for the moral underpinning of the law. A number of points arise from this.

When one is considering early intervention in the lives of children, it must be recognised that this intervention should be consensual and not coercive, except as against persons (parents and/or children) who have already committed an offence, or where the child in question is shown to be in need of care and the proposed measures are the least intrusive that can satisfy this need. A fortiori, coercive intervention in the case of adults should be limited to the cases I have discussed.

In relation to persons who have committed an offence, it may from time to time appear that rehabilitation of the particular offender would best be achieved by treatment which is far less onerous than the punishment suggested by considerations of retribution or general deterrence. However, the person by his or her conduct has made it fair and just that general deterrence be pursued through the imposition of proportionate punishment; and accordingly the objective of rehabilitation may sometimes properly be subordinated to considerations of retribution and deterrence.

Even in pursuing treatment and rehabilitation, notions of free will and responsibility may have an important role to play. According to psychologists involved in an innovative program with long-term prisoners at Long Larton Prison in the United Kingdom, one essential step in the rehabilitation of a violent offender is that he accept that he - not anyone or anything else, not society, not his parents, not his genetic inheritance, but *he himself* - was responsible for the events which put him in prison. Without that, they say, there is no hope of rehabilitation.[39] On a more literary and intuitive level, a persuasive case is made out by

Dostoyevsky in *Crime and Punishment* that recognition and acceptance of guilt, and indeed of deserved punishment, can be crucial in reform and rehabilitation.

This leads on to the more general point that, to be truly effective, the criminal law needs to be underpinned by a generally-accepted morality and a generally- accepted view that the law is in accord with that morality and is *fair*; so that compliance is secured to a large degree by wide recognition of moral obligations to do what the law requires and to co-operate with those who uphold and enforce the law. Persons who fail to recognise these obligations, whether by reason of disregard for moral considerations generally or disrespect for the law in particular, may do so partly because of brain malfunction; but recognition of these obligations needs to be promoted by means other than just treatment of disordered brains.

One significant reason for disrespect for the law is that the law can be seen as protecting the rich and oppressing the poor, even though it may apply the same rules to everyone. As Anatole France put it 100 years ago, 'the majestic equality of the law ... forbids the rich as well as the poor to sleep under bridges, to beg in the streets and to steal bread.'[40] I think it is understandable that many of the disadvantaged in society should regard appeals to law and to morality as attempts to impose on them standards which they justifiably do not accept, and which in any event are given only lip-service by others.

No doubt many factors contribute to such attitudes. One may be a perception that many of the most wealthy in society accept all the benefits that are provided by society and law and government, and at the same time fail to recognise their own moral obligation to make a contribution to the cost of those benefits which is fair having regard to their resources; but on the contrary take extreme measures in order to contribute as little as possible, by means such as setting up complicated systems of trusts and companies in their own countries and abroad with just this purpose. Denunciation of such practices as a significant immorality, coupled with serious and effective measures to defeat them, could make a substantial contribution towards increasing respect for the law. In general, I believe it is heavily incumbent on the better-off in society, if they wish to rely on the law and moral standards to protect their position, to make sure that they themselves act legally and morally in their dealings in society, so that demands made of the disadvantaged are not seen as unjustified and hypocritical.

Certainly we should try to ensure that the law is not such as to oppress the disadvantaged in society, particularly identifiable disadvantaged minorities such as indigenous Australians; inter alia by seeking (1) to attack inequalities in society by social programs, (2) so far as possible to make the content of the law such as not to weigh more heavily on disadvantaged minorities, (3) to ensure that law enforcement agencies do not target such minorities, (4) to use powers to refrain from recording convictions in trivial cases, and (5) upon conviction to find alternatives to gaol for such disadvantaged minorities wherever possible.

If principles such as these are followed, it could be possible without hypocrisy or embarrassment to combat disregard for moral considerations generally, and to support the requirements of the criminal law, by reference to fundamental moral principles with which there could be general agreement: for example, the Golden Rule (GR), both in its negative and (more circumspectly) its positive form.

GR negative, 'do not do to others anything that you would wish not to be done to you', could I think command wide acceptance. It has intuitive appeal, and it directly supports many of the prohibitions of the criminal law, and also the moral imperatives of honesty and fulfilling one's commitments. More indirectly, it supports moral principles of fairness such as one referred to earlier: if you would wish not to have other people taking advantage of you by freeloading on a system to which you contribute, then you should not freeload on a system to which others contribute - even if you believe there are other people again who are also freeloading.

GR positive, 'do to others what you would wish to be done to you', perhaps needs more circumspection. It could be understood as justifying, even requiring, that one impose one's likes and

dislikes on other people. It could even be understood as in some circumstances requiring sexual harassment or sexual assault! So before you do something to another for the reason that you would wish it to be done to you, it is necessary to be satisfied that it is wished by that other person as it would be by you.

The other area of circumspection in relation to GR positive concerns the tension between self-interest and altruism. Presumably GR positive means that you should do to others what you would wish to have done to you *if* you were in their position; and so it could require one to give to the poor until substantial equality is achieved;[41] and that probably would not have universal acceptance. A more moderate version of GR positive, which could perhaps be widely accepted, would require one to keep no more for oneself and one's family than is reasonable in all the circumstances, and to give anything left over to those who need it: there could be significant disagreement about how much is reasonable, but at least this version would tend in favour of charity and against extremes of self-indulgent extravagance.

Along with principles like GR, one could look for agreement about what are *virtues* which should be encouraged in oneself and others; and in a moral climate where GR is accepted and such virtues are recognised and promoted, one could also hope that people could be encouraged to give as much attention to their responsibilities as they do to their rights, and to identify with their community and with humanity in general. And moral education of children could be more effectively undertaken. In this regard, some thoughts sent by Mary Warnock to the seminar referred to in endnote 1 are most pertinent:

> The importance of moral education cannot be exaggerated. Children, as Mill understood, need to be taught to be moral. What they need to be taught is, specifically, to *want to be good*. Moral pluralism has been much exaggerated. Most religions, for instance, and most differing communities have more in common, in moral demands, than is often allowed. Human beings tend to dislike and fear the same things; there are numerous shared values between humans and other animals, but especially among humans. The real difference is between people who are interested in morality (who want to be good) and those who are not, that is who have no concept of morality at all. That it is difficult to produce more than a 'basic moral consensus' may be true. But if such a basic consensus that was truly moral could be introduced to children at school, and manifestly upheld by their teachers, using properly moral language, then this would be a huge step forward.

Thus, at the same time as we welcome neuroscience and apply it in our efforts to deal with crime in our society, we should I believe also be promoting ideas such as these, as well as seeking the retention and refining of reasonable notions of free will and responsibility.

* Chief Judge in Equity, Supreme Court of New South Wales.

ENDNOTES

1 For example, by giving a talk on free will at a session of the 1994 conference of the Australian Neuroscience Society; giving a talk on science and criminal responsibility at the 1995 conference of the Institute of Australasian Psychiatrists; giving a similar talk to a plenary session of the second Towards a Science of Consciousness conference in Tucson, Arizona in 1996, published in the book of the conference, Hameroff, Scott and Kaszniak (eds), *Towards a Science of Consciousness II* (MIT, Cambridge MA, 1998); and participating in a seminar on neuroscience, responsibility and the law organised by the Scientific and Medical Network at Drynachan Lodge near Inverness in 1997, along with neuroscientists Susan Greenfield and Steven Rose, forensic psychiatrist Peter Fenwick, law professor Alexander McCall Smith, and philosophers Mary Midgley and Bill Fulford, among others. (It was this seminar that prompted me to write this paper.)

2 (1990) 171 CLR 30, at 40-3.

3 [1935] AC 462.

4 (1843) 8 ER 718 at 722.
5 *R* v *Radford* (1985) 43 SASR 266 at 275.
6 *Bratty v Attorney-General for Northern Ireland* [1963] AC 386 at 413-4.
7 [1984] 1 AC 156.
8 *Russell* (1993) 70 A Crim R 17.
9 It has been said that a person will not involuntarily commit a crime under hypnosis; but perhaps this could happen if the person were given under hypnosis a *belief* that the facts were such that the act in question would not be a crime - for example, a belief that certain property was one's own and had been stolen by another (when that other was in fact the true owner).
10 [1991] 2 QB 92.
11 *Veen v the Queen [No 1]* (1979) 143 CLR 458.
12 (1988) 164 CLR 465.
13 *Veen [No 2]* at 476-7.
14 Berlin, *Four Concepts of Liberty* (Oxford University Press, Oxford, 1969) at 113.
15 (1962) *48 Proceedings of the British Academy* 1-25.
16 Hart, *Punishment and Responsibility* (Oxford University Press, Oxford, 1968).
17 Blakemore, *The Mind Machine* (BBC, London, 1988) at 257, 269-271.
18 Dennett, *Elbow Room* (Oxford University Press, Oxford, 1984) and *Consciousness Explained* (Allen Lane, New York, 1991); Patricia Churchland, *Neurophilosophy* (MIT, Cambridge MA, 1986); Paul Churchland, 'Eliminative materialism and propositional attitudes', in Lycan (ed), *Mind and Cognition* (Blackwell, Oxford, 1990) at 206-23; and Crick, *The Astonishing Hypothesis* (Simon & Schuster, London, 1994).
19 (Little Brown, London, 1996).
20 Clift, 'Mind, behaviour and the law', (1993) *53 Network* 14-5.
21 Gordon, 'Developmental infant research, organisation of personality and "criminal" behaviour', paper presented to the 1999 Australian Supreme and Federal Courts Judges Conference; and Taylor, Bagby and Parker, *Disorders of Affect Regulation: Alexithymia in Medical and Psychiatric Illness* (Cambridge University Press, Cambridge, 1997).
22 (Michael Joseph, London, 1995).
23 Libet, 'Neuronal vs subjective timing for a conscious sensory experience', in Buser Rougel-Buser (eds), *Cerebral Correlates of Conscious Experience* (North-Holland, Amsterdam, 1978); and Libet, Gleason, Wright and Pearl, 'Time of conscious intention to act in relation to onset of cerebral activities (readiness potential): the unconscious initiation of a freely voluntary act' (1983) *106 Brain* 623-42.
24 Braithwaite and Pettit, *Not Just Deserts* (Oxford University Press, Oxford, 1990); Posner, *The Problems of Jurisprudence* (Harvard University Press, Cambridge MA, 1990).
25 (Cambridge: Cambridge University Press, 1984) 361-3.
26 *Law, Ideology and Punishment* (Kluwer, Dordrecht, 1991) 148-9.
27 *Times Literary Supplement,* 26 June 1998, 8-10.
28 (Oxford University Press, Oxford, 1993) at 127-8.
29 At 296-8, 303.
30 (1993) 66 A Crim R 178 and (1993) 35 NSWLR 428.
31 At 188.
32 Greenfield, 'How might the brain generate consciousness?', in Rose (ed), *From Brains to Consciousness* (Penguin, London, 1999).
33 Ellard, 'Some notes on non-insane automatism and the will', (1995) *69 Australian Law Journal* 833-840.

34 Note that Braithwaite and Pettit, in their powerful (and basically consequentialist) attack on retributive theories of justice, *Not Just Deserts* (Oxford University Press, Oxford, 1990), assume that sense can and should be given to concepts such as 'morally culpable', 'blameworthy', 'fault', 'blameless', 'intend', and 'culpability' (all at 99).
35 *R v Martineau* [1990] 2SCR 633.
36 Lewis, 'The humanitarian theory of punishment', (1953) *6 Res Judicatae* 224-230. Lewis wrote (at 230) that he sent this article to an Australian periodical because he could 'get no hearing for it in England.'
37 Pinker, *How the Mind Works* (Allen Lane, London, 1997) at 55-6.
38 I have given here the merest outline of a defence of free will. I explore these and related ideas further in *The Mind Matters* (Oxford University Press, Oxford, 1991), 'Nonlocality, local indeterminism, and consciousness', (1996) *9 Ratio* 1-22, and 'Hume's mistake', in Libet, Freeman and Sutherland (eds) *The Volitional Brain* (Imprint Academic, Thorverton, 1999).
39 Beriff, 'Confrontation on E Wing', (1996) BBC TV Film.
40 *Le Lys rouge* (1894) ch.7.
41 See for example Kagan, *The Limits of Morality* (Oxford University Press, Oxford, 1989).

……