[Using the 19th century tale of Dr Jekyll and Mr Hyde as a leitmotif, the author gives an extended analysis of the philosophical bases of the two elements in the act of the accused which juries in Victoria are generally instructed are essential to criminal responsibility: consciousness and voluntariness. Drawing on a suggestion by the late Professor Brett he shows how the concept of personal identity—present consciousness of the self and consciousness or memory of the past life of that self—as defined by John Locke in his Essay Concerning Human Understanding (1694) underlies much of the modern legal thinking. Locke's theory that memory of past actions is the basis for present responsibility is suggested as the intellectual force lying behind the decisions in recent cases where convictions were quashed because the accused was so intoxicated he was held incapable of having formed any intention at all—an analysis that the author finds unintelligible. In bringing the underlying philosophical concepts of identity, consciousness and volition to bear upon the current search for criteria for criminal responsibility, the author discusses possible approaches the courts might adopt. He praises the clarity and good sense of the position taken in Joyce's case by the former Chief Justice of the South Australian Supreme Court, the Honourable John Bray.]

INTRODUCTION

This paper is an exploration of the diverse paths between two propositions in the analysis of criminal responsibility. The first is derived from an unusual restatement of a familiar problem by the late Professor Brett. An individual accused of a criminal offence, typically though not always an offence of violence, may say, 'I did this, but only because I was drunk'. Professor Brett went on to suggest:

What he is saying, in effect, is that this crime was not committed by him, but by a 'different person temporarily occupying his body'. His plea is of the same order as that made by a person who commits a crime while in, say, a state of hysterical fugue, or of amnesia; or by a person in his sleep; or by a person under hypnotic influence. In all these cases the defendant is saying in effect the same thing: 'The person who committed this crime was not really me, the person who is now before the court'.

Traditionally the issue of criminal responsibility in these cases has been made to depend on the questions: 'Was his act voluntary?'; 'Was he conscious?'; 'Did he intend the consequences?'; 'Did he realize, or foresee, what would happen?'; and similar formulations. It is rare for the courts to make explicit reference to the problem of personal identity which lies behind the claim to exculpation. The traditional questions avoid the obvious issue suggested by Professor Brett's examples. The cases, with very few exceptions, involve actions by the accused which are manifestly purposive and bear all the marks of intentionality. The value of this section of the Inquiry lies in the clarity with which the underlying issue is formulated: can these purposes, or these intentions, be attributed fairly to the individual before the court?

* J.D. (Chic.), LL.B.: Senior Lecturer in Law, University of Melbourne.
2 Ibid. 196.
That was one point of departure. The other was suggested by the judgment of the Full Court in the South Australian case of *R. v. Joyce* in 1970:

In our view, if the personality is divided and one of the divided parts was conscious of the act and wills it, the actor is responsible for it and the defence of automatism is not open.

It is an uncompromising statement, consistent with the approach taken by the South Australian Supreme Court in other cases of the time. The spirit of intellectual rigour in which the Court gave its answer owes much to the influence of the former Chief Justice, the Honourable John Bray, who presided over the Court in those years. The judgment of the Court in *Joyce* is one of enduring importance in the Australian canon of cases on criminal responsibility.

The area which lies between Professor Brett’s statement of what is essentially a moral issue, and the answer which *Joyce* provided in terms of legal responsibility, is uncharted. I have drawn on three texts in this attempt to mark some paths. The first, continued as a leitmotif throughout most of the paper, is ‘The Strange Case of Dr. Jekyll and Edward Hyde’ taken from the novel by Robert Louis Stevenson. The second is the discussion of personal identity in the *Essay Concerning Human Understanding*, by John Locke. The third is an account of criminal responsibility by an eminent American psychiatrist, Phillip Roche. The choice of texts may appear idiosyncratic. My only justification is that elements of error, exaggeration or caricature derived from these sources may serve to mark some of the possible paths between Professor Brett’s statement of the problem and the answer suggested by the Court in *Joyce*.

This paper is offered as a tribute to the eminent scholar who proposed the issue and to the Court which gave so articulate an answer.

**THE CAREW MURDER CASE**

It was a murder of great brutality. A maidservant at an upper window saw the victim, Sir Danvers Carew, accost Edward Hyde in the street, perhaps to ask for directions. There was no provocation at all. Hyde attacked Carew with a heavy walking stick and beat him to the ground. He continued the attack until the stick was broken and then trampled on the body. Hyde was never brought before the courts. He took his own life some months later. Dr Henry Jekyll, whose relationship with Hyde provides the problematic basis for this paper, had ceased to exist some time before Hyde’s dissolution.

Had it been possible to put Hyde on trial he would have been convicted of murder. One might doubt that conclusion only if it had been possible for Jekyll to return. But suppose it had turned out differently. What if Jekyll had survived Hyde? This question provides the basis for proposing a far more difficult issue. Could Dr Jekyll have been convicted of the murder of Sir Danvers Carew?

4 Ibid. 192-3.
6 The murder of Sir Danvers Carew is the central episode of evil in *The Strange Case of Dr. Jekyll and Mr. Hyde*, by Robert Louis Stevenson. The novel was first published in 1886. References in this paper are to *Dr. Jekyll and Mr. Hyde and Other Stories*, edited by Jenni Caldwell (Penguin English Library, 1979).
Stripped to its bones and arranged in chronological order, the relevant events of the case are as follows:

— Dr Henry Jekyll, who was a wealthy, respected and on the whole virtuous man, discovered an elixir which had the effect of releasing his *alter ego*, Edward Hyde. The term ‘*alter ego*’ is used for want of a better and is intended to bear no very particular meaning.

— After this discovery Jekyll and Hyde shared a life of alternating existences over many years. Except for the very last months of their shared life it was always necessary to take the elixir in order to effect a transformation.

— Hyde, when he was in the ascendant, enjoyed a life of unrelieved depravity. The author, for the most part, provides no information as to the nature of Hyde’s diversions. We may assume that he was at least sexually licentious. There is one incident, near the opening of the story, however, which indicates his indifference to human suffering. Hyde was walking at a brisk pace in the street when he collided with a little girl who was rounding a corner. She fell and he continued at the same pace and walked over her body. Witnesses were appalled at the inhuman callousness of his actions. One said he was like a juggernaut. The girl was not seriously hurt.

— This incident did nothing to deter Jekyll from continuing his double life. Some ten months later, however, Jekyll suffered the first involuntary transformation. He went to bed as Jekyll and woke as Hyde.

— Jekyll resolved never again to take the elixir. For a time he led a life of exemplary severity. Then his resolve weakened. On October 18th, the day of the murder, Jekyll succumbed to temptation and took the elixir again. Having been caged so long, Hyde was in a mood of unprecedented savagery and killed Carew. There was no motive for the murder. Afterwards Hyde sought refuge by resuming the identity of Jekyll.

— For two months more Jekyll led a blameless life. He entertained his friends. He was considerate, cultivated and civilized. But, on January 9th, there was another involuntary transformation. Jekyll then retreated into seclusion.7

— Thereafter the involuntary changes grew more frequent. Jekyll’s existence could only be maintained by increasingly frequent use of the elixir. The supply was soon exhausted. All attempts to compound more failed. One vital ingredient was unobtainable. Jekyll ceased to exist.

— Under the threat of imminent detection Hyde took cyanide.

The nature of the differences between Dr Jekyll and Mr Hyde are of importance. They were quite distinguishable in appearance. Jekyll was a large well fed and handsome man with ‘every mark of capacity and kindness’. Hyde was small, swarthy, agile, hirsute and considerably younger than Jekyll. Those who saw him were at a loss to describe him with any particularity. His appearance hinted at some indefinable deformity. One observer described him as ‘troglodytic’. There is a

— The date of this incident is given in Dr. Lanyon’s narrative, *ibid.* 74, and confirmed by Dr. Henry Jekyll’s posthumous account, *ibid.* 92. The letter received by Dr. Lanyon on that day was dated considerably earlier, however, on 10th December. This date appears to have been an error on the author’s part, explicable perhaps as a consequence of the speed with which the story was written and prepared for the printer.
reference to ‘ape like’ behaviour. Jekyll and Hyde did, however, have identical handwriting. And in his written communications Hyde adopted, when it suited him to do so, the prose style one would expect of Jekyll. There is a more fundamental point of relationship however. Hyde and Jekyll shared a common memory. Each, it appears, was conscious of the others’s actions. Jekyll was able to give an account of Hyde’s actions. In doing so, he assumes the role of a spectator.

One might pause here and ask why Stevenson said so little about the nature of Hyde’s diversions. Late Victorian prudery perhaps? Did Hyde’s tastes run to the forbidden varieties of sexual indulgence? Vladimir Nabokov remarks that prudery alone would hardly account for the author’s reticence. It is rather more likely that the artistic difficulties of making sexual licentiousness seem sufficiently evil would have deterred Stevenson from any explicit portrayal.

A certain amiable, jovial, and lighthearted strain running through the pleasures of a gay blade would . . . have been difficult to reconcile with the mediaeval rising as a black scarecrow against a livid sky in the guise of Hyde.⁸

To dwell on the worse sins of cruelty or sadism would have been even more destructive to the fabric of the story. For Jekyll is presented as a good man, whose own inclinations to evil are muted and momentary. He was always aware of the nature of Hyde’s diversions and Hyde could only appear when Jekyll chose to invoke him. If Hyde killed, tortured or maimed during the years when Jekyll enjoyed the privileged status of a spectator the story would altogether lose its moral point. For Jekyll then could lay no claim to goodness. He would be no different from any other calculating criminal who dons a disguise before perpetrating his crimes. The most that we can grant is that Jekyll knew that Hyde was capable of any crime in the sense that he lacked any moral restraints. But it is apparent that Jekyll trusted Hyde to restrain himself, for other reasons, from the worst excesses of which he was capable.

The novel won immediate fame. The underlying moral argument must have struck familiar chords at the time and, in the century which followed its publication, the composite figure of Jekyll and Hyde has acquired an iconographic significance. For most readers, I would guess, Dr Henry Jekyll continues to enlist a measure of moral sympathy. Whatever the degree of his fault he was not as bad as Hyde. And if one makes that judgment one is drawn on, it seems, to concede the essence of Dr Jekyll’s claim that he was not the same person as Hyde. The sheer familiarity of the moral plot of the story engages the reader and it is all too easy to overlook the extreme fragility of the mechanism. For Dr Jekyll was deceiving himself, wasn’t he? There is a significant passage in his posthumous testament in which he describes the first of the involuntary transformations. When Jekyll awoke to find himself transfigured the experience was one of icy horror. He retreated in terror to his laboratory, took the elixir and returned to breakfast in his more familiar guise. Here it is Jekyll’s emotions, not those of Hyde, which are described. And whom shall we say it was who conceived the intention to take the elixir? Dr Jekyll was no mere spectator in this instance. One is still inclined to grant

the sincerity of Jekyll's account of his double life. And if he deceived himself there is still a measure of mitigation to be found in the fact that it was necessary to do so. Better that than the frank avowal of evil, or so one might say. But self deception is an even more curious excuse than the denial of identity with Hyde. If there are not two selves but one, who is deceiving whom?9

The case presents, in an extreme form, problems concerning the relationship between responsibility and personal identity. Dr Jekyll never claimed that he was blameless. He did, however, deny that he was fully responsible for Hyde's actions. He did deny that he and Hyde were the same person and bore the same measure of moral responsibility. This form of excuse is encountered in more mundane cases. Yet the story is no more than a starting point. It presents the issue of personal identity in a form approaching caricature. Stevenson's plot depended on a curious reversal of the usual claim to a dissociation of selves. The most usual pattern is one where the accused maintains that he was unaware of, or at least unable to recall, the actions of the alien self. There is, on the other hand, no case in the legal literature of bodily transformation. The moral or legal relevance of bodily transformation is a matter of some uncertainty and, perhaps, scarcely worth the effort of elucidation: the exigencies of the plot required a substantial difference in the appearance of Dr Jekyll and Mr Hyde. For the sake of discussion I shall propose two variants, or mutants, of the original Dr Jekyll: Jekyll MK I and Jekyll MK II. All three of my characters are alike in one respect. Hyde's actions are, in some radical sense, inconsistent with what is known of the Jekyll personality and mode of behaviour.

**JEKYLL MK I**

- Jekyll and Hyde have the same bodily appearance.
- Jekyll and Hyde have a shared memory.

**JEKYLL MK II**

- Jekyll and Hyde have the same bodily appearance.
- Jekyll has no memory of Hyde's actions.

Other combinations are possible. The psychiatric literature on multiple personality records instances like Jekyll MK I and Jekyll MK II. There are also cases where the memory relationship is asymmetrical. A Hyde may be conscious of Jekyll's actions, though Jekyll remains unaware of Hyde.10 But these further variants do not add significantly to the range of issues raised by my examples.

**A SHORT ANSWER TO THE PROBLEM**

In legal contexts Jekyll's claim that he is not responsible, or not fully responsible, for Hyde's actions attracts a variety of responses. It is usual, however, for

---

9 See the excellent study by Haight M. R., *A Study of Self-Deception* (1980) and the conclusion at page 129: "'self-deception' is contradictory, and so never more than a metaphor; but it may be a metaphor that we need".

10 The empirical literature is collected in Haight M. R., *ibid.* Appendix A.
Jekyll to put his claim in a modified form. The plea that Jekyll is not responsible for the actions of an alien self is commonly recast as a plea that, ‘he could not control himself’. The curious incoherence between the ‘he’ and the ‘him’ in this formulation passes unremarked. The simplest legal response to the plea is to base responsibility on Jekyll’s last act — the taking of the elixir. Herbert Fingarette has provided the most articulate account:

one who knowingly and unreasonably risks loss of his powers of rationality with respect to criminal law, whether by chemical or any other means, should — I propose — be viewed as committing a crime of criminal negligence.11

The term ‘rationality’ is at the heart of Fingarette’s theory of liability. I shall return to it presently. For the moment, however, it is sufficient to indicate Fingarette’s application of this approach to the case of a defendant who killed while under the influence of L.S.D.:

To find the proper basis of culpability, we must turn from the intoxicated act itself to its context of origin. Lipman, while sane and sober, deliberately set about to intoxicate himself, and in doing so he foreseeably, materially, and unreasonably increased the risk of substantial loss of his mental capacity for rational control of his conduct with respect to the law (and indeed with respect to any rational norms.)12

Fingarette’s conclusion is that the defendant must be found guilty of an offence of gross negligence — in this instance manslaughter. More generally, in crimes of so-called ‘basic intent’, such a defendant can be described as ‘reckless’ and this will provide the necessary mental element required for conviction.13

There is a strong tendency in Fingarette’s published work on the subject to seize on the fact of voluntary extinction of the capacity for rational control as a sufficient indicator of guilt in itself. So far as crimes of negligence are concerned it is an argument which outflanks all save the most esoteric forms of opposing argument.14 To equate voluntary extinction of the capacity for rational conduct with recklessness is a more dubious manoeuvre. It is common now in England, however, where courts rely on the device of ‘basic intent’.

Where recklessness must be proved, a more restrictive approach is possible. In Australian courts, a finding of recklessness would require proof that the defendant realized that the criminal act was a likely or probable consequence of intoxication.

In whatever form it takes, however, the essence of this approach to the problem of responsibility involves an inquiry as to the degree of the defendant’s culpability at the time the intoxicant was consumed.

Dr Jekyll could hardly complain if his guilt were to be judged on this basis. In England and in Australia the crime of murder requires proof of an intention to kill

12 Ibid. 278.
14 It might be argued, in reliance on the position adopted by H.L.A. Hart, that an individual who lacked the capacity to take care should not be blamed for negligence. See Hart H.L.A., ‘Negligence, Mens Rea, and Criminal Responsibility; in Hart H.L.A., Punishment and Responsibility: Essays in the Philosophy of Law (1968) 136. If the individual voluntarily impairs that capacity, however, it may be doubted whether the argument is applicable. See ibid. Notes (on the essays) 261-2.
or inflict grievous bodily harm. Actual realization that those consequences were likely will also suffice. But murder is not a crime of negligence or 'basic intent'. It is clear, from the account we have, that Hyde's attack on Sir Danvers Carew was unprecedented in its savagery. The incident involving the little girl gave no warning that a crime of this magnitude was likely. Had Jekyll realized the risk it is sufficiently certain that he would not have undertaken his last voluntary transformation. Jekyll never attempted to dissociate himself from all guilt. He did, however, claim that he was not guilty of the murder of Sir Danvers Carew. At the time he took the elixir he neither intended to kill nor realized that Hyde was likely to cause serious harm.

I have not attempted to distinguish among my gallery of Jekyll mutants in setting out this simple approach to the problem. If all or any of them escape conviction for murder there remains the problem of explaining why this is so. The simple, or pragmatic, approach has the virtue of allowing a conviction of some offence. It would probably allow the conclusion that Jekyll was guilty of manslaughter. But Jekyll would never deny that he bore some measure of guilt for Hyde's actions. The question is whether he is entitled to avoid the charge of murder.

It should not escape notice that the simple or pragmatic approach would treat Jekyll in much the same way as an accomplice whose principal goes beyond the scope of the agreement. Jekyll indeed referred to himself as a 'conniver' with respect to Hyde's actions before the killing of Carew.15 There is, in this simple approach, an implied concession that Jekyll and Hyde were distinct persons.

THE DENIAL OF IDENTITY

There are other forms of legal argument which can be brought to bear on the problem. Some of these will be treated below. All, however, avoid any explicit consideration of Jekyll's claim to separate identity from Hyde. The criminal law analysis depends on concepts of voluntariness, consciousness, intentionality and realization. These have no more than an oblique bearing on the claim that Jekyll is not to be identified with Hyde for the purpose of allocating responsibility for murder. They go to the question whether Jekyll could control Hyde or, if they are treated as one, whether Jekyll could control 'himself'. It is possible that the use of this conceptual vocabulary will be found, in the end, to be unavoidable. It is nevertheless a translation of Dr Jekyll's claim. An examination of what is involved in this process of translation reveals some underlying features of these familiar legal concepts.

There is an account of personal identity in the second edition of John Locke's Essay Concerning Human Understanding, published in 1694. No-one nowadays would hold to the theory in its entirety. It does, however, provide the starting point for modern philosophical analysis. The philosophical debate is divided on the question whether Locke's theory is capable of being saved by amendment or is fundamentally erroneous.16 I have chosen to deal in caricatures and I will accord-

15 The Strange Case of Dr Jekyll and Mr Hyde 87.
Regina v. Jekyll, sub. nom. Hyde

ingly ignore many of the refinements of that debate. Locke’s account of personal identity, whatever its ultimate merits, is peculiarly relevant to issues in current criminal jurisprudence. It was suggested by Professor Brett that, in a more general sense, much of the contemporary analysis of criminal responsibility can be traced back to Locke.¹⁷

The section on personal identity made its appearance in the second edition of Locke’s Essay. It was written at the suggestion of his friend, William Molyneux, of whom I shall have more to say later. Locke’s examination of personal identity had a particular purpose. He was concerned with the idea of the ‘person’ as a ‘forensic’ term. Persons were, in this sense, to be conceived as the objects of rewards or punishments — as moral agents. The term ‘man’, which by the usual extension includes ‘woman’, is quite distinct from the person in Locke’s account. The identity of a ‘man’ consists in,

nothing but a participation of the same continued life, by constantly fleeting particles of matter, in succession vitally united to the same organised body.¹⁸

The man, in Locke’s sense, provides no more than a residence for the person. But it is possible at least to imagine that the same residence might be inhabited by a succession of persons:

should the soul of a prince, carrying with it the consciousness of the prince’s past life, enter and inform the body of a cobbler, as soon as deserted by his own soul, every one sees he would be the same person with the prince, accountable only for the prince’s actions.¹⁹

There are undoubtedly difficulties in the way of this contention which need not detain us here.²⁰ For it is the criterion of identity for persons which is our present concern. Four propositions derived from the Essay indicate the contours of personal identity as an ethical or ‘forensic’ concept.

(1) ‘Consciousness’ is a ‘reflex act of perception’ accompanying action or thought. We ‘cannot think without being conscious of it’.²¹

(2) But ‘consciousness’ extends also to the past. It is more usual nowadays to speak of ‘memory’ and most modern accounts of Locke tend to refer to memory rather than consciousness of the past. It is apparent that Locke thought of consciousness of the past, like consciousness of the present, as a form of perception.

(3) Personal identity, or selfhood, is constituted by present consciousness of the self and consciousness, or memory, of the past life of that self.

¹⁷ Brett P., op. cit. 48.
¹⁸ John Locke, An Essay Concerning Human Understanding (2nd ed. 1694), Book II, Chapter 27, Section 6.
¹⁹ Ibid. I. Section 15.
²⁰ See Williams B., Problems of the Self (1973) 11-2: ‘Suppose a magician is hired to perform the old trick of making the emperor and the peasant become each other. He gets the emperor on his throne and the peasant in the corner, and then casts the spell. What will count as success? Clearly not that after the smoke has cleared the old emperor should be in the corner and the old peasant on the throne. That would be a rather boring trick. The requirement is presumably that the emperor’s body, with the peasant’s personality, should be on the throne, and the peasant’s body with the emperor’s personality, in the corner. What does this mean? In particular, what has happened to the voices? The voice presumably ought to count as a bodily function; yet how would the peasant’s gruff blasphemies be uttered in the emperor’s cultivated tones, or the emperor’s witicisms in the peasant’s growl? A similar point holds for the features.
²¹ John Locke, op. cit. Book II, Chapter 27, Section 12.
The last of the propositions raises a different point. I include it for its obvious bearing on issues of responsibility and for its anticipation of arguments which have currency in contemporary discussions of criminal responsibility. Locke argued that the two principal functions of the mind were 'perception' and 'volition or willing'. Will connotes the power to act, or to refrain from action. The will was, however, under the governance of the 'mind'. Locke accordingly proposes an early version of the 'policeman at the elbow' test:

Nor let any say, he cannot govern his passions, nor hinder them from breaking out, and carrying him into action for what he can do before a prince or a great man, he can do alone, or in the presence of God, if he will.

It follows from the propositions concerning personal identity that, 'Socrates waking and sleeping is not the same person'. And,

to punish Socrates waking for what sleeping Socrates thought, and waking Socrates was never conscious of, would be no more of right than to punish one twin for what his brother twin did, whereas he knew nothing, because their outsides were so alike that they could not be distinguished.

Where there are, 'distinct incommunicable consciousnesses at different times, it is past doubt that the same man would at different times make different persons'. Locke invokes in support of his argument the practice of the courts in the 17th century. The 'sober man', he says, is not punished for what the madman did: they are two distinct persons. That last example may appear to sit a little oddly with Locke's criterion of personal identity. We do not nowadays think of madness as a defect of memory. Recall, however, that Coke defined madness in terms of absence of 'memory and understanding'. And Sir Matthew Hale, who died less than twenty years before the Essay was written, speaks of 'absolute madness' as a 'total deprivation of memory'. The modern conception of the insanity defence was a far later development.

Though memory of past actions is said to provide the basis for present responsibility, it is apparent that Locke was not prepared to allow mere forgetfulness to divide a person from responsibility for his past. Though his account is far from clear, it appears that personal identity was to extend to the limits of what could be remembered by the person. Loss of memory was to be distinguished from mere forgetfulness.

If we now return to Dr Jekyll and Mr Hyde, and the variants on the theme of dissociation which I have sketched, it would seem that Dr Jekyll, the protagonist in the original story, was the same 'person' as Hyde according to Locke's criterion. For they shared a common memory. It may be that one could describe them as different 'men', though Locke's definition of 'man' is probably capable of

---

22 Ibid. Book II, Chapter 6, Section 2.
23 Ibid. Book II, Chapter 21, Section 53.
24 Ibid. Book II, Chapter 27, Section 19.
25 Ibid. Section 20.
28 Perry J., 'The Problems of Personal Identity', in Perry J., op. cit. 3, 16ff. As Perry indicates, more sophisticated amendments of John Locke's position have also been proposed by modern philosophers.
melding them: there was but a single ‘life’ and only one body, though it was changed in form. But even if they were different ‘men’, they would be the same ‘person’, and would not be distinguishable in terms of moral culpability. Locke’s prince remained accountable for his previous actions though he had taken up habitation in the body of a cobbler. If the prince committed murder, presumably the cobbler’s body must hang. In the case of Jekyll MK I the same conclusions would follow without the need to consider distinctions between the man and the person.

However Jekyll MK II, who was not conscious of Hyde’s past actions, would escape moral responsibility. There is nothing necessarily absurd in this, though the conclusion and the distinctions may run counter to one’s intuitions. Locke accepted however that his analysis had some strange results. Since no-one is now likely to accept Locke’s theory of personal identity in its entirety, the justification for using it here is that it may enable the nature of our moral intuitions to be revealed more clearly. There are two consequences of Locke’s analysis which are worth noticing. First it seems a little cavalier to dismiss Dr Jekyll’s claim, or the claim made by Jekyll MK I, on the sole ground that they shared a common memory with Hyde. The claim to avoid full moral responsibility for Hyde’s actions really rests on the ground that Jekyll was, by all accounts, a good man by ordinary standards, whilst Hyde was the epitome of evil. The second consequence of Locke’s analysis is related to the same puzzle. If memory, or consciousness, is the sole criterion for personal identity, Jekyll MK II must escape full moral responsibility for the murder of Carew. I have assumed that Jekyll MK II was a reasonable and virtuous man. But suppose he was just as bad as Hyde? Raymond de Sousa puts the following variation on the story:

When Dr. Jekyll changes into Mr. Hyde, that is a strange and mysterious thing. Are they two people taking turns in one body? But here is something stranger: 

Or. Juggle and Dr. Boggle too, take turns in one body. 

but they are as like as identical twins! 

You balk: why then say that they have changed into one another? Well, why not: if Dr. Jekyll can change into a man as different as Hyde, surely it must be all the easier for Juggle to change into Boggle, who is exactly like him.

We need conflict or strong difference to shake our natural assumption that to one body there corresponds at most one agent.

Yet Locke’s argument would still compel the conclusion that an evil version of Jekyll MK II cannot be held responsible for what Hyde did. It makes the dismissal of the plea made by Jekyll MK I appear all the more arbitrary. The issue is anticipated in Aristotle’s Ethics where a distinction is drawn in terms of voluntariness. Aristotle classified ‘acts done through ignorance’ as ‘involuntary’ or ‘not voluntary’. Involuntariness provided an excuse. But those who acted in ignorance and do not feel compunction or regret afterwards are not said to have acted involuntarily. They are merely ‘non-voluntary’ agents.

All this is by way of setting the scene for an examination of some more modern variations on a theme introduced by Locke. Before leaving the Essay on Human Literature...


30 The Nicomachean Ethics (translated by J. A. K. Thomson, Penguin Books, 1953) Book III, Chapter 1, 80. Aristotle implied that one would neither pity nor pardon in the absence of manifest regret or compunction. The argument is linked with his treatment of intoxication and blameworthiness. See also ibid. 90, 91, 100.
Understanding, however, it is necessary to note a very direct link between the Essay and modern analyses. Locke distinguished between morality and law. He accepted that distinctions which God could be expected to make might well prove impracticable for courts. He was prepared to allow the madman a defence. But he would not do the same for the drunkard or sleepwalker:

Human laws punish both, with a justice suitable to their way of knowledge; because, in these cases, they cannot distinguish certainly what is real, what counterfeit: and so the ignorance in drunkenness or sleep is not admitted as a plea. For, though punishment be annexed to personality, and personality to consciousness, of what he did, yet human judicatures justly punish him, because the fact is proved against him, but want of consciousness cannot be proved for him. But in the great day, wherein the secrets of all hearts shall be laid open, it may be reasonable to think, no one shall be made to answer for what he knows nothing of; but shall receive his doom, his conscience accusing or excusing him.31

William Molyneux, at whose suggestion the chapter on personal identity was written for the second edition of the Essay, criticised an earlier draft of this passage. The exchange of letters on the subject may have strained Locke’s patience a little: there is perhaps a hint of irony in his assurance to Molyneux that he could ‘distinguish between the censures of a judicious friend, and the wrangling of a peevish critic’.32 Molyneux was a distinguished man in his own right: a lawyer, metaphysician, member of the Irish Parliament and the author of works celebrated in his own time. He was also the gadfly critic, embalmed forever in the amber exuded by the wounds which he inflicted on the Essay in the course of its growth. In his first letter Molyneux conceded that the courts might justly punish both the somnambulist and the drunkard. He went on to add, however, that were he to be a member of a jury he would feel morally entitled to acquit the somnambulist, whatever the law might say. There was, in his view, a moral distinction between the cases. Drunkenness, unlike somnambulism, was a voluntary condition. Though courts might deny a defence in both cases, the reasons for doing so were different. Difficulties of proof might explain the rule for somnambulists, but ‘drunkenness being a crime, one crime cannot be alleged in excuse for another’. In his reply Locke repeated the argument that there could be no moral condemnation in the case of a drunken act that was ‘totally and irrecoverably forgotten’. He makes no further comment on the somnambulist but invokes another example. Is there, he asks, any moral distinction between a man who contracts a fever caused by intemperate drinking and, ‘in the frenzy of his disease (which lasted not perhaps above an hour) committed some crime’, and a man who commits a crime, ‘in a drunken frenzy without a fever’? Both were, in Locke’s view, without consciousness of what they had done and were accordingly without moral guilt. The argument provided Molyneux with an irresistible opportunity for reply. In his next letter he took the obvious point: ‘How comes it to pass that Want of Consciousness cannot be proved for a Drunkard as well as for a Frenetic?’ One may be as easily counterfeited as the other. Yet the law allows the frenetic a defence and denies the

31 John Locke, op. cit., Book II, Chapter 27, Section 22.
Regina v. Jekyll, sub. nom. Hyde 379

drunkard's claim to exculpation. Locke did not reply for three months. When he did so the Essay was already printed and bound. The passage which Molyneux had criticised was essentially unaltered. In the end, however, Locke conceded a part of the argument: 'I agree with you that drunkenness being a voluntary defect, want of consciousness ought not to be presumed in favour of the drunkard'. The drunkard, unlike the frenetic, was not entitled to a determination of the issue by human courts of justice. His fault precluded any right to such indulgence. But Locke never resiled from his claim that the amnesic drunkard was blameless before God. He could hardly do so, for that would be to admit that the entire theory of identity was erroneous. The distinction between human and divine justice fore­shadows the position adopted in the early nineteenth century by Bentham who also distinguished between principle and expedient practice in his discussion of drunk­eness.

Neither Locke nor Molyneux descended to the level of detail involved in distinguishing between murder and manslaughter. Given the tendency of the time to consider drunkenness as an aggravation, rather than an excuse or mitigation for crime, one may be reasonably confident that Locke would have conceded that drunkenness, no matter how extreme, would provide no legal defence to murder. If it is assumed that this was an accurate account of the law in Locke’s time it would follow that Jekyll, in all his mutations, would be judged in exactly the same way as the extreme drunkard. In the case of Jekyll MK II, however, Locke’s conclusion would have involved a reluctant compromise between moral innocence and practical expedience engendered by the difficulties of proof in courts of justice.

It is a pity that the no less problematic issue of somnambulism was not pursued in the correspondence between Locke and Molyneux. Molyneux conceded from the outset that the sleep walker had no legal defence. As a lawyer, familiar no doubt with the practice of his time, he was probably correct. He merely argued that the somnambulist and the drunkard were morally distinguishable. It is apparent from his letters, however, that he was inclined to believe that the cases should be legally distinguishable as well.

OUR NEOLITHIC FOREBEARS

Locke’s distinction between the man and the person has been subjected to attack from a variety of philosophical directions. I shall not deal with these. In order to short circuit that discussion I will simply adopt, somewhat arbitrarily, Strawson’s conclusion that, ‘The criteria for personal identity are certainly multiple’. The virtue of that position, which is by no means idiosyncratic, is that it leaves open a

33 Locke did, however, add in the Errata to the completed Essay an amendment in clarification of his views. The passage from the Essay quoted in the text includes this amendment.

34 The Correspondence of John Locke, op. cit. Volume 4, Letter 1693, Locke to Molyneux, 19 January 1694, 784, 785, displays Locke’s awareness of the point: ‘nay, it is an argument against me, for if a man may be punish’d for any crime which he committed when drunk, whereof he is allow’d not to be conscious, it overthrows my hypothesis’.

35 Jeremy Bentham, An Introduction to the Principles of Morals and Legislation (edited by Harrison W., 1960) 284. The ground of principle of which Bentham rested the drunkard’s excuse is, of course, diametrically opposed to Locke’s theories of personal identity and consciousness.

36 See, for example, Williams B., Problems of the Self (1973), Personal Identity and Individuation 1.
variety of issues which seem incapable of philosophical resolution. Strawson goes on to say,

I am not denying that we might, in unusual circumstances, be prepared to speak of two persons alternately sharing a body, or of persons changing bodies &c. But none of these admissions counts against the thesis that the primary concept is that of a type of entity, a person, such that a person necessarily has corporeal attributes as well as other kinds of attributes.37

In this way it is possible to preserve for discussion that particular aspect of personal identity with which Locke was primarily concerned — the relationship between personal identity and the criteria by which we justify the imposition of punishment or the conferring of rewards. The more problematic aspects of Locke’s distinction between the man and the person can be avoided. In this particular context consciousness, and memory, can still be viewed for the sake of argument as essential elements of personal identity.

I have sought to preserve this somewhat attenuated version of Locke’s thesis because it provides an illuminating contrast with a radically different kind of approach to moral responsibility. My text here is provided by Philip Roche in his published version of the Isaac Ray Award Lectures for 1957.38 I choose this because it is the work of a distinguished representative of one influential line of development in forensic psychiatry. It has too the same quality of suggestive caricature as my other texts. The line of descent can be traced back to Freud. This is, however, a far cruder set of hypotheses. And to be fair, Roche conceded the charge of crudity. In this model, ‘normality’ is marked by ‘the extent to which one’s behavior is governed by conscious choices’.39 This area of conscious choice is the domain of the ego. By contrast, ‘One bound to unconscious dominance is bound to the magical orientation of one’s Stone Age ancestors’.40 Or again, the dominance of the unconscious marks, a return to ‘the prelogical projection system of our neolithic forebears’.41

Here it is the ego and the id which are personified. Roche remarks in a disarming aside that:

Some may point a finger at psychiatry for its own kind of demonology, its amoral id. But, in defense, we can say that at least our psychoanalytic id is purely metaphor for the child in us and for that part of us close to our original nature . . . . It may be that so far as our self-knowledge is concerned, William James spoke a disquieting truth when he maintained that ‘scientific theories and postulates are mere fashions and the ancient fashion of postulating natural forces as persons — the demon theory — is . . . certain to return in the long run’.42

Fully conscious choice implies an awareness of the unconscious forces of the id. ‘Normality’ is a concept which admits of degrees. One who is fully self-aware, or fully normal, is engaged in the constant discipline of appraising the demands of reality, the rationalisations by which behavior is justified, and the unconscious component of ‘primitive, magical wish thinking’. Consciousness and conscious control are also a measure of a ‘relative freedom of will’.43

37 Strawson P., Individuals (1964) 133.
39 Ibid. 23.
40 Ibid. 24.
41 Ibid. 18.
42 Ibid. 238.
43 Ibid. 23.
It is, in short, another kind of bifurcation of the self, founded on a metaphor:

we act as if the bad unconscious contends against the good conscious, ‘as if’ within the mind there is a perpetual struggle between the cave man within us and the socially domesticated gentleman. 44

There is, however, no distinction between the person and the man as there was in Locke. Nor is there any single criterion, such as memory, to which reference can be made. The cave man and the gentleman are metaphorical epithets: in a larger sense the ‘self’ is a unity comprised by id, ego and super ego. But the interest of the model, for my present purpose, lies in its reliance on metaphor. Gregory Bateson, in a comment on theories of this kind, wrote:

if we exclude the unconscious processes from the ‘self’ and call them ‘ego-alien’, then these processes take on the subjective colouring of ‘urges’ and ‘forces’; and this pseudo-dynamic quality is then extended to the conscious ‘self’ which attempts to ‘resist’ the ‘forces’ of the unconscious. The ‘self’ thereby becomes itself an organisation of seeming ‘forces’. 45

One can see this tendency in another metaphor employed by Roche:

The psychic forces which we postulate as conscious and unconscious do not operate separately but seem to make their exertions together in relative proportions like those in the parallelograms of forces which conjoin in unitary resultants of behavior. 46

It is a deterministic model. When Roche refers to ‘relative freedom’ of will he refers merely to the extent to which action is mediated by argument, rationality and the demands of social reality. Rational, conscious action is nonetheless determined. Like Locke, however, he is concerned with the issue of reward and punishment. For he argues that consciously determined behavior is capable of modification by threats of punishment or promises of reward. Unconsciously determined behavior, on the other hand, is governed by ‘archaic patterns, inflexible and repetitive’ and relatively uninfluenced by threats and promises. 47

This is not, as it stands, an analysis of moral responsibility. For Roche the whole apparatus of moral analysis on which the doctrines of criminal law have been constructed is pre-scientific — a persisting remnant of primitivism. For the most part, Roche is concerned with the linguistic mechanisms which sometimes enable and more often impede, the translation of psychiatric concepts into the primitive or magical language of responsibility. Much of his analysis is devoted to the legal concept of insanity, and the various tests which fix its area of operation. The tests I shall ignore. There is a more fundamental point to be made. The insanity plea is something more than a defence to criminal liability. Until recently it also provided a definitional boundary for an area of discourse. It was only in this area, and in the related context of pleas of unfitness to stand trial, that criminal courts were required to attempt a translation of psychiatric analyses into legal concepts. Evidentiary rules, for the most part, precluded the use of psychiatric expertise in cases where the defendant chose to present himself as a normal, sane, or reason-

44 Ibid. 27.
46 Roche P., op. cit. 24.
47 Ibid.
able man. We may not have reached the limits to which the concept of insanity can be extended. It is quite apparent, however, that the use of psychiatric testimony can no longer be confined within this area. The various kinds of plea which fall under the heading of automatism will almost always require the presentation of expert evidence if the issue is to go to the jury. Analyses of the kind presented by Roche have no essential link with insanity or disease of the mind. Stress or acute intoxication, physical illness or physical injury may all, in Roche’s terms, affect the balance of conscious and unconscious forces. He refers, for example, to ‘acute alcoholic intoxication which dissolves the outer shell and exposes the inside man’.48

There is an obvious sense in which this analysis seems to fit the case of Jekyll and Hyde. Yet the nature of the link between the analysis and conclusions expressed in terms of moral or legal responsibility remains obscure. One could ask, of course, whether Jekyll was sane. He was not obviously insane, though Stevenson’s account of the case suggests the possibility of a variant in which another mutant of Jekyll suffers from the delusion that the consumption of an innocuous substance makes all the difference to his behaviour.49 I suggested at the outset that Hyde was certainly guilty of murder, whatever one might conclude about Jekyll. There is, perhaps, a slight doubt on this point. Suppose Hyde had not cheated the courts by taking his own life and suppose, too, that Jekyll was beyond recall. Hyde is described most often by Stevenson as the epitome of evil. Yet his act of killing Carew was also described as an act of ‘moral insanity’. That is merely an expression of course. If it were possible to imagine a person who was the epitome of evil, could such a being be held morally responsible for his actions? Hyde was perfectly capable of restraining himself, had he perceived a serious risk of detection. He was nothing if not calculating. He was, however, quite incapable of remorse,50 or any distinctively moral response. There is at least an argument that Hyde inhabited that disputed territory between guilt and insanity which is called psychopathy.51 Roche refers to the dilemma:

To some extent we all lead double lives; most of us succeed in confining our magic to our dream life; in others the magic comes to wakeful symbolic expression in atavistic sexual and aggressive behavior and we have not yet made up our minds whether such people are insane or plain wicked.52

But this talk of sanity and insanity does nothing to elucidate the relationship between the psychiatric analysis and moral, or criminal, responsibility. For

48 Ibid. 9. for a more careful statement of the argument, see Halleck S., Psychiatry and the Dilemmas of Crime (1967) 159ff.
49 It is a distinctive feature of Stevenson’s novel that the elixir produced major and observable physical changes when it was taken. But Henry Jekyll was admirably clear in his appreciation of its true effect: ‘Had I approached my discovery in a more noble spirit, had I risked the experiment while under the empire of generous or pious aspirations, all must have been otherwise, and from these agonies of death and birth I had come forth an angel instead of a fiend. The drug had no discriminating action; it was neither diabolical nor divine; it but shook the doors of the prison house of my disposition; and, like the captives of Philippi, that which stood within ran out.’ Op. cit. 85.
50 See Thalberg, Enigmas of Agency: Studies in the Philosophy of Human Action (1972) 205-6, ‘Remorse is a moral emotion par excellence. It is the only emotion that has as its objects nothing but intentional actions which the agent has performed in violation of his moral principles’.
51 For an analysis which combines clarity and intellectual sophistication, see Walker N., McCabe S., Crime and Insanity in England Volume II (1973) 205-37.
52 Roche P., op. cit. 19.
Roche, a request for such a link would be met with the reply that this is not the province of psychiatry. A scientific discipline — the description which he insists on applying to psychiatry — cannot incorporate notions drawn from the primitive word magic of criminal jurisprudence. There is nevertheless one suggested link between these two forms of discourse. It is to be found in the proposition that those whose actions are dominated by unconscious forces are 'uninfluenced in communication through appeal, argument, rewards and punishments and are exceptionally alterable only through the arduous work of analysis'.

If it is taken literally the last suggestion is obviously false. We do not confine our judgments of non-responsibility to those who cannot be influenced by threats or rewards. And, conversely, institutions for the criminally insane are commonly administered, like other 'total institutions', by judicious use of threats and promises. There is, however, another sense in which Roche's analysis can be translated into the language of moral or legal responsibility. It is to be found in his emphasis on conscious, or rational, control of behavior. The concept of 'rationality', has been the subject of exhaustive analysis by the philosopher, Herbert Fingarette. The idea is at least implicit in the last quotation from Roche. It is more obvious in the work of another eminent American forensic psychiatrist, Seymour Halleck who remarks, 'Ultimately most of our decisions to call people mentally ill are based upon judgements of unreasonableness'. The offender's behaviour may 'suggest goals or motivations which seem so unacceptable and so incomprehensible that they are readily looked upon as unreasonable'. Or the offender's goal may be 'reasonable', in the sense that it involves the acquisition of property or status, but the execution of the crime is marked by a failure to pursue the goal in a 'logical, consistent or effective manner'. That will in turn suggest that the individual may be 'driven by motivations which are not apparent and which deviate from those which society would consider reasonable'. Though the behaviour is 'unreasonable', it does have a meaning and it is purposive. It is not random or inexplicable. Roche, Halleck and, for that matter Dr Henry Jekyll, would all agree that Hyde was in Jekyll and not in the bottle. But the meaning or purpose of the behavior is obscured by its apparent unreason:

When the society asks a psychiatrist to explain unreasonable behavior, it is really asking him to discover the rationality (those factors which consciously or unconsciously influence the individual) which lies behind a behavior the community cannot accept or understand.

The reference to 'rationality' in this quotation from Halleck does no more than indicate that the behaviour has a rationale of which the offender may, or may not,
be aware. Herbert Fingarette uses 'rationality' in a more restrictive sense, corresponding to Halleck's notion of 'unreasonableness', in formulating a criterion for responsibility:

accountability is lessened to the extent that there is incapacity for rational assessment of the nature of one's act with respect to public norms of wrong. 60

To appreciate how this criterion is supposed to coincide with moral, or legal, judgments it is necessary to consider one of Fingarette's examples. Jones and Smith are walking in the woods. Jones picks up a stone. Then he sees Smith and it occurs to him that Smith would make a perfect target. He throws the stone at Smith in order to test his aim. Smith is injured and Jones steals away in order to escape detection. We are told that, 'The suffering of others, in its immediate emotional-moral relevance, is beyond his ken'.61 Jones is not responsible, according to Fingarette, because he is blind to the moral meaning of his action:

If the meaning of law is at bottom rooted in moral meaning, then genuine responsiveness to law requires genuine responsiveness to moral meaning, to moral relevance which is plain and pressing, as in the case of throwing a rock at Smith.62

The criterion of rationality has no necessary connection with disease of the mind. Fingarette is concerned to explain why such a disease will sometimes be relevant to issues of responsibility. But rationality, in Fingarette's sense, may be absent, though the individual is sane.

The major problem, as Roche and Fingarette both acknowledge, is that the criterion of rationality would provide a ground for moral, and legal, exculpation of individuals who appear to be paradigmatic examples of wickedness. 63 If Fingarette's stone-thrower is not accountable, what of Edward Hyde? The stone-thrower was described as a schizophrenic, but that hardly advances the argument. Fingarette is concerned to explain why his schizophrenia is morally relevant. Edward Hyde was equally incapable of 'genuine responsiveness to moral meaning'. And, like the schizophrenic, he was a mere victim of fate so far as his moral condition was concerned. He was called into existence at the whim of Dr Henry Jekyll.

These morally unsympathetic cases, and similar examples which can be readily imagined, may simply illustrate that Fingarette's concept of rationality provides no acceptable criterion for responsibility. Even if it were accepted, there would be a practical limit to its application. There is a parable elsewhere in the work of Robert Louis Stevenson which suggests an alternative form of judgment for Edward Hyde.

Once upon a time the devil stayed at an inn, where no-one knew him, for they were people whose education had been neglected. He was bent on mischief, and for a time kept everybody by the ears. But at last the innkeeper set a watch upon the devil and took him in, in fact.

The innkeeper got a rope's end.

60 Ibid. 201. n. 15.
61 Ibid. 188. In the example Jones is said to be schizophrenic. But that is not the ground of exculpation: the disease is merely invoked to provide an explanatory context in which Jones's attitude to other people becomes credible.
62 Ibid. 189.
A criterion of rationality might provide a basis for distinguishing between different forms of custodial disposition. In legal contexts the distinction between guilt and insanity is familiar. Fingarette’s stone-thrower might be found not guilty but insane for he did suffer from a disease of the mind. It would require a considerable act of faith, however, to suppose that acts of callous and unnecessary cruelty are always associated with mental illness. It is not impossible to imagine a sane version of the stone-thrower. The tension between the argument that such an individual is not morally responsible, and the argument that he is simply wicked, finds its expressions in the dispute over the status of psychopathy as a disease of the mind.

Another philosopher, Joel Feinberg, has presented a more restrictive variant of the rationality criterion. Like Roche and Fingarette he makes the point that we are at first inclined to ‘consider the senselessness of a crime [as] a kind of moral aggravation’. Yet senselessness can be seen as a possible ground for mitigation or exculpation. He suggests an example:

The fetishist’s shoplifting is not rational and self serving; he attains no economic objective by it. But neither does it hurt anyone he hates nor help anyone he loves; it neither gains him good will and prestige, nor satisfies his conscience, nor fulfills his ideals. It is, in short, not interested behavior.

What the fetishist does appear senseless to others. But the point of the example, as it is developed by Feinberg, is that the episode is senseless from the fetishist’s point of view as well. He does not know why he does these things: his life seems incoherent, even to himself.

Moreover, the ‘senseless’ desires, because they do not cohere, are likely to seem alien, not fully expressive of their owner’s essential character. . . . it is as if he were acting on somebody else’s desires. And, indeed, the alien desires may have a distinct kind of unifying character of their own, as if a new person were grafted on to the old one.

Feinberg argues that punishment is inappropriate in these cases. The argument does not, however, depend on any assumption that the fetishist is incapable of controlling himself, or that punishment will not deter him from these acts in future. It is rather that to blame or condemn him is unfairly to identify him with his alien self. The fetishist puts his moral character in issue and claims that his senseless actions cannot be attributed to his real self. The schizophrenic stone-thrower, or a psychopathic variant of the stone-thrower, could not put their claim in quite this way. There is no ‘real self’ to provide the contrast.

---

66 Ibid.
67 Ibid., 288. It would be a delicate exercise in psychosocial analysis, however, to distinguish among fetishes and the variety of forms which the impulse to collect things may take. It is not quite right to categorise the philatelist, however passionate his involvement, as a stamp-fetishist.
This is closer to the argument advanced by Jekyll. Yet he was, in the beginning, unlike the fetishist. Jekyll led a double life for years. There was nothing senseless or irrational about it. He enjoyed Hyde and he cultivated him and, until near the end, he had no reason to fear exposure. There was none of the pathos of Feinberg’s fetishist during those years. Many rational people would envy such a life. Only the murder of Carew by the exultant and uncaged Hyde can be said to have been senseless from Jekyll’s point of view. In one sense, Jekyll could not ask why Hyde should kill Carew. He knew Hyde’s capacity for evil. What he could not understand, however, was the reason why Hyde abandoned himself to a homicidal impulse when prudence should have restrained him. Jekyll must have trusted Hyde until the last months of their joint life. There were two apparent reasons for doing so. Though Hyde had no moral sense, he was calculating. He needed the guise of Jekyll as his refuge. Moreover the possibility of his reappearance, until those last months, depended on a certain level of restraint in his behavior. We may say that Hyde gained the upper hand. It is a parable of Jekyll’s moral deterioration. The end result is mutual loathing, and it is here that Feinberg’s argument takes hold. Jekyll’s denial of identity with Hyde is a plea for confirmation of Jekyll’s continuing existence as an essentially moral being.

None of these writers would suggest that the individual who responds to his unconscious impulses, or who behaves ‘irrationally’, is to be excused on the ground that his actions were compelled or determined. The individual is taken to have engaged in intentional conduct which is the outcome of a choice, albeit an irrational choice. Nor is the criterion for responsibility to be found by asking whether the individual was capable of being influenced by threats of punishment. Roche, to be sure, is equivocal on the point. Halleck, Feinberg and Fingarette are not. They argue that moral and criminal responsibility require something more than the mere possibility of deterring the individual by threats of punishment. Roche, to be sure, is equivocal on the point. Halleck, Feinberg and Fingarette are not. They argue that moral and criminal responsibility require something more than the mere possibility of deterring the individual by threats of punishment. The criteria for moral responsibility suggested by these writers are more generous than anything to be found in Locke. Cases where the self is divided so that, in Locke’s sense, there is no continuity of consciousness or memory would provide examples of irrationality. But this criterion for testing responsibility also extends to cases where there is no amnesia. Locke, on the other hand, expected the individual to govern those passions of which he was aware.

There is another distinction between the psychoanalytic texts and Locke. The division of the self to which Roche refers is always presented as a metaphorical description of the phenomena. It could hardly be otherwise since, in Roche’s terms, the dominance of either the caveman, or the gentleman, is always relative. There are no distinct persons, resident in the same body. The caveman and the gentleman are aspects of the same self. A psychoanalytic approach to the case of Jekyll and Hyde would involve the hypothesis of a transcendent self which unites

69 There is a clear distinction here between the position of these writers and classical utilitarianism. Bentham, op. cit. 284, based exculpation on the ground that punishment would be inefficacious in certain instances of wrongdoing by the insane, the intoxicated or by infants.
these manifestations of the conscious and unconscious mind. This transcendent self is, in one sense of the word, ‘responsible’ for all of its consciously and unconsciously determined manifestations — much in the way that Freud suggested that we are all of us responsible for our dreams.70

Locke’s invocation of memory as a criterion for personal identity may make little sense from a psychoanalytic point of view. It neglects a sense of the idea of responsibility for one’s actions which is immanent in the Freudian tradition. There is, nevertheless, an area of living controversy in the application of these divergent traditions to issues of criminal responsibility.

**LOCKE: MEMORY AND THE REPRESSION OF MEMORY**

If it were possible to imagine a case of, ‘distinct incommunicable consciousnesses’ there would be, according to Locke, two or more distinct persons, neither of whom could be held responsible for the actions of the other. His examples tend to reflect the idea of possession by an alien self. The example of the prince and the cobbler is not explicable as a mere dissociation of personality. His account, as one might expect, takes no cognisance of the phenomena of repression. But mere forgetfulness for incidents in the past would not be taken to destroy the continuity of selfhood. The forgetful individual can be reminded or often he can, by effort of attention, recall parts of the past which seemed lost. Some modern attempts to restate Locke’s theory go further and take into account the hypothesis that painful memories may be repressed. In these cases too, one might say in amending Locke’s account, there are not distinct and incommunicable consciousnesses.

When Phillip Roche suggests that those whose actions result from the primitive promptings of the unconscious are not fit subjects for punishment, he adds a rider. They can, on occasion, be affected and their lives can be altered, ‘through the arduous work of analysis’. The metaphors of possession by an alien self have no place in the analytic enterprise. Whilst Roche argues that these individuals should not be liable to institutionally administered forms of punishment, it may be an object of therapy to bring them to the point of accepting responsibility for all the manifestations of self. The metaphor of the caveman and the gentleman has a forensic rather than a therapeutic rationale.

The assumptions of Freudian analysis have been examined by Alistair MacIntyre who stresses the descriptive element in Freud’s achievement:

> In the pre-Freudian — the novelist’s, use of ‘unconscious’, what is unconscious is what the agent does not recognise, although others may. In Freud’s use what is unconscious goes unrecognised for what is both by the agent and by everyone else — until Freud fashions his own diagnostic technique.71

Conduct which appeared irrational and inexplicable is revealed as an expression of unconscious intentions and purposes. The unconscious is not merely a receptacle for repressed material. Freud emphasised that the necessary economies of mental life also required a concept of the ‘preconscious’ — that which could be

---


recalled by the individual in appropriate circumstances. It is the phenomenon of repression, however, which gives rise to what Freud perceived as a paradox in the therapeutic endeavour:

The patient, who is suffering so much . . . who is ready to undertake so many sacrifices in time, money, effort and self discipline in order to be freed from those symptoms — we are to believe that this same patient puts up a struggle in the interest of his illness against the person who is helping him. How improbable this must sound!72

The psychoanalytical setting, and the techniques adopted by the psychoanalyst enable the repressed material to be brought into consciousness. In the end the patient will avow those unconscious intentions and purposes which have been repressed:

We instruct the patient to put himself into a state of quiet, unreflecting self-observation, and to report to us whatever internal perceptions he is able to make — feelings, thoughts, memories — in the order in which they occur to him. At the same time we warn him expressly against giving way to any motive which would lead him to make a selection among those associations or to exclude any of them, whether on the ground that it is too disagreeable or too indiscreet to say, or that it is too unimportant or irrelevant, or that it is nonsensical and need not be said. We urge him always to follow only the surface of his consciousness and to leave aside any criticism of what he finds . . . 73

In this classical exposition the skills of the psychoanalyst include the art of a highly specialised form of interrogation. Freud’s account deals with only one of a variety of techniques which may be employed. During the late sixties and early seventies the hallucinogens, L.S.D. and mescaline, were frequently used to hasten the process of analysis. The use of these drugs in psychotherapy was common in the United States. In Australia their use was frequent in Victoria where more than 4000 individuals were given L.S.D. for therapeutic purposes. An American analyst, Dr Sidney Cohen explained the rationale: ‘Under L.S.D. the most devastating of buried memories have been recovered and within a single session thoroughly relived and resolved’.74

We may suppose that these techniques are powerful. Cohen also refers to the use of the hallucinogenic drugs by intelligence organisations for the purposes of interrogation and ‘brainwashing’.75 The sinister processes of political interrogation are an elaborate, and often effective, travesty of techniques employed in the consulting room. It is a commonplace in the literature of political interrogation that the subject may be brought to the point of sincerely avowing responsibility for imaginary harms. That is, often enough, the object of the enterprise. The purposes of psychotherapy are, by contrast, beneficial to the patient. There is, nevertheless, a parallel. There is no guarantee, other than the patient’s avowal, of the truth of the memories recovered in analysis. And an avowal, no matter how sincere, or therapeutically useful, is no guarantee at all.76

The possibility that our memories, or our consciousness, of past states of mind may be false has been advanced as an objection to Locke’s account of personal

73 Ibid.
75 Ibid. [131ff. For a more extended account see Marks J., The Search for the ‘Manchurian Candidate’: The C.I.A. and Mind Control (1979).]
identity. It was an objection anticipated by Locke, though his reply to it rests on a contorted and pathetic affirmation of faith:

that it never is so, will by us, till we have clearer views of the nature of thinking substances, be best resolved into the goodness of God, who, as far as the happiness or misery of any of his sensible creatures is concerned, will not, by a fatal error of theirs, transfer from one to another that consciousness which draws reward or punishment with it.

These reflections on psychoanalysis and political interrogation can be contrasted with the processes of the criminal law. The net of protective rules which constrains police investigations, and the trial of the accused, can be counted as so many impediments to the recovery of buried memories. Forcible measures are precluded. The use of drugs without the consent of the suspect is anathema. And one can hardly expect an accused to indulge in the uncritical revelations appropriate to a psychiatric consulting room. There is every inducement to deny an inculpative interpretation of behaviour and to repress any consciousness of incriminating behaviour.

When Locke concluded that the sleepwalker and the drunkard must be punished by human courts of justice though they may have no memory of their offences, he did so on the ground that, 'want of consciousness cannot be proved in such cases'. That was an expression of the prevailing view of the time which cast the burden of proving absence of a guilty mind on the accused. His argument anticipates a more modern theme however.

Questions about the allocation of the burden of proof are secondary. The first issue is whether an issue is capable of proof. An accused may lie about his mental state at the relevant time or seek to conceal it by silence. Here he knows the truth and it is the object of the trial to establish a version of the facts which coincides with the knowledge which he conceals. But denials of intention, memory or other elements which go towards the establishment of guilt may rest on self deception. The accused may give every appearance of sincerely believing that his account is true though the evidence indicates its falsity. Self deception, in its more extreme forms, may be seen as a manifestation of unconscious strategies which the individual is powerless to acknowledge. The paradox of self deception, if we admit that there is such a thing at all, rests on the metaphor of the divided self. It assumes that the self deceiver really does know the truth if only he would admit it, even to himself. There is, however, no way of ensuring that this will happen by any means which can be countenanced in a system of criminal justice. The investigative processes of the criminal law provide no exact counterpart to the 'arduous work of analysis', or to the more sophisticated forms of political interrogation. Nor should it be assumed that these are necessarily more effective in ascertainment of truth. They may be more effective as ways of eliciting admissions. But the individual who is brought to the point of avowing guilt may merely have been led to the point of adopting yet another unconscious strategy of self deception. In these judgments of responsibility the idea of truth is itself problematic. And it is the

77 Flew A., Locke and the Problem of Personal Identity (1951) 26 Philosophy 53.
problematic nature of truth in this area which lies behind Locke’s argument about the difficulties of proof in human courts of justice.

There is, it seems, a distinctive element in judgments of legal responsibility. We are driven to accept that guilt may be established beyond reasonable doubt no matter how sincerely the accused may protest his innocence. A jury may have to decide, for example, whether a gun was fired intentionally or went off by accident. Where the facts are ambiguous the jury must give the accused the benefit of the doubt. It is also possible, however, for the jury to conclude that the facts are not ambiguous and that they compel the conclusion that the accused fired intentionally. The sincerity of his denial must be taken into account. It is not conclusive. Nor does the verdict necessarily entail the conclusion that he consciously lied to the court in asserting that it was an accident.

In these cases, one might say, the individual does not remember at all. He creates an alternative version of the past and represses the truth. Alternatively he may claim that he has no memory of the events in issue. He can, in either case, dispute his guilt but he cannot acknowledge it. The conclusion that he is guilty may, and perhaps should, leave one with a persisting sense of uneasiness. Sir Walter Moberly refers to a common element in arguments which seek to justify the imposition of punishment:

‘punishment’ is only possible where it can be felt by the culprit to be deserved. If there are men wholly incorrigible, they will be capable of no sense of shame or disgrace. Punishment will only operate as deserved pain where, in spite of his misdeed, the wrongdoer still has some vestige, at least, of moral health.79

There is an implied reference to the problem of the psychopath here. But the psychopath is thought to be incorrigible because he lacks a sense of moral error. Those who cannot, or will not, acknowledge their guilt may be incorrigible in another sense of the word. For they cannot, or will not, acknowledge what they have done even to themselves. Nor is the problem confined to cases where the individual suffers from some disorder of memory. Dr Jekyll would concede that he was responsible for what Jekyll had done: he denied responsibility for what Hyde had done. Does it follow that there is no moral justification for legal punishment in these cases? In a passing reference to the problem of dual personality Moberly remarks that, ‘He is both Jekyll and Hyde; and as Jekyll, he can approve and respect that moral order which, as Hyde, he has outraged’.80 That is all the more reason, of course, for Jekyll to deny that he is Hyde. The justification for punishment which Moberly advances is curiously revealing:

The extent to which it may truly be said of all wrongdoers that ‘they know not what they do’ is one of the most vexed questions of moral psychology. But, beyond doubt, there often is some refusal to attend to unpleasant consequences. The wrongness of an act and its unpleasant consequences are known in a sense, but there is a reluctance to look them in the face, which may be partly subconscious, but which may entail something like a resolute effort of will worthy of a better cause. It is this condition of mind which may be broken down by punishment.81

The last sentence has a slightly sinister ring. Yet the individual who persists in his denials in the face of the evidence, and in spite of punishment, must be allowed

80 Ibid. 115.
81 Ibid. 207-8.
to do so. Compelled acknowledgement of guilt is a characteristic of totalitarian regimes. Human courts cannot arrogate to themselves the privileges of divine justice and it is for this reason that acceptance of liability to punishment as one’s desert cannot be advanced as a necessary element in the justification of legal punishment.

THE LEGAL MATRIX

If we put aside issues over the insanity defence and definitions of mental illness, there appears to be no explicit recognition of excuses based on a division of the self in judicial decisions. Dr Jekyll’s plea must be formulated in terms of consciousness, voluntariness and the capacity to form, or have, relevant states of mind. There are three issues at least which are suggested by the preceding material. The first is whether the characteristic modes of legal analysis can add anything to an understanding of Jekyll’s claim that he was morally distinguishable from Hyde. The second is whether it is politic or practicable to provide him with a defence, or qualified defence, to a charge of murder. The third issue is more subtle. We may find that Jekyll’s claim is recognised by courts, though it would be formulated quite differently. And this may be so whether Jekyll’s claim is morally justified or not. To what extent has this occurred?

It is possible to distinguish, in a general way, between two kinds of exculpatory analysis in the legal material. The first involves a range of excuses which depend on a denial of consciousness and memory with respect to the criminal act. The second involves the denial of voluntariness. Often, of course, the two run together, as they do in the traditional analysis of somnambulistic homicide. But the distinction is apparent in the usual form of direction to juries in Victoria. The jury is instructed that the acts of the accused must be shown to have been ‘conscious and voluntary’ if guilt is to be attributed to him. There is a considerable degree of tension between these concepts of consciousness and voluntariness. They correspond, very roughly, with the themes which I have drawn from John Locke and Phillip Roche. It is this tension which I want to explore in the remainder of this paper.

A. Conscious Action

Consciousness and its derivatives are words of multiple application. I will take a small range of usages, excluding all those in which consciousness is ‘raised’, ‘lowered’, or attributed to races, categories of sexual gender, nations or other collectives. Legal analyses appear to be primarily concerned with those usages in which the word, or a derivative, can be qualified by the prefix ‘un’. There are at least three relevant senses:

the conscious mind
(b) (to be) conscious
(c) (to be) conscious
of something.

the unconscious mind
(to be) unconscious
(to be) unconscious
of something.

The first sense is clearly an alien in this company, if only because it assumes an acceptance of psychoanalytic theories of the self. The second and third senses depend less on theory and more on common habits of usage. They differ from the first in another way. References to the 'conscious mind' and the 'unconscious mind' suggest the existence of some definable borderline between the two concepts.83 By contrast, consciousness, or consciousness of something, are relative concepts. One can be completely, deeply or partially conscious or unconscious. And one can be fully, or dimly, conscious of some fact, risk or moral precept. To say that an individual is conscious of something, or unconscious of something, has a faintly archaic air. These expressions are probably less common than the use of the related terms, ‘awareness’ and ‘unawareness’. For most purposes, however, these locutions are equivalent.84

The fact that one can be more or less conscious or aware of something provides a contrast with the concept of intentionality. It is possible for someone to say that he ‘fully intends’ to kill his enemy before he does so. He means that this is no idle threat. He may announce that he ‘half intended’ to kill his enemy if his resolve weakened at the end. But if he does it, we will not allow him to quantify the degree of his intention.

This distinction between the language of intention and the language of consciousness suggests the existence of some problems in the elucidation of the notion that certain states of consciousness may preclude the capacity to form intentions. For the moment, however, I shall put these aside.

In the case of Jekyll and Hyde related by Stevenson there appears to be no ground for saying that Dr Jekyll was unconscious at any stage. Nor, if the identity problem is disregarded, is there any reason for saying that he was unconscious of what ‘he’ was doing. So also with Jekyll MK I, who shared his memory with Hyde. But Jekyll MK II would no doubt claim to have been unconscious when Hyde killed Carew. In legal terms it would be a plea of automatism. If it was accepted, the effect would be to push the enquiry on the issue of guilt back to the stage where the elixir was consumed. There is, however, no compelling reason for accepting this failure of memory as evidence of automatism. The mere fact that Jekyll MK II could not remember the killing of Carew does not seem quite enough. Locke, it is true, argued that this would constitute an excuse — as a matter of morality. The limits of the self as a moral agent were said to extend only so far as the person’s memory extended. In an American case, U.S. v. Olvera,85 there is a sophisticated rejection of Locke’s argument:

83 See the discussion in Sigmund Freud, *The Ego and the Id* (1923), which provides a summation of Freud’s theories on psychic topography.
84 Careful attention to habits of usage will reveal a wealth of possible distinctions however. For an examination of the finer shades of meaning see White, *Attention* (1964) Chapter IV.
85 (1954) 4 U.S.C.M.A. 134; 15 C.M.R. 134. The judgment is one of a series given by the United States Court of Military Appeals on the issues of intoxication, amnesia and criminal responsibility.
If anything, the guilt feelings which serve to produce the repression actually reflect the circumstance that the individual, in a very real sense, originally ‘knew’ the deed to be wrong — and yet chose to perpetrate it. Indeed, he does not wish to live with the recollection — and thus, to avoid anxiety, he buries it mentally.

It is, one might say, the same ‘person’ throughout. And there is at least a tendency to hold him responsible not only for the act, but also for the fact that he has repressed his memory of it. One may be disinclined to excuse an individual who is, in effect, seeking to evade responsibility for his actions. On the practical side, it is worth remarking that the phenomenon of repression may be comparatively common. In a study of 50 convicted English murderers, O’Connell found that 20 had either no memory of the actions which led to their arrest, or very impaired memory.86

One could amend Locke’s account. There is a commonly drawn distinction between psychogenic and organic amnesias and it might be said that the individual who has repressed his memory of an incident ‘could’ remember, though some form of specialised treatment might be necessary to enable him to do so. Compare, however, the facts of another American case, Cummins v. Price.87 Raymond Cummins was a spurned suitor who apparently killed his girlfriend and then attempted suicide. A gunshot wound to the head produced anterograde amnesia for the incident. It is quite possible, in cases of this kind, that the organic damage to the brain will put all memory of the incident beyond the possibility of recall. The case is reported only for the decision that Cummins was fit to stand trial despite his amnesia. On the issue of responsibility for the killing, however, such a defendant might advance a moral argument suggested by Helen Silving:

Since . . . . ‘intent’ exists today only as a recollection, it is hardly possible to separate the intent from the recollection. That ‘intent’, phenomenologically, is a recollection.88

This has obvious Lockean overtones. It would follow that Cummins could not be held responsible for murder because ‘his’ intention cannot be proved. It has ceased to exist. Equally obviously it is not an argument which would be accepted by a court.89 It might be rejected for reasons of policy, as Locke suggested. Loss of memory is too easily feigned. There are, one might add, no sharp distinctions between organic amnesia, psychogenic amnesia and feigned loss of memory. It is also possible to meet the moral argument more directly. In the Olvera case the court pointed out that the defendant was ‘certainly able to analyse rationally the probabilities of his having committed the offence in light of his own knowledge of his character and propensities’. Moreover:

he will presumably be able to remember whatever punishment he may receive — with the result that its deterrent effect will be present as to future danger of criminal promptings.90

Though the argument is stated shortly and there is a shade of brutality about it, one might see this as an attempt to provide some moral justification for legal

89 See, for example, Tsigos [1964-65] N.S.W.R. 1607.
condemnation. It is reminiscent of Moberly’s treatment of the issue. For the most part, however, Locke’s argument that memory is a criterion for moral responsibility has not been the subject of articulate response from the courts. It is simply accepted that amnesia alone provides no basis for a defence. Jekyll MK II would be required to go further and establish that he was ‘not conscious’ at the time of the killing.

There is an exceptional category of cases, however, in which Locke’s argument may find an analogy in the criminal law. Roles are assigned to individuals and, not infrequently, legal duties are imposed as a consequence of those roles. The parent must tend the child, the soldier must return from leave and the prisoner must remain in custody. An omission to perform these duties may amount to an offence. Ignorance of duty is no excuse. But a loss of memory for those elements of personal identity which constitute the role should perhaps be inconsistent with a finding of guilt. The prisoner who cannot remember who he is may simply wander off, with no intention of escaping. He is not conscious of his identity as a prisoner. The question whether he is ‘conscious’ or ‘unconscious’, appears irrelevant.91

If these exceptional cases are put aside the plea of unconscious action, or automatism, is a familiar one. It has prompted a surprisingly large number of decisions, particularly in Victoria.

There are two kinds of case which immediately enlist moral sympathy. There is the case of the somnambulistic homicide, best exemplified in the curious case of Cogdon,92 in which a mother killed her adolescent daughter whilst dreaming. There is also the case of the defendant who is hit on the head and thereafter commits a criminal act whilst in a state of ‘post-traumatic automatism’. Footballers who commit some criminal act after being hit on the head provide sympathetic examples.93 Epilepsy, cerebral arterio-sclerosis and other disorders of organic origin have also been pleaded as a basis for automatism. In this last group of cases, however, the issue tends to involve the insanity plea and the tests for distinguishing cases of (sane) automatism. If issues involving the question of the defendant’s sanity are put aside the most striking development is to be found in the Victorian decision in Keogh94 that voluntary intoxication might provide a basis for a defence of automatism. This extension depends on the analogy between sleepwalking and extreme intoxication which formed the basis of the dispute between Locke and Molyneux. It is in this context, too, that the question whether the defendant was ‘capable’ of forming an intention to kill, injure, rape or steal has most often arisen.95 For if the defendant who acted in a state of acute intoxication is

91 See Scott [1967] V.R. 276, where the appellant who was a member of a prison working party claimed that he had been hit on the head by another prisoner. Thereafter he remembered nothing for two days when, finding himself at large, he decided not to return to captivity. The judgments are, however, couched in terms of ‘consciousness’ and ‘voluntariness’.
92 (1950) Unreported decision of the Supreme Court of Victoria; noted in Morris N., Somnambulistic Homicide: Ghosts, Spiders, and North Koreans (1951) 5 Res Judicatae 29.
supposed to be 'unconscious', then it would follow that he was incapable of intending anything at all.

There is a simple objection to this form of analysis: it is unintelligible. It is not at all difficult to understand how intoxication may be relevant when one is attempting to discover what an individual intended when he acted. The effects of drugs vary of course. But many intoxicants will make the user more prone to accidents, mistakes, blunders and failure to realize the consequences of his action than he is when not under their influence. Some intoxicants may induce hallucinations. Sometimes memory may be impaired. But these well known consequences of intoxication have no obvious relationship to the question whether the defendant was conscious, or capable of forming intentions. I can imagine the state of someone who is so intoxicated as to be incapable of forming intentions. But I cannot set this imaginary figure into motion and make him do something. Consider the actions of the defendant in the High Court case of O'Connor. He rifled the glove box of a car belonging to an off-duty policeman. He took a knife. A witness summoned the policeman who attempted to question O'Connor. He ran away. He was caught and arrested and then stabbed the policeman. A struggle ensued during which he attempted to stab the policeman again. Or consider the actions of the defendant in the leading English case of Majewski. He engaged in a fierce struggle with the police, shouting, 'You pigs. I'll kill all you fucking pigs, you bastards'. Both O'Connor and Majewski were intoxicated at the time. But can it be suggested seriously that these defendants were unconscious, or that they were incapable of acting intentionally? There are, undoubtedly, cases of involuntary movement during unconsciousness. A stuporous drunk may roll out of bed, or perhaps overlay a child. One would not ascribe intentions in such cases. But these are far removed from the circumstances which are advanced in most cases involving the plea of automatism.

H. L. A. Hart describes 'unconscious action' by referring to 'outward movements [which] appear to be co-ordinated as they are in a normal action'. Professor John Smith writes of action which is only 'apparently purposive'. Again it is possible to understand these descriptions in their application to some simple cases. A drunk may stumble over a dog in a darkened hallway. To the dog's owner it may look like a deliberate kick. The owner may be mollified - he may accept the excuse when he is told that the drunk did not realize the dog was there and did not mean to kick it. The perhaps apocryphal story of the drunken Irish nurse who mistook a baby for a log of wood and put it on the fire is an extreme example of a predictable action set in an unusual context.

---

96 Ibid.
98 Hart H.L.A., op. cit. Acts of Will and Responsibility 90, 109. The position taken in the paper on the criteria for involuntariness is modified in the Notes. See ibid. 255-6. Involuntary, or 'fundamentally defective', actions are there described as 'those where the bodily movements occurred though the agent had no reason for moving his body in that way ... ' In cases of somnambulism, fugue states and the like, however, the amended criterion is scarcely adequate to Hart's purposes. For a psychoanalytic account will normally provide us with an account which makes D's behaviour intelligible in terms of his 'reasons' for acting even if he is unable or unwilling to avow them. The issue will then require the invocation of some test for determining whether they are truly 'his' reasons. And so we return to the metaphors of the divided self.
example of this kind. One can say in these cases that the drunken individual was not aware, or conscious of, certain circumstances or consequences of his action. His mental state could be described as one of impaired or altered consciousness. The existence of that state makes the denial of intention more credible. Here the ‘appearance’ of a purpose or intention to harm belies reality. But none of this involves a denial of capacity to form intentions, or an assertion that the individual was in a state of unconsciousness. O’Connor and Majewski made no attempt to account for their actions in terms of accident, blunder or mistake. All the evidence suggests that they were conscious of some (if not all) the circumstances in which they acted, and perfectly capable of acting intentionally.

There is, in these cases, a consistent tendency to confuse the ideas of ‘being conscious’ and ‘being conscious of’ something. The error may be traced back to Locke, who defined consciousness as a ‘reflex act of perception accompanying’ mental or physical action. The optical metaphor was natural enough, and it has had a long and distinguished list of adherents. Indeed Freud wrote of consciousness as a ‘sense organ for the perception of psychical qualities’. Locke also believed that this inner act of perception was an essential element of action. There is, it should be added, a no less distinguished list of philosophers, and psychologists, who have rejected that thesis.

If an individual did not realise, or if he was not aware or conscious of, the fact that his pistol was aimed at another human being he could not, of course, be said to have intended to kill him. This is not the point of Locke’s thesis, however. For Locke there could be no intention to kill unless a further condition was satisfied. The individual must also perceive that it is his intention to kill. It may be granted that such ‘acts of perception’ do occur. One can attend to one’s state of mind, reflect on one’s intentions, decide to abandon them or resolve anew to carry them out. Perhaps these are things one always could do, if one were so minded. But it is also true that intentional action can be undertaken without reflection and without any apparent act of inner perception.

If an inward act of perception is supposed to be an essential element of intentional action then it follows that the individual must be aware of it at the time. And, in Locke’s account, this awareness or consciousness constitutes personal

---

3 Ryle G., The Concept of Mind (1949) 159: ‘The metaphor of ‘light’ seemed peculiarly appropriate since Galilean science dealt so largely with the optically discovered world. ‘ Consciousness’ was imported to play in the mental world the part played by light in the mechanical world. In this metaphorical sense, the contents of the mental world were thought of as being self-luminous or refugent.
4 Sigmund Freud, The Interpretation of Dreams (Avon 1967) 654. See also Sigmund Freud, The Ego and the Id (1923). Charles Mercier provided an elaboration in Criminal Responsibility (1905) 49: ‘The capacity of the mind for holding ideas is limited. It cannot simultaneously contemplate many things. It can only bring one thing at a time into the full glare and illumination of consciousness. Neighbouring ideas, — associated ideas — if present at all, are remitted to a twilight, which becomes more obscure as the distance from the central focus of attention increases.
5 Gilbert Ryle, op. cit. provided the major modern statement of the sceptical case.
identity. If the individual denies the occurrence of this inner act it may be possible to conclude from his demeanour that he is lying. If he cannot recall it, however, one of two conclusions would follow in terms of Locke's analysis. One might say, contrary to all outward appearances, that there was never really an 'act' at all. It was only the delusive appearance of an act. That conclusion would follow from the way in which Locke defined the concept of human action. Or, in the alternative, Locke's thesis on personal identity would lead to the conclusion that there was an act, but an act which cannot be attributed to this 'person'. It may belong to some other 'person' who does remember it. Or it may exist in some limbo reserved for acts which cannot be attributed to the 'man' but not the 'person'. It is this idea that an inward act of perceiving one's state of mind is a necessary element of intentional action which may account for some recent decisions. An O'Connor, or a Majewski, pleads that his words and actions present only the delusive appearance of purpose or intentionality; some inward act of self perception was absent. The oscillation between the two possibilities suggested by Locke is the source of the ambiguity in the statement that one who pleads automatism denies that it was 'his act', which caused the harm. He may be denying that his movements are to be counted as an act. Or he may be denying that he is truly its author.

There are, no doubt, other contributing factors. However, an underlying reliance on a theory of action and personal identity derived from Locke goes some way towards explaining an otherwise baffling tendency of courts to accept the possibility that individuals can steal, bellow imprecations and engage in fierce struggles with policemen when they are not conscious.

Dr Jekyll and Jekyll MK I, both of whom remembered Hyde's actions, could hardly deny that they were conscious at the time. But what of Jekyll MK II who had no memory of those episodes? Mere failure of recall does not entail lack of consciousness at the relevant time. What is the additional elusive element which makes all the difference? If the preceding arguments against Locke and his modern descendants are accepted, it is a chimaera. It is, however, possible to sketch an outline of the kind of answer which might be derived from decided cases. The recognised examples of automatism, in the sense of unconscious action, seem all to involve natural processes, organic disorders or the use of drugs which are capable of producing states of complete unconsciousness. States, that is to say, in which even 'apparently purposive' action cannot occur. In the usual case where automatism is pleaded, however, the individual has manifestly not reached this state. He may have advanced some way along the continuum. The capacity for alert discrimination, or restraint, is characteristically diminished. He is far from being completely unconscious, however. Alcohol, for example, can produce a state of immobile stupor if it is consumed in sufficient quantities, yet ambulatory and sometimes very vocal drunks can plead, admittedly with variable success, that they were not conscious at the relevant time. There is a corresponding unwillingness on the part of courts to allow a plea of unconscious action in cases where the evidence merely indicates that the individual was subjected to great psychological stress which resulted in the commission of an offence which he cannot remember.6

Dr Jekyll never revealed the formula of the elixir and, in any event, its action was the result of an unidentified contaminant in one of the ingredients. The physical effects of the transformation were of the utmost severity: a grinding of the bones, deadly nausea and horror of the spirit. Even if it had also produced amnesia, it would not necessarily follow that Jekyll MK II was ‘unconscious’ in the elusive sense suggested by the decided cases. There is no indication that the elixir was capable of producing a state resembling sleep, coma, or drunken stupor. If, as seems likely, the defence of automatism has been extended to various disorders by finding an analogy with sleep something more than pain, stress and amnesia would be necessary before the analogy would be extended to Jekyll MK II.

One last reflection on amnesia and unconsciousness. Two doses of the elixir were necessary to effect the change from Jekyll to Hyde and back again. In the second of my hypothetical examples Jekyll MK II has no memory for what Hyde has done. Suppose it were the second dose which obliterated Jekyll’s memory of Hyde’s actions. We might then say that Jekyll MK II was nevertheless conscious of what Hyde was doing when Hyde was in the ascendant. He would resemble Raymond Cummins, who shot himself in the head after killing his girlfriend. Suppose, however, that the first dose operated prospectively, so that Jekyll MK II would not perceive what Hyde was doing. One might be more tempted to find an analogy with established categories of automatism if this were the case. But how could one tell which of the two versions of the drug’s effect was correct? And, to invoke Locke again, why should it influence our moral judgment of the case?

B. Voluntary Action

The line of Victorian cases to which I made earlier reference required proof that the defendant’s action was ‘conscious and voluntary’ before guilt could be attributed to him. That reference to an independent requirement of voluntariness provides a second possible line of defence. Moreover, the question whether the individual was capable of forming intentions at least implies that conditions other than lack of consciousness might produce such an incapacity. Dr Jekyll, or Jekyll MK I, for example, might claim that Jekyll could form no intentions when Hyde appeared.

There are evident signs that the equation of automatism with unconscious action is near the point of logical breakdown. In the New Zealand case of Burr,7 North P. said that automatism cannot require evidence of absolute unconsciousness, ‘because you cannot move a muscle without a direction given by the mind’. He suggested that the defence was founded on evidence that ‘all the deliberative functions of the mind’ are absent ‘so that the accused person acts automatically’.8 It is a vague and tautologous formula. It also serves to introduce the theme of voluntariness.

I shall exclude from consideration the idea of involuntary movements such as tics, fits, spasms, certain kinds of reflexes and the like. These simply do not count

8 Ibid. 745.
as intentional action. So, also, in those cases where an individual is pushed so that he cannons into another, or where his hand is seized and used as a weapon to strike another. It is the sense in which acts are described as voluntary, or not voluntary, with which I am concerned. Like consciousness, voluntariness is a matter of degree. One may be more or less free to act or refrain from action. Goads or impediments to action may be more or less pressing or prohibitive. In determining whether an act was voluntary one must have regard to a variety of rules and standards which vary with the context of application. If the question is whether a confession was made voluntarily, the answer will depend on the rules which govern permissible procedures of interrogation. If it is whether an assailant who was provoked acted voluntarily the rules and standards will be quite different. To say that an act was not done voluntarily is not to say that there was no intention to do it. The blackmailer’s victim, if he pays, does so intentionally. If he loses his self-control and attacks the blackmailer it is nonetheless an intentional act.9

These examples all involve external, or environmental, pressures on the individual. Where the pressure arises from internal sources — from the workings, let us say, of the unconscious mind — there is a tendency to hark back to theories of the will, and volitions. Here once more Locke provides a source. He argued that will or volition is a power of the mind and its exercise is described, like the exercise of consciousness, as an inward act. This inward act is again supposed to be an indispensable accompaniment to ‘acts’ in the external world. Modifications of this theory can be traced through the work of Dr Thomas Brown to Austin’s Lectures and thence to the jurisprudential literature.10 Stephen referred to volition, or the exertion of the will, as an ‘internal crisis’, of which ‘we are conscious but which cannot otherwise be expressed’. The inherent obscurity of the theory led H. L. A. Hart to dismiss it as ‘a misleadingly antiquated piece of philosophical psychology’.11 Again I am less concerned with finding an acceptable solution to this debate than in tracing its consequences. The theory of the will which I have sketched with such brevity, is the evident source of much of the modern case law on denials of voluntariness. The frequently cited passages from the judgment of Barwick C.J. in Ryan12 are illustrative:

It is basic, in my opinion, that the ‘act’ of the accused, of which one or more of the various elements of the crime of murder as defined must be predicated must be a ‘willed’, a voluntary act ... It is the act which must be willed, though its consequences may not be intended ... 13

[A]n accused is not guilty of a crime if the deed which would constitute it was not done in the exercise of his will to act. The lack of that exercise which precludes culpability is not, in my opinion, limited to occasions when the will is overborne by another, or by physical force, or the capacity to exercise it is withdrawn by some condition of the body or mind of the accused ... If voluntariness is not conceded and the material to be submitted to the jury wheresoever derived provides a substantial

9 Threats or duress may, however, affect the issues which arise when it is necessary to decide what it was that D intended. See, for example, Hurley and Murray [1967] V. R. 526; and the debate over Steane [1947] K. B. 997.
10 See the brief discussion in Hart H. L. A., op. cit. 97ff.
11 Ibid. 101.
12 (1967) 121 C. L. R. 205.
13 Ibid. 213.
14 Ibid. 214.
basis for doubting whether the deed in question was the voluntary or willed act of the accused, the jury's attention must be specifically drawn to the necessity of deciding beyond all reasonable doubt that the deed charged as a crime was the voluntary or willed act of the accused.15

If this internal accompaniment to action does not occur, the theory requires one to conclude that there is nothing more than the delusive appearance of an act. When the argument is put in this form, the way is opened to a more direct invocation of arguments based on a division of the self. Here, however, the source is not to be found in Locke, but in the psychoanalytic material.

*Ratahi*, a decision of the Victorian Court of Criminal Appeal,16 is an extreme case. Ratahi was convicted of rape, indecent assault and felonious assault. He appealed against his conviction on the ground that the trial judge had excluded certain psychiatric evidence as to his mental state.17 The Full Court quashed his convictions and ordered a new trial. It was said that the evidence provided a basis on which a jury might have been led to doubt whether Ratahi's actions were 'voluntary or willed'.18 Had the jury taken this view they might have acquitted entirely. The Court described Ratahi's behaviour as 'incongruous'. His victim was a stranger to him. He attacked her from behind as she was returning from a visit to her doctor. He undressed her and raped her. Then she was pulled to her feet and taken some distance before she was again raped. Then Ratahi kicked her twice in the head. After that he helped his victim to her feet, took her back to the scene of the first rape and helped her to dress. He told her he was sincerely sorry for what he had done. Ratahi remembered these events. He was able to give a coherent account at his trial which was substantially the same as the account given by his victim. There was no evidence of intoxication. But the defence sought to introduce evidence to the effect that the sexual attack was a secondary regression to a catastrophic rage reaction, while Ratahi was in the impaired state of consciousness known as depersonalisation and that sexual desire itself did not initiate the attack.19

Depersonalisation was described as,
a state during which, although he could perceive what was going on, it was as if the acts were being done by another and that he was an observer and in a sense unable to interfere or resist.20

The Full Court quoted at length from the judgment of Barwick C.J. in *Ryan* and concluded:

where a defence of automatism is raised the primary enquiry must be directed to the question whether the overt acts of the accused relied on by the Crown as constituting the offences were voluntary acts of the accused.21

It was, as I said at the outset, an extreme case. In other jurisdictions attempts to base a defence of automatism on evidence of dissociation, or depersonalisation,

15 Ibid. 216-7.
16 Unreported judgment of September 8th, 1976.
17 The trial court placed heavy reliance on *Joyce* [1970] S.A.S.R. 184, in reaching this decision.
18 The Court of Criminal Appeal also remarked that the evidence led on Ratahi's behalf was not sufficient to raise the issue of insanity. The Court suggested, however, that evidence of a disease of the mind might have been found had the psychiatric evidence been heard in full.
19 *Ratahi*, ibid. transcript p.4.
20 Ibid. 3.
21 Ibid. 10.
are often rejected by the courts.\textsuperscript{22} It is possible that the Victorian Supreme Court might adopt a more restrictive approach nowadays.\textsuperscript{23} 

Ratahi is, nevertheless, one of a number of Victorian cases in which the denial of ‘conscious and voluntary’ action has been accepted as a possible defence in circumstances where it would be quite impossible to contend that the defendant was ‘unconscious’ at the time of the offence.

The core of logic in Ratahi, and similar decisions, is to be found in the recognition that the automatism defence is not explained by invoking the idea of action performed in a state of unconsciousness. The cases are too diverse and the criterion of consciousness too uncertain. The concepts of involuntariness, lack of the ‘will to act’ or lack of volition seemed to provide a more adequate, and a more accommodating, rationale. There is a price to pay, however, for this piece of rationalisation.

Once consciousness is abandoned as a touchstone it is necessary to find some other criterion which will guide decisions on the issue of voluntariness. ‘Altered states of consciousness’, depersonalisation and dissociation can be produced in many ways. There is no obvious or necessary link with anything which could be termed a ‘disease of the mind’. Transient organic disorders, psychotropic drugs, and psychological stress can all produce such states. The state may or may not have been ‘voluntarily incurred’. That last expression, moreover, indicates the existence of yet another realm of dispute over the meaning of voluntariness.\textsuperscript{24}

At least three approaches to the problem of devising criteria for determining responsibility in these cases are possible. The first and most ruthless approach would be to abandon all reference to these mysterious mental entities and judge the accused according to his intentions and beliefs at the relevant time. This would involve no return to ‘objective’ tests of liability. The determination of intention requires the trier of fact to consider all observable manifestations of behaviour before, during and after the event. That necessarily includes the defendant’s own account of his intentions and beliefs, if he is able to give one. The effect would be to exclude, however, defences resting on a denial that an intentional act was preceded by a volition, or accompanied by the will to act.\textsuperscript{25} It is, perhaps, too


\textsuperscript{25} The following discussion takes no account of the insanity defence. The presence of some condition which can be described as a ‘disease of the mind’ allows a far wider range of exculpatory rationales to be considered by courts.
ruthless a solution. The classical problem of the somnambulist may be thought to demand a more sympathetic answer. Mrs Cogdon left her bed, fetched an axe from the woodheap, went to her daughter’s bedroom and killed her with two accurate, forceful blows. She told the Court that she had done these things in the course of a dream. In her dream she had seen North Korean soldiers all around the house and one of them on the bed, attacking her daughter.²⁶ It is difficult to avoid the conclusion that she acted intentionally. One can still ask, however, what her intentions were. In Mrs Cogdon’s case no-one doubted the sincerity of her account of her dream. If it is taken at face value one might conclude that she did not intend to harm her daughter but rather to kill the enemy soldier who attacked her. Her acquittal might be justified on the ground that her actions would have been necessary and praiseworthy had her dream been true. It is a possible explanation for the decision though hardly a convincing one. Bray C.J. treated it dismissively in Joyce.²⁷ To make the somnambulist’s guilt turn on the content of a dream seems arbitrary, even capricious. The somnambulist may not be able to recall any dream. If a dream is remembered, the symbolic events recounted by the dreamer may not fit any of the categories of legal exculpation.²⁸ Moreover, the question whether an account of a dream is true, as distinct from sincere, appears to be essentially unanswerable.

If one is unwilling to countenance an approach which might occasionally result in the conviction of somnambulists it is tempting to swing to the other extreme and to allow full play to the possibility of basing defences on evidence of dissociation or depersonalisation. This second approach is commonly expressed by asserting that voluntariness is a fundamental element of criminal responsibility. But a fundamental element of this nature cannot be limited to sympathetic cases like Cogdon. Cases of uncharacteristic or incongruous behaviour, such as Ratahi must also be left to the jury to determine on the issue of voluntariness. Intoxication by drink or drugs would also fall within the principle. It would only be necessary to provide some basis, either from expert evidence or from common experience, for the hypothesis that the accused’s behaviour was the product of unconscious forces. The attraction of this approach is that it avoids the need to confront the difficult issues of consciousness and intentionality. The proposition that Mrs Cogdon was unconscious when she killed her daughter can only be supported by dogmatic assertion.²⁹ The suggestion that she intended to kill an enemy soldier, rather than her daughter, rests on a refusal to consider the probable psychiatric explanation of her actions: in a post-Freudian age it is hardly convincing to suppose that Mrs Cogdon simply made an unfortunate mistake.

²⁶ Cogdon (1950) Unreported decision of the Supreme Court of Victoria; noted in Morris N., Somnambulistic Homicide: Ghosts, Spiders and North Koreans’ (1951) 5 Res Judicatae 29.
²⁸ Note, however, the orthodox view that defendants who plead self-defence, or provocation, founded on a mistaken apprehension must adduce evidence that the mistake was reasonable in the circumstances. Cf. the curious exception in Wardrope [1960] Criminal Law Review 770.
The 'voluntariness' approach is extreme in its attempt to ground legal defences on wide and ultimately indefinable grounds. It bears more than a passing resemblance to a process of judgment which Locke considered more appropriate to divine rather than human justice. There is a third, more conservative, approach. We might accept that the somnambulist has a defence simply because this category of exculpation is deeply embedded in history; all must sleep; and because fear of what sleep may release in the way of dreams or uncharacteristic action is a part of the common human condition. Mrs Cogdon's defence would not depend on the content of her dream if this view were to be accepted. The apparent conflict between her avowed and her 'unconscious' purpose would not affect the issue of guilt.

The defence would depend rather on the fact that somnambulism has a special relationship, of great antiquity, to issues of moral responsibility. The sleepwalker's defence cannot be extended to cases of explosive reaction under stress, intoxication, or organic disorder, on the simple ground that these states may be physiologically akin to sleep, or that they also allow unconscious impulses to be expressed in purposive conduct. In this approach the somnambulist's defence cannot be rationalised as an expression of some more fundamental requirement of voluntariness as a prerequisite for guilt.

This last approach to the issue involves an inevitable element of the arbitrary. The practice of exculpating the somnambulist can be extended by analogy to other sympathetic cases. So, for example, individuals who have been hit on the head and who thereafter engage in criminal acts during a period of 'post traumatic automatism' might be considered morally and legally indistinguishable from the sleepwalker. It is a vulgar error to suppose, however, that the grounds for drawing the analogy consist entirely of facts and suppositions drawn from the study of physiology and psychiatry. This process of reasoning by analogy requires a judicious balancing of similarities, dissimilarities and a consideration of issues of morality and policy. There is a difference between this process of reasoning by analogy, where recourse is had to the historical development of moral principles, and proceeding on the assumption that we can formulate some single principle or criterion for diverse cases.

When Locke asserted that the sleepwalker and the drunkard alike must suffer conviction, though each was morally innocent, he based his argument on the difficulties of proving 'consciousness' or the absence of that state. The dispute with Molyneux indicates that Locke might have been driven to concede that human courts should make the attempt to grapple with those difficulties in the case of the somnambulist. If the argument about the difficulties of proof is accepted, however, the concession would be made as an exceptional indulgence. The argument about proof is not without force. But there is a stronger reason for refusing to recognise that consciousness and voluntariness are fundamental and unqualified requirements for legal guilt. The defences based on them, particularly in their more florid modern form, are incongruities and anomalies in the existing structure of our criminal law.

The occasional tendency of courts to elevate the principle of voluntariness to a position of pre-eminence would allow Dr Jekyll and all his mutants to find the basis
for a legal defence. Expert evidence would be necessary of course. One would assume that it would be admissible.\textsuperscript{30} The issue would go to the jury. The likely result of their deliberations might be expected to depend less on the obscure logic of consciousness, or voluntariness, and more on the nature of the distinctions between the unfortunate Dr Jekyll and his neolithic forebear. The practical consequences of their decisions might also be expected to influence the verdict. I have distorted the moral meaning of Stevenson’s tale by concentrating on the issue of Jekyll’s guilt. In the novel Jekyll was transmuted into Hyde and in the end only Hyde remained. The hypothetical basis of my discussion — that Jekyll survives Hyde and Jekyll alone must be judged — is not only false to the story but may also be less true to life. A jury might well convict a Jekyll for the Hyde in him. Though the defences resting on a denial of conscious and voluntary action have the potential for considerable width of application, judges have defended them on the ground that juries may be expected to exercise common sense and record a conviction.\textsuperscript{31} There is at least the suggestion that juries are to make practical judgments whilst courts enunciate idealistic principles.

\textbf{VOLUNTARINESS: A FUNDAMENTAL PRINCIPLE?}

I think it is fair to say that the case law adds very little to our understanding of the \textit{moral} strength or weakness of Jekyll’s plea. It is possible to give the plea a legal form and rationale, in some jurisdictions at least. But the moral issue is remitted to the jury with ‘consciousness’ and ‘voluntariness’ as the sole criteria. Neither of these concepts adds anything to moral understanding of the issues.

The last of my questions had to do with the problems of policy and practicability which would arise from allowing Dr Jekyll a defence, or qualified defence, to a charge of murder. The issue is one of principle. Is the defence suggested by \textit{Ratahi} and similar cases congruent with other criteria for criminal responsibility? Even if Dr Jekyll has a moral claim, there may be good reasons of principle for distinguishing between moral accountability and legal accountability.

I suggested earlier that there is a distinction between ascriptions of intention and ascriptions of voluntary action. To say that someone acted intentionally is to make a statement about his behaviour, or state of mind, which is relatively free of moral evaluation. We may condemn him \textit{because} he acted intentionally. Moral evaluation is not a prerequisite for ascribing intention. If one asks whether his action was voluntary, however, the question requires a consideration of the pressures which induced, or compelled, him to act. Where those pressures have some obvious external source the answer will depend on a moral evaluation of his conduct in response to stress. It is essential to that evaluation of his conduct that one should be able to say, on occasion, that he could have, and should have, resisted those pressures. In law it is this element of moral evaluation by reference to criteria or

\textsuperscript{30} The effects of the elixir were not a matter of common experience or knowledge. See \textit{Darrington and McGauley} [1980] V.R. 353; \textit{Carr} (1982) 5 A.Crim.R. 466.

standards which accounts for the limits on the defences of provocation, self-defence, duress and necessity. There is a passage in the judgment of Windeyer J., in Parker, which expresses the idea:

What is insisted upon, if provocation is to avail as a defence, is that the action of the accused should be a normal reaction of an ordinary man. It may be that, on psychological analysis, the impulsive act of a sane man and an insane impulse are similar, in that in each case there is an act done, without deliberation or volition, in immediate reaction on the presentation of a situation. But law looks on them differently, whether or not it is psychologically proper to do so... to rely upon provocation it was necessary to shew that the provocative conduct of the deceased aroused in the prisoner an intent, in the legal sense, to do the act he did, not that it robbed him of the capacity to form an intent.

The ‘objective test’ in provocation, and objective tests in the other defences, are not designed for the purposes of determining whether the defendant lost his self-control, or was subject to ungovernable passion or temptation. The question is whether the loss of self-control occurred in circumstances which explain and excuse, or mitigate, his loss of control. And the tests entail the consequence that some defendants who could not control their conduct will be convicted. Moreover these defences are limited in their application. Provocation is only available in cases where the defendant sought to kill. Excessive self-defence is limited in the same way. Duress and necessity, by contrast, are general defences. However they have only limited application to murder.

The rules which govern these defences have their imperfections. They are the subject of continuing processes of reform and refinement. Consideration of those rules does indicate, however, that a claim of inability to exercise self-control, no matter how sincere, is not a sufficient basis for exculpation. The defendant can, of course, avoid the criteria and standards which govern the voluntariness defences by pleading insanity. Some utterly untenable pleas of provocation, for example, might be treated in this way. But the defence of insanity cannot accommodate all cases where loss of self-control was legally inexcusable.

If the principles which govern these defences are compared with the course of reasoning in the automatism cases there is a manifest incongruity. To allow Ratahi the possibility of a complete defence is to discriminate grossly against those who will be judged by the standards and criteria applied in cases where provocation, self-defence, excessive self-defence, duress and necessity are pleaded. In terms of voluntariness there is no ground for distinction. Real jealousy of a real rival is not less compelling than impulses which emanate from a disordered imagination. If the occurrence of ‘volitions’ is an essential element of voluntary action there is no reason to suppose that they are present when action is prompted by external stress and absent when the source of stress is less readily observable. So far as we know, Ratahi was not responsible for his state of mind. He may well have been a

32 (1963) 111 C.L.R. 610, 653.
33 See, for example, the Reports of the Victorian Law Reform Commissioner: Report no. 9 Duress, Necessity and Coercion (1980); Report no. 12 Provocation and Diminished Responsibility as Defences to Murder (1982).
35 See Halleck, op. cit. 210-1.
victim of misfortunes beyond his control which shaped and moulded him. But the same may be said of many individuals whose response to provocation is grossly excessive. It is even more incongruous to allow the possibility of a voluntariness defence where the defendant has chosen to take a mind altering drug.

It would be possible to argue that this incoherence in the structure of defences to murder merely indicates the need for reform and rationalisation of the rules which define that offence. It may be that the comparison of automatism cases with those involving provocation, or excessive self-defence illustrates the defects in those doctrines when measured against the voluntariness principle. Originally the definition of murder may have been more open textured. Blackstone thought that the principal distinction between murder and manslaughter was to be found in the reflection that, 'manslaughter arises from the sudden heat of the passions, murder from the wickedness of the hearts'.

Though the criminal jurisprudence of the nineteenth century, and in particular, attempts to codify the criminal law, tended to reduce the meaning of malice aforethought to an analytical formula, it is possible to trace a residual element which resisted articulation. There could not, in the nature of things, be any precise analysis of the wickedness of the heart, or of Foster's requirement of a 'wicked, depraved, malignant spirit'. There are occasional late appearances of this residual element of malice aforethought as, for example, in the judgment of Taylor J. in Howe. It can also be glimpsed in the dispute on the question whether the jury has a right, as distinct from a power, to return a verdict of manslaughter where the evidence points to murder or nothing.

Current proposals to 'subjectivise' the plea of provocation, to extend the defence of duress to murder, and to provide a defence of diminished responsibility can be seen as attempts to avoid the overly rigid formulations of the recent past. It seems plausible to suggest that some of the cases in which the 'conscious and voluntary' direction was given, such as Haywood, Tait, and Allwood can be explained in a similar fashion. It may be, in other words, that much of the problem lies in the definition of murder.

Not all of the cases can be explained in this way, however. The proposition that conscious and voluntary action is a necessary element of criminal responsibility is expressed quite generally. The incongruity of any general doctrine of this kind is apparent when one considers that provocation is no defence at all to any offence which does not require proof of an intention to kill. The jury will be formally directed to disregard provocation in these cases. So also where a plea of self-defence, duress or necessity, fails to match the objective tests which govern the availability of those defences.

The conclusion which can be drawn from the legal material is that Dr Jekyll's plea, if it is to be recognized at all, really goes to the inarticulate element in

36 Blackstone's Commentaries 190.
38 (1958) 100 C.L.R. 448.
40 See references in n. 32 and n. 33 supra.
murder. It is not a denial of intention and it does not match the standards and criteria which should govern the availability of the voluntariness defences. If Hyde stole, robbed, raped or set fire to buildings, Jekyll or any of his variants should be guilty of those offences.

That conclusion does less than full justice to the moral quality of Dr Jekyll’s plea. For there are, in a sense, standards and criteria on which he would seek to rely in his attempt to dissociate himself from Hyde. Ultimately he asks that the goodness of Jekyll should be taken into account in judging the degree of his responsibility for any unprecedented acts by Hyde. But that is, after all, the essence of the task which must be undertaken at the stage when Jekyll is sentenced for his offences. And it is this possibility of distinguishing between the stages of guilt determination and sentencing which accounts for many of the differences between moral and legal condemnation. It also provides a possible ground for distinguishing the offence of murder from other criminal offences in those jurisdictions where the penalty is mandatory.

I have argued that the metaphor of the divided self lies behind much of the case law concerned with the issue of ‘conscious and voluntary’ action. The tendency towards incoherent extensions of these categories of exculpation is apparent when defences such as self-defence, provocation, necessity and duress are compared with the defence of automatism and its variants. There is a far simpler and more compelling example of this incoherence. Consider the cases where the offender’s crime only comes to light years after the event. The offender may have suffered the torments of guilt. True they were not sufficient to induce him to confess and throw himself on the mercy of the state. But let us suppose that he has turned over a new leaf and led an exemplary life. He has buried the past, rather than merely repressing his memory of the past. It is a far more convincing example of altered identity than anything suggested in the cases on conscious and voluntary action. Yet he has no formal defences when the issue of guilt comes to be decided. Here the formalised rules of law preserve an inscrutable silence. It is true that the authorities may be unwilling to exact the full price for past misdeeds. The offender might not be charged at all, or he might be charged with a less serious offence than his crime warrants. The jury might return a merciful verdict. The sentencing judge will take into account the alteration of character during the intervening years. But these expedients were also possible in the cases where consciousness and voluntariness have been in dispute. There is no apparent justification for distinctions which bear more heavily on the ‘reformed offender’ than the individual who pleads automatism on the ground that his criminal conduct was a manifestation of his worse self.

There is another resolution to the debate between Locke and Molyneux which provides the epilogue to this paper. The issue was considered by the Full Court in South Australia in Joyce. Subsequent decisions may have impaired the authority of the discussion. I do not think they match the rigour of its analysis, or take
anything from the logic of its argument. The defendant had stabbed his girlfriend
to death. He sought to rely on the defence of automatism. Psychiatric evidence of
‘dissociation’ was advanced in support of the plea. The expert evidence drew on
the metaphor of the divided self. It was said that there were two David Joyces, one
of whom was conscious of what he was doing, one of whom was not. It was in
response to that evidence that the Court advanced the proposition quoted at the
outset:

In our view, if the personality is divided and one of the divided parts was conscious of the act and
wills it, the actor is responsible for it and the defence of automatism is not open.

The Court addressed itself to the problem of the drunkard and the sleepwalker in
obiter:

If it were otherwise we do not see why a man who kills while intoxicated through drink or drugs
should not be able to escape, not on the ground that he lacked the necessary intent to constitute the
crime of murder, which at the most would reduce his guilt to manslaughter, but on the ground that
the act was involuntary or not his act, because the higher part of his personality was put to sleep.
Such persons, like the present appellant, . . . may fail to remember what they have done after
returning to a state of sobriety, but that does not of itself exclude will or intention at the relevant
time.46

However the sleepwalker, the dreamer and those who suffer a blow to the head
before engaging in criminal conduct, were to be distinguished. Here a defence
might be advanced. There is a careful dissection of the reasons for this conclusion.
The Court emphasised the precedents for the plea in these cases. There is an echo
of Molyneux in the remark that, ‘Everyone must go to sleep and a man does not
choose to be knocked on the head’.47 Perhaps, the Court suggests, the dreamer
‘should be judged as if the dream were reality’. Or it might be that in such cases
‘whatever directs the acts is so subordinated to the conscious will that the act can
fairly be said to be involuntary’.48 There is, in these passages, a scrupulous
attention to the need to find some special justification for a defence in such cases.
They are not treated as exemplifications of some more general legal doctrine.

It may be as Joyce suggests, that the exceptional cases involving sleepwalkers,
and individuals who have been knocked on the head, are best described by
invoking the concept of voluntariness. It is a convenient enough label for cases in
which the individual’s freedom of choice is impaired. But the extent to which an
impairment of that freedom will affect the determination of legal responsibility
must depend on articulate standards and criteria. It is an issue of legal policy which
cannot be resolved by dwelling on the mysteries of ‘will’, ‘volition’ or ‘conscious-
ness’. Professor Brett, in his analysis of the issues, proposed a different resolution
of the problem of the divided self from that suggested in Joyce. But he also warned
that, ‘our first step in discussing this problem must be to put aside any temptation
to resolve it by talking of involuntary activity. The phrase is so vague and
imprecise as to cause confusion rather than enlightenment’.49 The Full Court in
Joyce made the same point in a slightly different fashion when it discussed the
example of somnambulistic homicide:

46 Ibid. 193.
47 loc.cit.
48 loc. cit.
49 Brett P., op. cit. 196.
we find it hard to see how an act can be committed with a purpose and still be committed unconsciously. Can such acts still be involuntary and unconscious in the eyes of the law? We share, with respect, the difficulties expressed so energetically by Windeyer J. in Ryan v. R. 50

*Joyce* was cited to the Victorian Court of Criminal Appeal in *Ratahi*. The Court declined to follow the South Australian case. It was the judgment of Barwick C.J. in *Ryan v. R.* which formed the basis for the Court’s decision in *Ratahi*.