

Self-Control in the Modern Provocation Defence[†]

RICHARD HOLTON* AND STEPHEN SHUTE**

Abstract—Most recent discussion of the provocation defence has focused on the objective test, and little attention has been paid to the subjective test. However, the subjective test provides a substantial constraint: the killing must result from a provocation that undermines the defendant's self-control. The idea of loss of self-control has been developed in both the philosophical and psychological literatures. Understanding the subjective test in the light of the conception developed there makes for a far more coherent interpretation of the provocation defence. It also makes clear just how radical various proposals for reform of the defence would be.

The basic structure of the provocation defence, by which a murder charge can be reduced to manslaughter, is straightforward. As all criminal lawyers know, it consists of a union of two tests: a 'subjective' test, which asks whether the defendant was provoked to lose his self-control and whether he committed the killing as a result of that loss; and an 'objective' test, which asks whether the provocation was sufficient to make a reasonable person do as the defendant did. Recent judgments, and much recent academic discussion, have been preoccupied with the objective test. Yet this focus of interest has not led to consensus. Indeed, it is fair to say that understanding of the objective test is in some disarray.

We contend that this stems largely from asking the objective test to do too much work: work that should have been done by the subjective test. The subjective test is far more complex, and far more demanding, than has been generally realized. It requires not simply that agents must kill in response to provocation, but that they must kill *as a result of losing self-control* in response

[†] The central ideas in this article have been developed from a paper delivered by the authors at a conference on *Human Nature in Law and Political Morality* held at the University of Cambridge in 2002, and at the annual conference of the Australian Society of Legal Philosophy, held at the University of Sydney in 2003. The authors would like to thank the audiences on those occasions, and John Gardner, Timothy Macklem, Roger Shiner and an anonymous referee, for comments that greatly improved the article; and the Arts and Humanities Research Council for granting them each research leave that allowed work on it.

* Professor of Philosophy, Department of Linguistics and Philosophy, Massachusetts Institute of Technology.
E-mail: holton@mit.edu.

** Professor of Criminal Law and Criminal Justice, School of Law, University of Birmingham.
E-mail: S.C.Shute@bham.ac.uk.

to provocation. We aim to take seriously the idea of loss of self-control.¹ Once we do so, we find that the subjective test provides a substantial hurdle, for the simple reason that only an agent who antecedently possessed self-control can lose it as a result of provocation. To borrow Muddy Water's words, you can't lose what you ain't never had; nor, we might add, can you lose what you have already lost.

Irascibility and drunkenness plausibly fall into the class of characteristics that should result in the defendant failing the subjective test. Irascible agents have never gained self-control,² drunks have already lost it. There is, therefore, no need to insist that the objective test must be formulated in such a way as to exclude them. To say this is not to suggest, of course, that the law should dispense with the objective test altogether. It retains a crucial role. It is just to say that much of the work that it has been called upon to do can be better done by the subjective test. Furthermore, getting clear on the subjective test will bring a better understanding of the objective test, since it too concerns the loss of self-control. We return towards the end of this article to the question of how work should be apportioned between the two tests; we start by examining the subjective test.

1. *The Subjective Test Within the Classic Definition*

In English law, provocation remains a common law defence.³ If we are to understand its structure, therefore, we must look to case law for assistance. The first port of call is the so-called 'classic definition'⁴ of provocation in *R v Duffy*. There Devlin J (as he then was) described provocation as:

some act, or series of acts, done [or words spoken] ... which would cause in any reasonable person, and actually causes in the accused, a sudden and temporary loss of self-control, rendering the accused so subject to passion as to make him or her for the moment not master of his mind.⁵

¹ In contrast, the Law Commission has recently recommended that the subjective element be removed completely from the defence: see *Partial Defences to Murder*, Law Com No. 290, Cmnd 6301 (2004); and *Murder, Manslaughter and Infanticide*, Law Com No. 304, HC 30 (2006). As Horder has rightly observed, this change 'detaches the plea of provocation from reliance on one of the key justifications for its legal recognition'; see J. Horder, 'Reshaping the Subjective Element in the Provocation Defence' 25 (2005) *OJLS* 123, at 133. However, Horder is sceptical about how much weight the subjective criterion can put on the idea of loss of self-control, claiming that it 'might be thought to be more metaphorical than psychologically descriptive' (126). Notwithstanding disagreements here, we have considerable sympathy with many of Horder's broader contentions, and return to discuss them at the end of this article.

² Or at least, that is true of one sort of irascible agent; we discuss below the possibility of agents who have self-control that is habitually undermined.

³ See *R v Ahluwalia* [1992] 4 All ER 889, CA, per Lord Taylor CJ at 894, and *Attorney General for Jersey v Holley* [2005] 3 WLR 29, PC, per Lord Nicholls of Birkenhead, at 32.

⁴ Per Lord Taylor CJ in *R v Ahluwalia*, above n 3 at 894.

⁵ [1949] 1 All ER 932. In the Court of Appeal, Lord Goddard CJ (see 933) described Devlin J's summing up as 'as good a definition of the doctrine of provocation as it has ever been my lot to read, and I think it might well stand as a classic definition ...' In *R v Ahluwalia*, above n 3, at 894, Lord Taylor CJ noted that the bracketed words had to be added and the seven missing words removed to allow for the changes made by the Homicide Act 1957.

Devlin J's statement places loss of self-control firmly at the heart of the provocation defence, a view given legislative force by section 3 of the Homicide Act 1957.⁶ It also expands upon the notion of loss of self-control by adding two adjectives—'sudden' and 'temporary'—to the definition. This raises three questions that will be at the heart of our account: (i) what is self-control? (ii) what do we mean when we speak of someone *losing* their self-control, and what characteristics do they display? and (iii) what difference, if any, is made to the scope of the provocation defence by the addition to the definition of the words 'sudden' and 'temporary'?

2. *What is Self-Control?*

In the classic definition given earlier, Devlin J says that to lose one's self-control is to be no longer master of one's mind. If that is right, then self-control itself must consist in mastery of one's mind.⁷ But in what does this mastery consist? There are three main accounts to be found in the philosophical and psychological literatures:

- (i) Self-control consists in the ability to bring one's actions into line with one's second-order desires: i.e. it is the ability to ensure that the desires upon which one acts are the desires upon which one desires to act;
- (ii) Self-control consists in the ability to bring one's actions into line with one's judgments about what it would be best to do;
- (iii) Self-control consists in the ability to bring one's actions into line with one's resolutions and commitments.

These are clearly competing accounts (though they will classify many cases the same way⁸), so we need to adjudicate between them. But, before doing so, let us focus on two things that they have in common. First, note that all three accounts agree that self-control consists in a *particular kind* of control over one's actions—in each case the obvious contrast is with actions that are driven purely by one's (first-order) desires. Accordingly, loss of self-control involves a loss of this particular kind of control. It does not require that the agent 'goes berserk', loses control of her body, or fails to know what she is doing.⁹ The agent will still be an agent when self-control is lost, and her acts will still

⁶ Section 3 states: 'Where on a charge of murder there is evidence on which the jury can find that the person charged was provoked (whether by things done or by things said or by both together) to lose his self-control, the question whether the provocation was enough to make a reasonable man do as he did shall be left to be determined by the jury; and in determining that question the jury shall take into account everything both done and said according to the effect which, in their opinion, it would have had on a reasonable man'.

⁷ Compare Tindal CJ's description of the provocation defence in *R v Hayward* (1833) 6 C & P 157, at 159: '... while smarting under a provocation so recent and so strong, that the prisoner might not be considered at the moment the master of his own understanding...'

⁸ Consider, for example, a dieter who wishes that he didn't want to eat more, who judges that it would be better if he ate less, and who has resolved to do so. When he exercises self-control by refusing some tempting food, he meets the conditions set by all three accounts of self-control.

⁹ See, for instance, J. Horder, *Provocation and Responsibility* (Oxford: Clarendon Press, 1992) at 102.

be intentional, driven by a desire for revenge, or whatever. What is lost when one loses self-control is control over *which* mental elements drive one's actions; and that, to return to the formulation of Devlin J, is to lose a certain kind of mastery over one's mind.¹⁰

Second, note that whilst our primary focus is on provocation, the notion of self-control has a much wider application. The three accounts that we are examining were initially proposed to characterize self-control in the face of *temptation*: resisting a cigarette for instance, or sticking to a diet (some of this literature is discussed below). At a push, we might try to take temptation as the essential factor and assimilate provocation to it—the defendant succumbs to the temptation to kill. But the move is most unnatural. Moreover, there are other cases that are different again. Self-control can be needed, for instance, to overcome the fear that would otherwise make one flee situations of danger; or in cases where elation risks leading one to disregard prudence. These are not naturally described as cases of either provocation or temptation. Nevertheless, there is clearly something that these diverse cases have in common: they all involve an attempt to overcome certain inclinations ('passions') that would otherwise move us. We might conjecture that there is a similar mechanism at work in each. As we shall see, this conjecture is borne out by the empirical evidence.

Bearing in mind these points in which the three accounts agree, let us now turn to the ways in which they differ. The first account, which comes from Harry Frankfurt, has been highly influential in philosophy.¹¹ It has also, however, been widely criticized; indeed, in later work, Frankfurt himself has stepped back from it.¹² We too think that it cannot be right. Why should second-order desires (desires about one's first-order desires) be privileged? For a start, might not there be desires of a still higher order: third-order desires, say, or fourth? Should not these count as still more privileged?¹³ A natural response to this worry is to say that self-controlled agents must get their desires and actions into line with their *highest-order* desires (so long as

¹⁰ Contrast the comments of the trial judge, Peter Pain J., in *R v Richens* [1993] 4 All ER 877, at 882, who characterised the loss of self-control as: "a complete loss of control, to the extent where you really do not know what you are doing. In one leading case it has been described as the defendant "no longer being master of his mind". It might be slightly better put that his mind is no longer master of his body, because he has so far lost control, he really does not know what he is doing. In ordinary language, we talk of "seeing red" or "going berserk". These comments were held by the Court of Appeal (see 884) to constitute a material misdirection.

¹¹ H.G. Frankfurt, 'Freedom of the Will and the Concept of a Person' in *The Importance of What We Care About* (Cambridge: Cambridge University Press, 1988) 11–25.

¹² In his later work, Frankfurt replaces the idea of higher-order desires with that of those desires with which one *identifies*, whether or not these are higher-order. See 'Identification and Wholeheartedness' in *The Importance of What We Care About*, above n 11 at 159–176; and 'The Faintest Passion' in *Necessity, Volition and Love* (Cambridge: Cambridge University Press, 1999) 95–107. An account of self-control based on conformity to these desires is much closer to the second account that we discuss here.

¹³ G.L. Watson, 'Free Agency' 72 (1975) *Journal of Philosophy* 205. For a related argument about the explanatory power of second-order beliefs, see S. Shute, 'Knowledge and Belief in the Criminal Law' in S. Shute and A. Simester (eds), *Criminal Law Theory: Doctrines of the General Part* (Oxford: Oxford University Press, 2002) 171–206.

these are at least second-order). But this response just makes a second and deeper worry all the more pressing: higher-order desires have no special authority just in virtue of being higher order. They can be as whimsical or foolish as any other desires; and the agent can be fully aware that this is so. Consider someone who does not like strawberries, but wishes that she did: we wouldn't think that self-control requires her to bring her first-order desires and actions into line with her second-order desires.¹⁴ Or consider the style-conscious teenager who does not want to smoke, but wishes that he was the kind of person who did: how cool he would then be. Isn't it more plausible to think that self-control consists in resisting this romantic higher-order desire rather than in conforming to it? This is a thought, moreover, with which the teenager might well agree.

What, then, of the second account, the idea of self-control as the ability to bring one's actions into line with one's judgment of the best? This approach has a much longer pedigree, dating at least to Aristotle. For Aristotle, the continent person (the *enkrates*) feels the pull of temptation or anger or whatever, but resists acting on it as a result of her judgment that such action would not be best; the incontinent person (the *akrates*) does not resist.¹⁵ The account certainly has intuitive appeal. However, some clarifications and modifications are in order. First, we need to distinguish between, on the one hand, those agents who do what they judge best because they have no desire to do otherwise, and, on the other, those for whom such action requires a struggle against their desires. In both cases the agents act in ways that conform to their judgments about what they ought to do, yet only the latter employ self-control as we ordinarily conceive of it: the former have no need of it.¹⁶ Second, we need to be careful about which of the agent's judgments are relevant for the purposes of this account. It is a feature of anger, temptation and the like that they corrupt judgment.¹⁷ In the heat of anger we may judge that our opponents deserve what we are doing to them, a judgment that we would not have made in a calmer moment. In this sense, our actions and our judgments about what is right may be in harmony in the heat of passion. Yet this does not mean that we have exercised self-control when we self-righteously react in ways that we would, in calmer moments, judge excessive. Indeed, quite the

¹⁴ J. Kennett 'Decision Theory and Weakness of Will' 72 (1991) *Pacific Philosophical Quarterly* 113, at 126–7.

¹⁵ The main discussion is in *Nicomachean Ethics*, Book 7. Note that the judgment on which the *akrates* fails to act need not be a moral judgment; it is rather the broad practical judgment that a certain action is best all things considered. See D. Davidson, 'How is Weakness of the Will Possible?' in his *Essays on Actions and Events* (Oxford: Clarendon Press, 1980) 21–43.

¹⁶ Aristotle recognized this in his distinction between actors who exhibit continence (*enkrateia*), and those who exhibit temperance or sound-mindedness (*sophrosune*). See *Nicomachean Ethics* at 1152a1ff. Aristotle regarded the latter—persons whose passions and reason work in harmony—as the more virtuous.

¹⁷ This corruption goes very deep. Even focusing on the benefits to be gained from resisting temptation tends to make agents more likely to succumb, since, under the influence of the temptation, the rewards are judged less valuable. See R. Karniol and D. Miller, 'Why not wait? A cognitive model of self-imposed delay termination' 45 (1983) *Journal of Personality and Social Psychology* 935 and W. Mischel, 'From Good Intentions to Willpower' in P. Gollwitzer and J. Bargh (eds), *The Psychology of Action* (New York: The Guildford Press, 1996) 197–218.

reverse: we might think that self-control requires us to keep our judgments free from corruption of this form, or, if we cannot do that, at least to develop strategies that enable us to avoid acting on them when they are corrupt.¹⁸ It seems, then, that the relevant judgments are those that are not corrupted by anger or temptation or the like. We might try to identify them in different ways: as those made before the anger arose, or after it subsided, or that would have been made had the anger not arisen. We should take care, though, not to idealize too far. Self-control is still control relative to some of the agent's *actual* judgments of what is right; it is not relative to what the agent would judge if she were truly virtuous or knowledgeable.

Summarizing these judgments as the agent's *considered* judgments, we can now give our more careful formulation of the second account:

(ii*) Self-control consists in the ability to bring one's actions into line with one's considered judgments about what it would be best to do, where these judgments depart from one's desires.

This is very plausible. But if it is correct, it might seem that the third account—self-control as the ability to abide by one's resolutions—must be wrong.¹⁹ For the two are surely different, and where they differ it seems that the second is better than the third: if we make a considered judgment that an action is right, we can show self-control by acting on it, even in the absence of a prior resolution; and if we resolve to do what we judge to be wrong, we do not exhibit self-control by sticking to the resolution. It might seem, therefore, that the third account is better understood as an account of *strength of will* rather than as an account of self-control.²⁰

However, to leave things at this is to miss out on the close relation between the two notions. For, whilst they are conceptually distinct, it is plausible that

¹⁸ This assumes that, even when we have lost our self-control, we can still make judgments, even if they are distorted. In contrast, according to Horder, legal theorists of the eighteenth and nineteenth centuries thought that the effect of the passions (including anger) was to produce a state akin to madness in which we act without judgment—an account which, according to Horder, had its origins in Hobbes. See his *Responsibility and Provocation*, above n 9, esp. Ch. 5. Much recent work on the cognitive basis of anger suggests that our assumption is more accurate than Hobbes': anger is grounded in judgments about the harm inflicted and the appropriate response, even if these judgments are so fleeting that agents are often not aware of them. See, for instance, A Beck, *Prisoners of Hate: The Cognitive Bias of Anger, Hostility, and Violence* (New York: HarperCollins, 1999) 31ff.

¹⁹ We know of nowhere where this account is clearly given, but something along these lines is implicit in much psychological work. See, for instance, the discussion of self-control as the ability to stick to long run preferences that is involved in discussions of hyperbolic discount curves: G. Ainslie, *Picoeconomics* (Cambridge: Cambridge University Press, 1992); and H. Rachlin, *The Science of Self-Control* (Cambridge MA: Harvard University Press, 2000). In a slightly different tone, see Walter Mischel's work on the self-control involved in the ability to delay gratification, work that is summarized in 'From Good Intentions to Willpower', cited above (n 17) and discussed below. For philosophical discussions of the importance of resolutions, see E. McClennen, *Rationality and Dynamic Choice* (Cambridge: Cambridge University Press, 1990); M. Bratman, *Faces of Intention* (Cambridge: Cambridge University Press, 1999); and R. Holton, 'Intention and Weakness of Will' (1999) 96 *Journal of Philosophy* 241 and 'How is Strength of Will Possible?' in S. Stroud and C. Tappolet (eds), *Weakness of Will and Practical Irrationality* (Oxford: Clarendon Press, 2003) 39–67.

²⁰ For an account of strength of will along these lines, see R. Holton, 'Intention and Weakness of Will', above n 19. There it is argued that it is a mistake in contemporary philosophy to identify weakness of will with *akrasia*, and hence to see it as the contrary of self-control.

strength of will is the standard *mechanism* by which self-control is achieved: we do the thing that we believe to be right, rather than what we currently want to do, *because of* a prior resolution. Certainly, empirical studies support the idea that we are much more effective in overcoming temptation if we have formed a prior commitment, especially if this includes specific intentions for how that commitment is to be implemented.²¹ But the connection plausibly goes deeper than that. We have explained that judgment may be corrupted by anger or the like. Nevertheless, the self-controlled agent will be able to avoid acting on these corrupted judgments and will instead be able to bring her actions into line with her *considered* judgments. What access will an agent have to these judgments? She might try to imagine what she would think if she were calm or what advice she would give to a third party who found herself in the same situation. But, in the heat of the moment, the results of such exercises are likely to be corrupted too—even in the unlikely event that she would be able to perform them. Instead, all that agents are likely to have are prior resolutions about what they should and should not do. Of course, we do not mean that self-controlled agents will have a set of specific resolutions capable of dealing with any specific circumstance. Rather, we mean that they are likely to have a set of more general commitments: not to rise to a bait, not to resort to violence except in self-defence, and so on.²²

The suggestion, then, is that there is a distinct faculty of self-control that enables agents to do what they judge best in the face of strong inclinations to the contrary; and that this faculty standardly works by employing prior resolutions. Agents who lacked self-control could still perform intentional actions: it is just that their actions would be driven by their immediate inclinations.

In suggesting this, we have not gone far beyond common-sense ('folk') psychology: that is, we have not gone far beyond the theory of mind that is embedded in ordinary everyday practices, and in the law. However, when we turn to more systematic psychology, we also find support for this view. Furthermore, the empirical literature contains suggestive findings on how the faculty works: it appears that self-control operates in many ways like a muscle. None of these arguments is conclusive; but, together with the common-sense considerations, they give rise to a very persuasive case.

²¹ For references to studies on the general effectiveness of intentions (or 'goals', as they put it, which seem to amount to the same thing), see R. Baumeister, T. Heatherton and D. Tice, *Losing Control* (San Diego: Academic Press, 1994) 62–3. For the added efficacy of implementation intentions, see P. Gollwitzer, 'The Volitional Benefits of Planning' in P. Gollwitzer and J. Bargh (eds), *The Psychology of Action* (New York: The Guildford Press, 1996) 287–312.

²² Arguably, there is an even closer connection between strength of will and self-control, which is that only a prior commitment would have the necessary motivating force to overcome contrary desires. The issue turns on the doctrine of internalism: on whether a moral judgment can by itself be capable of moving us to action, or whether it needs something else, such as a desire or an intention. We cannot pursue the issue here.

A. *Developmental Evidence*

We do not know of any empirical studies on the development of the ability to exercise self-control in the face of provocation; but there is extensive work on the development of self-control in the face of temptation. Walter Mischel and his colleagues have tested children on their ability to delay gratification to achieve greater reward.²³ For instance, they are told that they will receive one cookie if they ring a bell, which they are free to do at any time; but that they will get two if they refrain from ringing the bell until an adult comes in. The researchers found that ability to wait comes in around the age of four or five and by the age of six almost all children have it, though to markedly different degrees. Strong self-control is a very good predictor of later success in a wide range of academic and social skills.

B. *Effects of Depression, Fatigue etc.*

The ability to abide by a resolution is systematically affected by a range of features. Reformed alcoholics are far more likely to relapse if they are depressed, or anxious, or tired.²⁴ Moreover, these states affect one's ability to abide by *all* of one's resolutions: resolutions not to drink, not to smoke, to eat well, to exercise, to work hard, not to watch daytime television, or whatever. It is, of course, possible to explain this by saying that these states (depression, anxiety, fatigue etc.) systematically strengthen all of one's desires to do these things (or weaken all of one's resolutions not to do so); but it is surely a more economical explanation to say that they affect one's ability to act in line with one's resolutions. For why else would there be such systematic effects? Moreover, whilst stress makes dieters want to eat more, it tends to have the opposite effect on those who are not on a diet.²⁵ Once again, then, the empirical evidence seems to point to a separate faculty of self-control.

C. *Ego Depletion*

Finally, let us mention the remarkable empirical literature on what is known as 'ego depletion'. A set of experiments has suggested that self-control comes in limited amounts that can be used up: controlling oneself to eat radishes rather than the available chocolate cookies in one experiment makes one less likely to control oneself to persist in solving unsolvable puzzles in the next;²⁶

²³ For a summary, see W. Mischel, 'From Good Intentions to Willpower', above n 17. Much of Mischel's work focuses on the strategies used to obtain self-control.

²⁴ See R. Baumeister, T. Heatherton and D. Tice, *Losing Control*, above n 21 at 151ff. The same is true of those who are dieting (ibid. at 184ff), or trying to give up smoking (ibid. at 212ff) or taking drugs (M. Muraven and R. Baumeister, 'Self-Regulation and Depletion of Limited Resources: Does Self-Control Resemble a Muscle?' 126 (2000) *Psychological Bulletin* 247, at 250).

²⁵ See M. Muraven and R. Baumeister, 'Self-Regulation and Depletion of Limited Resources', above n 24, at 251.

²⁶ See R. Baumeister, E. Bratslavsky, M. Muraven and D. Tice, 'Ego-Depletion: Is the Active Self a Limited Resource?' 74 (1998) *Journal of Personality and Social Psychology* 1252.

suppressing one's emotional responses to a film makes one less likely to persist in squeezing a handgrip exerciser.²⁷ Again, it is possible to think that what happens here is that people's desires are affected: that suppressing one's emotional responses reduces one's desire to persist with squeezing the handgrip exerciser (or increases one's desire to desist). But why should there be effects on such disparate desires? And why do some activities (those that require self-control) bring about these effects, whilst others (such as doing mathematical problems) do not?²⁸ A much better explanation is that one's action is determined not simply by the strength of one's desires and one's resolutions, but also by a separate faculty of self-control; and that, like a muscle, this faculty gets tired when it is repeatedly exercised. Moreover, there is some evidence that the faculty is like a muscle in another respect. Whilst exercise decreases its effectiveness in the short-term, in the long run it actually increases it.²⁹

As we have said, none of these considerations is compelling on its own. But together they do make a persuasive case for the existence of a real faculty of self-control, one which works by keeping a person to her considered judgments in the face of changing desires.

3. *Loss of Self-Control*

We are now in a position to broach the second of our questions: to see what it is to *lose* self-control, and to see what results such a loss would engender. If the foregoing account is broadly correct, self-control is a specific ability working through a specific mechanism. An agent could, in theory, live without it; indeed it seems that animals and very young children generally do. Agents who lacked self-control would still perform intentional actions—they would still be agents—but these actions would be driven by their current desires or other inclinations. What they would lose is the ability to control those inclinations: to bring their actions into line with what they judged best.³⁰

Few, if any, adults live entirely without self-control. However, some have more than others; and the self-control that they have can be lost. We have seen already that states like stress and anxiety can destroy self-control. Importantly for our concerns, so does anger. We do not simply mean that anger will sometimes *overwhelm* our self-control, in the sense that we have insufficient self-control to deal with it. We mean that it can actively *undermine* self-control: the effect of anger can be to destroy the self-control that we would

²⁷ See M. Muraven, D. Tice and R. Baumeister, 'Self-control as a Limited Resource: Regulatory Depletion Patterns' 74 (1998) *Journal of Personality and Social Psychology* 774.

²⁸ *Ibid.* at 781–2.

²⁹ See M. Muraven, R. Baumeister and D. Tice, 'Longitudinal Improvement of Self-Regulation Through Practice: Building Self-Control Strength Through Repeated Exercise' 139 (1999) *Journal of Social Psychology* 446.

³⁰ In speaking of current desires we do not exclude the possibility that these might be desires for the future; the point is that they are desires which are currently held, rather than desires which the agent knows he once held, or will come to hold.

otherwise have. Anger thus undermines the very thing that is supposed to control it.³¹

This gives rise to an important result. If an agent suffers a provocation and responds with some violent act, there can be two quite different explanations of why she acts as she does. The first is that she lacked sufficient self-control to handle the provocation: it simply gave rise to violent inclinations that her self-control was unable to restrain, inclinations that therefore moved her to perform the violent act. The second explanation, in contrast, accepts that the agent did have sufficient self-control to handle the violent inclinations. However, the provocation acted in two ways: as well as giving rise to the inclinations, it also undermined the self-control that would otherwise have restrained them. Without the latter, undermining effect, the agent's self-control would have been sufficient to prevent the violent response; with it, it was not. Our view is that it is only in cases in which this undermining has occurred that we can properly speak of an agent *losing* her self-control as a result of the provocation. So, it is only in these cases that the modern provocation defence can properly be invoked.

Whilst the distinction here, between cases that involve undermining of self-control, and those that do not, is conceptually clear, it does present an evidential problem. How are we to know whether a given violent act falls under one explanation or the other? The problem stems from the fact that the very counterfactual tests we normally employ to determine causal chains have no obvious purchase here. We want to know whether the *loss* of self-control was necessary for the violent response. That is, we want to know whether the violent response would have happened even without the agent's self-control having been undermined.³² But we cannot ask: would the agent have refrained from violence if she had been provoked in this way but had not received the stimulus that caused her to lose her self-control? For, *ex hypothesi*, the provocation *is* the stimulus that caused her to lose her self-control.

So how are we to answer this question? If we knew enough neuro-science we might think of looking inside the agent's brain to see what had happened: were there the kind of brain changes that suggest that self-control was lost?³³ But we

³¹ Self-control is thus an instance of what is known in philosophy as a 'finkish' disposition: a disposition that can be disabled by the very stimulus to which it is supposed to react. (The term was introduced by C. B. Martin; it was finally published in his 'Dispositions and Conditionals' 44 (1994) *The Philosophical Quarterly* 1, but was in widespread use long before that. For helpful discussion, see S. Mumford, *Dispositions* (Oxford: Clarendon Press 1998) 82–92.

³² See *R v Thornton* [1992] 1 All ER 306, CA, *per Beldam LJ*, at 314. 'In every... case the question for the jury is whether at the moment the fatal blow was struck the accused had been deprived for that moment of the self-control which previously he or she had been able to exercise.'

³³ Cases like that of Phineas Gage, who apparently lost his self-control after suffering damage to his pre-frontal cortex, suggest that this is the region of the brain involved. Research here might well bring further evidence for the existence of a separate faculty of self-control; but it is not yet far enough advanced. For an accessible introduction to some of the evidence, together with some provocative conclusions, see A. Damasio, *Descartes' Error* (New York: Putnam, 1994). For some recent evidence of different pathways involved in immediate and delayed gratification, see S. McClure et al., 'Separate Neural Systems Value Immediate and Delayed Monetary Rewards' 304 (2004) *Science* 503. We draw very different conclusions from Damasio's work, and from other

cannot do that. So we have to use less direct methods for determining the right explanation. To say that these methods are indirect is not to say that they are generally unreliable. Nor are they in any way esoteric. They will turn out to involve just the kind of factors that judges and juries are likely to consider in provocation cases. We take this as evidence that, despite its apparent complexity, the account of self-control that we have offered is in fact both the common-sense one and the legal one.³⁴

4. *The Subjective Test Revisited*

With this account of self-control in place, we now return to the subjective test as it occurs in the provocation defence, and to the question of the evidence that can bear on it. The test breaks down into three sub-tests: (i) did the agent have self-control prior to the provocation? (ii) if so, was that self-control lost as a result of the provocation? and (iii) if so, was it this loss of self-control that was the cause of the homicide? We examine each in turn.

(i) Did the agent have self-control prior to the provocation? It is plausible, as we have already suggested, that everyone has *some* self-control. So the question becomes: did the agent have *sufficient* self-control? Sufficient, that is, to resist any inclination to kill, had the agent's self-control not been undermined by the provocation. One way of investigating this is to ask whether the agent generally has self-control with respect to the kind of provocation that he received. If not, then there are good grounds for thinking that the provocation did not cause a *loss* of self-control, since there are good grounds for thinking that sufficient self-control was lacking in the first place. The obvious example is the irascible agent who is unable to control his anger in response to the smallest set-back; the self-control he possesses is easily overwhelmed.³⁵

The modern law of provocation is consistent with the idea that, if the defendant generally lacks the relevant ability to control himself (i.e. his self-control is so weak that it can be overwhelmed by an inclination to kill without having been undermined by provocation), he will not be able to make use of the provocation defence. For, although the burden of disproving provocation lies with the prosecution, an habitually aggressive or violent person will be unable to avail himself of the provocation defence if he cannot establish that there was a reasonable possibility both that he had been provoked and that,

recent psychological and neuro-physiological work, to those drawn in A. Reilly, 'Loss of Self-Control in Provocation' 21 (1997) *Criminal Law Journal* 320.

³⁴ We restrict our discussion to the results of anger arising from provocation. Similar considerations plausibly result when the driving emotion is one of fear. We return to this below.

³⁵ See also *Luc Thiet Thuan v R* [1997] AC 131, PC, where the defendant's brain had been damaged, thus reducing his capacity for self-control.

as a result, he had lost his self-control. Where that is not the case, the trial judge is entitled to withdraw the issue from the jury.³⁶

We have described the irascible agent as someone who lacks sufficient self-control. But, given our account of self-control, there is surely another possibility: the agent who has sufficient self-control to constrain his anger, but whose self-control is always *undermined* by the slightest provocation. Such an agent would surely pass the subjective test (though he would be likely to fall to the objective). We concede the theoretical claim, but question the significance of the possibility. It is not that there would be no evidence to distinguish an irascible agent whose self-control was habitually undermined from one whose self-control was habitually overwhelmed. The former would show a systematic loss of self-control, which the latter would not—a point that we shall develop shortly. Rather, what is implausible is that we could ever have evidence that the irascible agent whose self-control is habitually undermined would, without that undermining, have a capacity of self-control that is sufficient to resist the provocation. For we have no evidence that that capacity is capable of anything much at all.

Irascible agents have a long-term lack of self-control. Other cases involve agents for whom, at the time of the provocation, self-control had already been temporarily lost: for instance, through drunkenness. There is good evidence that the effect of alcohol is primarily to remove self-control, rather than to induce violent behaviour *per se*.³⁷ So a drunken agent will also standardly fail the subjective test.³⁸

(ii) If there is evidence that the agent did possess the requisite power of self-control prior to the provocation, we then need to ask whether there is evidence that her self-control was *lost* as a result of the provocation. Since it is a *faculty* being lost, it should show up in a loss of self-control across the board. The contrast is with cases in which the faculty of self-control remains in place, but where it is ineffectual against the urge to retaliate: cases where, to use our earlier terminology, it is overwhelmed, not undermined.³⁹

³⁶ See *R v Miao (Hui)*, unreported [2003] EWCA Crim 3486, CA, 17 November 2003, *per* Rose LJ.

³⁷ For a review of the literature, see C. M. Steele and L. Southwick, 'Alcohol and Social Behavior I: The Psychology of Drunken Excess' 48 (1985) *Journal of Personality and Social Psychology* 18. Interestingly, the effect seems to be far less marked on women. See, for instance, P. Giancola et al., 'The Effects of Alcohol and Provocation on Aggressive Behaviour in Men and Women' 63 (2002) *Journal of Studies on Alcohol* 64.

³⁸ Note that, in cases in which the loss of self-control is caused by some factor other than the provocation, that factor need not *precede* the provocation. It may be contemporaneous with, or even subsequent to, it. If A provokes B and, mulling over the smart, B subsequently drinks heavily, as a result of the drinking loses his self-control, and then retaliates, it is not the provocation that was the cause of the loss of self-control (even though it was one of the causes of the retaliation).

³⁹ Overwhelming is what standardly happens in cases of temptation. Agents who succumb to temptation—eating more than they judge they should, starting an illicit affair, or whatever—will not normally lose self-control in other areas. Their self-control has been overwhelmed in one area, but it has not been undermined across the board. We further need to distinguish both cases of overwhelming and of undermining from cases in which the agent simply judges that the best course is to respond to provocation with violence, and does so. Once again, the provocation defence cannot be invoked, but now for the simple reason that there is no loss of control.

Of course, as we have seen, in losing the faculty of self-control the agent does not need to 'go berserk'.⁴⁰ Without self-control we can still have an intentional agent, one who will, for instance, make reasoned choices about the means of effecting the killing: she will be more likely to pick up a hammer than a cushion to use as a weapon, more likely strike the provoker than an innocent passer-by.⁴¹ Nevertheless, we would expect a loss of self-control to have a *systematic* effect on those actions that it would ordinarily inhibit. So, for instance, a lack of concern on the part of the agent for her own long-term welfare (manifested, say, in a lack of concern about being caught) would be evidence that self-control was in fact lost; in contrast, evidence of such concern (donning gloves before picking up the hammer to ensure that no finger prints were left) would be evidence that it was not.⁴² Thus, the Court of Appeal was right in *R v Ibrams and Gregory*⁴³ to point out that evidence of certain sorts of planning cuts against a claim of loss of self-control, although the Court was wrong to attempt to elevate this to the status of a rule of law. Such planning, however, must be of the kind that results from the distinctive exercise of the faculty of self-control; and planning or action which simply aims to ensure the success of the homicide is not of that kind. The fact that the accused in *R v Thornton*,⁴⁴ on receiving new provocation, went to the kitchen to arm herself with a carving knife, and then spent some time sharpening it, should not count against her claim that she had lost her self-control. In contrast, an elaborate plan to avoid detection should count against that claim.⁴⁵

There is a further feature that is brought by talk of *loss* of self-control, rather than of its *reduction* or *diminution*. This is the idea that loss of self-control is all or nothing. We mean this not in the sense that it is systematic, that if self-control is lost in one domain it is lost in all: we have already argued that that is quite true. Rather, what is meant is that loss of self-control cannot come in degrees, that it cannot be diminished without being entirely lost: one either keeps it intact, or loses it altogether. The model is not that of a dial that can be turned down, but of a rubber band that can be stretched and then

⁴⁰ See text above at n 9.

⁴¹ This is not to deny the existence of so-called 'displaced aggression', which does fall under the scope of the provocation defence. For a review of its strength, see A. Marcus-Newhall et al., 'Displaced Aggression is Alive and Well: A Meta-Analytic Review' 78 (2000) *Journal of Personality and Social Psychology* 670.

⁴² L. Berkowitz cites a revealing case in his study of violent offenders: a man who, in the course of a fight, had broken a bottle to make a weapon which he then put down on the grounds that it was too dangerous. Here we do have evidence that the man had not lost self-control. See L. Berkowitz, 'Is Criminal Violence Normative Behaviour?' 15 (1978) *Journal of Research in Crime and Delinquency* 148, cited in R. Baumeister, T. Heatherton and D. Tice, *Losing Control*, above n 21 at 231.

⁴³ (1981) 74 Cr App R 154.

⁴⁴ *R v Thornton* [1992] 1 All ER 306, CA.

⁴⁵ For the reasons explained in this paragraph, Horder's observation that 'rational forward-planning seems inconsistent with the spontaneity of response at the heart of a plea of provocation' (see 'Reshaping the Subjective Element in the Provocation Defence', above n 1) needs careful qualification.

will break. It is the model that we implicitly use when we say something inside us ‘snapped’⁴⁶ or ‘cracked’.⁴⁷

This idea is deeply embedded in the legal doctrine of provocation. But now we reach a problem, for some findings might be taken to indicate that this model is simply not consistent with the empirical facts. The problem comes from the idea that self-control requires effort. This, of course, is a familiar enough idea from first-person experience, and the empirical literature bears it out. The best evidence comes from measures of the physical arousal to which it gives rise. Ask someone to engage in self-control and he will show the standard signs of physiological arousal that accompany effort: changes in blood pressure, pulse, skin conductance, etc.⁴⁸ But effort comes in degrees. So might it not be that, when we talk of provocation causing loss of self-control, we should really be talking of it raising the effort needed if self-control is to be maintained? Indeed, some psychologists argue that it is almost never the case that one literally could not resist an inclination to retaliate to a provocation; it is simply that such resistance would require a very great deal of effort.⁴⁹ If this is right, it suggests that the all-or-nothing model is mistaken: since effort comes in degrees, and so does self-control.

We think that such an argument moves too fast. In the first place, it does not follow from the mere fact that self-control requires effort that it cannot be lost altogether. It could be that maintaining self-control requires effort up until the point at which it is lost, after which no amount of effort will bring it back. (Pulling on a rubber band requires effort, but that is our model for the thing

⁴⁶ See, for example, *R v Bratton*, CA, 17 July 1980, reported briefly as [1981] *Crim LR* 119 (*per* Watkins LJ): ‘Something in his mind snapped. He attacked his mother ferociously and killed her... For years, so he claimed, the appellant had been subjected to indignities in the home; he had been hectorated and domineered by his mother and treated as though he were still a small child about the house. She had indulged in sexual relations with [his] dogs... She seems to have been determined either to do away with one or both of the dogs... [T]here then came about the event of the fateful day, when all that had happened before welled up in front of him and suddenly he lost his self control and destroyed her’; *R v Williams*, CA, unreported, 1 July 1982 (*per* May LJ): ‘Because of the amount of drink that this man had taken, because of the relationship that there had been and still was between himself and the deceased and because of his fears that it might be coming to an end, as he said in his statement, “something snapped”’; *R v Parole Board, ex parte Bradley* [1991] 1 WLR 134 (QBD), (*per* Stuart-Smith LJ) at 137: ‘A little later he said: “I don’t know how to describe it. Something like snapped in my head and I sort of attacked her... I lost control, it was like as if I was somebody else. I put my hands on her neck and squeezed. I wanted to see what it would be like”’; *R v Humphreys* [1995] 4 All ER 1008, CA (*per* Hirst LJ) at 1012: ‘In her defence the appellant relied on the whole history, from the moment she first met Trevor Armitage, as a cumulative catalogue of provocative conduct against her, culminating with the jibe as to the inefficiency of her wrist cutting which, as already noted, was put forward as the trigger which snapped her self control’; *R v Lawson* [1998] *Crim LR* 883, CA (*per* Lord Bingham CJ): ‘[counsel for the appellant] paints a picture of a man who was driven beyond endurance by the weight of the burdens upon him until he finally snapped’; and *R v Connelly*, unreported [2002] EWCA Crim 170, CA (*per* Goldring J): ‘It was only after the third time that the deceased attacked him that he snapped. The provocation was intense, particularly for someone as immature as him’.

⁴⁷ See *R v Wheeler*, *The Daily Telegraph*, 25 September, 1986 (cited in Horder, *Provocation and Responsibility*, above n 9, at 109, n. 161).

⁴⁸ M. Muraven, D. Tice and R. Baumeister, ‘Self-control as a limited resource’, above n 27 at 774–5, contains a survey of the relevant literature.

⁴⁹ R. Baumeister, *Evil: Inside Human Violence and Cruelty* (New York: W.H. Freeman, 1997) 274–7; R. Baumeister, T. Heatherton and D. Tice, *Losing Control*, above n 21 at 230–2.

that can snap.) But suppose that we accept that the agent always could have maintained self-control if he had expended more effort: that, unlike the broken rubber band, the control lines still remain in place, even if they are far harder to manipulate. Does it follow that talk of the agent literally *losing* self-control is out of place? Should we not rather talk of self-control being *diminished*?

Still we think not. What is crucial for the subjective test is that the defendant loses self-control; whether or not he could have avoided losing self-control if he had exerted more effort is beside the point (although it is, of course, central for the objective test). Perhaps it is true that the defendant loses self-control because the effort needed to maintain it increased; but that is just to explain the way in which it was lost, not to deny that it has been lost.

What this line of thinking brings out is that the agent is to some degree compromised by his loss of self-control. Had he made more of an effort he could have maintained it, and this shows that some blame attaches to him, even if making more of an effort was, at that point, beyond him. Yet that seems right. Provocation is only a *partial* defence; if it brings no blame, should the agent not be *acquitted*? We shall return to this point later.

(iii) If self-control was lost as a result of the provocation, is there reason to think that this loss was the *cause* of the homicide? It is often said that loss of self-control is inconsistent with motives of revenge or punishment.⁵⁰ We suspect, however, that a desire for revenge or punishment is present in almost every case in which self-control is lost as a result of provocation. The crucial point is whether the agent would have been moved to kill by such a desire even in the absence of a loss of self-control. We said that, if the agent is to invoke a provocation defence, she must have had sufficient self-control prior to the provocation to resist any inclination to kill. But such self-control will only be employed if the agent believes that the best thing is not to kill—we characterized self-control as the ability to bring one's actions into line with what one judges best. It is quite possible for an agent who has self-control to embark on a plan to kill someone because he judges this the best thing to do: because he judges the victim deserves it, for instance. In such a case, if it is this prior decision that causes the killing, then a subsequent loss of self-control is beside the point. Since it is not the loss of self-control that causes the killing, the provocation defence cannot properly be invoked.⁵¹ It is here that we think that talk of motives, or better still intentions, of revenge and punishment are relevant to defeating a provocation defence: not because they are inconsistent with the defence, but because they can provide (defeasible) evidence that the loss of self-control was causally irrelevant.

⁵⁰ See, for instance, Lord Taylor CJ in *R v Ahluwalia*, quoted in the text at n54 below.

⁵¹ There are, of course, many different plausible accounts of the nature of causation; we are not here committing ourselves to any particular one of them.

5. *Why Must the Loss be Sudden and Temporary?*

We have so far said nothing to answer our third question, the significance of the words ‘sudden and temporary’ in the *Duffy* definition. This is sometimes read as imposing a requirement of contemporaneity: a requirement that the reaction follow *immediately* on the provocation. This in turn is sometimes held to be part of the very nature of provocation.⁵² This latter claim is surely wrong. If the provocation was sufficient to induce a loss of self-control, then (given that the agent will not have the capacity to wait on prudential grounds) it is likely that it will produce an immediate reaction. But there is nothing in the nature of loss of self-control that requires this. It might, for example, be that no opportunity to retaliate has yet presented itself. And, just as it is false to assume that once self-control is lost a reaction will follow immediately, so it is false to assume that a loss of self-control will occur, if it occurs at all, immediately a person is provoked. As many have come to realize, some people have slow fuses. It is perfectly possible to smoulder for a long time before finally losing control.

Could it be, though, that the words ‘sudden and temporary’ in *Duffy* embed an independent requirement of contemporaneity in the provocation defence, even if this does not follow from the nature of loss of self-control. Smith and Hogan, for example, came close to such a position when they said in the seventh edition of their textbook: ‘It seems that the function of these words [sudden and temporary] is to emphasise to the jury that there must have been no time in which D was able to (and, it is submitted, did) think and reflect between the final provocation and the fatal act’.⁵³ Even this guarded version of the contemporaneity requirement, however, became difficult to sustain following the decision in *R v Ahluwalia* where the then Lord Chief Justice, Lord Taylor, said:

Time for reflection may show that after the provocative conduct made its impact on the mind of the defendant, he or she kept or regained self-control. The passage of time following the provocation may also show that the subsequent attack was planned or based on motives, such as revenge or punishment, inconsistent with the loss of self-control and therefore with the defence of provocation. In some cases, such an interval may wholly undermine the defence of provocation; that, however, depends entirely on the facts of the individual case and is not a principle of law.⁵⁴

⁵² For instance, Horder claims that loss of self-control ‘by its nature leads to a sudden and immediate reaction to the stimulus of provocation’. See *Provocation and Responsibility*, above n 9 at 68.

⁵³ See *Criminal Law*, 7th edn (London: Butterworths, 1992) 355.

⁵⁴ See *R v Ahluwalia*, above n 3 at 897–8. In the 10th edn of his textbook on *Criminal Law* (London: Butterworths, 2002) Professor J.C. Smith wrote, at 368: ‘It seems that the words, “sudden and temporary,” imply only that the act must not be premeditated. It is the loss of control which must be “sudden,” which does not mean “immediate.”’

Clearly, Lord Taylor is here rejecting the idea that the provocation defence imposes a contemporaneity requirement; a delayed response is, at most, evidence that the defence is inapplicable for other reasons.⁵⁵ However, almost in the same breath that he rejected a contemporaneity requirement, Lord Taylor forthrightly defended the *Duffy* test. The expression ‘sudden and temporary loss of self-control’, he said, ‘encapsulates an essential ingredient of the defence of provocation in a clear and readily understandable phrase. It serves to underline that the defence is concerned with the actions of an individual who is not, at the moment when he or she acts violently, master of his or her own mind’.⁵⁶ This trenchant defence of *Duffy*, combined with a refusal to accept that the fatal act must have immediately followed the provocation, raises the question of how the words ‘sudden and temporary’ can underline the idea of an individual who is not master of his or her own mind.

We suggest that the primary force of these words is simply to underline the applicability of the model of loss of self-control that we have presented. If loss of self-control is *sudden*, then it is not a gradual process; rather, it happens in a stroke. Talk of sudden loss thus naturally suggests the idea of the broken elastic band, and of the loss being complete. Of course, there is no entailment here: there remains the possibility of sudden but partial loss. But a certain difficulty is logically precluded. Since the loss is sudden, there is no transitional period in which the agent has merely reduced self-control, and hence no problem about knowing how to treat a killing performed in that period.

It is less clear what idea is introduced by the insistence on *temporary* loss. Since we know of no cases in which this requirement has been invoked, we can only speculate. What this requirement presumably does is exclude non-temporary (i.e. permanent) losses of self-control from the ambit of the defence. What is the rationale for this? It could simply be that, when the loss is permanent, insanity and not provocation is the appropriate defence. But it is possible that the consideration here is also addressing one of the evidential questions raised above. Where self-control is evident as soon as the anger has subsided, we have good, though defeasible, grounds for thinking that it was indeed undermined by the anger, rather than having been absent in the first place. If so, though, it seems that temporariness should follow contemporaneity in being treated as an evidential consideration rather than a principle of law.

⁵⁵ What should we make of Lord Taylor’s insistence that motives of revenge or punishment are inconsistent with the defence, something that we have already denied? We suspect that what is being ruled out is a form of deliberate or planned revenge that would indeed hardly be compatible with a loss of self-control, since it would require the kinds of calculation that that state would preclude; however, this is hard to reconcile with Lord Taylor’s approval of Devlin J’s claim in *R v Duffy* (above n 5 at 932) that ‘circumstances which induce a desire for revenge are inconsistent with provocation, since the conscious formulation of a desire for revenge means that a person has had time to think, to reflect, and that would negative a sudden temporary loss of self-control which is of the essence of provocation’.

⁵⁶ See *R v Ahluwalia*, above n 3 at 895.

6. *Application of the Model to the Recent Controversy*

We have outlined an account of provocation that follows both the letter and the spirit of the law in taking the idea of self-control seriously. We now turn to apply this to some of the recent legal controversy. As we mentioned at the outset, much of this has focused on the issue of whether the different degrees of self-control that different people have should affect their eligibility to employ the provocation defence. The core of the debate has concerned the objective test, and its requirement that the provocation was sufficient to make a reasonable person do as the defendant did.

In asking whether a reasonable person could have acted as the defendant did, our aim is, of course, to ascertain whether the defendant's behaviour was itself reasonable. But we should not hold everyone to the same standard of reasonableness. Since different agents have different characteristics, what is reasonable for some might not be reasonable for others. If the objective test is to be sensitive to such differences, we must therefore in turn attribute these characteristics to the reasonable person with whom the defendant is to be compared. But that raises the question of whether this should be done with all of the defendant's characteristics. If the defendant is depressed or irascible or drunk, should we compare him with a reasonable person who is similarly depressed or irascible or drunk? Moreover, if we attribute these characteristics to the reasonable person who serves as our standard, is this to help us determine the *severity of the provocation* produced by the relevant action; or is it to help determine the *standard of self-control* that we expect in the face of such provocation?⁵⁷

It is now agreed that, when it comes to assessing the severity of the provocation, almost any characteristic can be relevant: what is provocative to, say, a reasonable Muslim might not be provocative to a reasonable atheist. Controversy starts over the issue of the characteristics that can be attributed to the reasonable person in order to determine whether that person might lose self-control in the face of a provocation of a certain severity. The position laid down in *DPP v Camplin*,⁵⁸ which was endorsed by a majority opinion of the Privy Council in *Luc Thiet Thuan v R*⁵⁹ and in *Attorney General for Jersey v Holley*,⁶⁰ was that only age and gender are relevant to this issue. While a person of tender years might not be expected to maintain the same level of self-control

⁵⁷ See A. Ashworth 'The Doctrine of Provocation' 35 (1976) *Cambridge Law Journal* 292, at 292 and *DPP v Camplin* [1978] AC 705 at 718, HL, where Lord Diplock drew a distinction between the 'gravity of the provocation' and the defendant's 'power of self-control'. The viability of this distinction was rejected by Lord Hoffmann and Lord Clyde in *R v Smith* [2000] 3 WLR 654, HL, *per* Lord Hoffmann at 673-4 and *per* Lord Clyde at 689, but has recently been approved by the majority of the Privy Council in *Attorney General for Jersey v Holley*, above n 3.

⁵⁸ See above n 57 at 718.

⁵⁹ [1997] AC 131.

⁶⁰ See above n 3.

required of a more mature agent, an irascible or drunk agent should be assessed by the standards of the even-tempered or sober.

Yet, in *R v Smith*,⁶¹ the House of Lords opted, by a majority, to broaden the class of relevant characteristics to include not just depression, the factor that was relevant there, but any characteristic that might be thought to have a 'just' bearing on the defendant's ability to exercise self-control, whether or not that characteristic amounted to a recognized medical disorder. Clearly the word 'just' was expected to do a large amount of work. Furthermore, it was to be left entirely to the jury, with no further guidance from the trial judge, to determine which characteristics were relevant. This gave the jury considerable latitude over the scope of the defence: might it not conclude that even drunkenness or irascibility were relevant?

Concerns here are greatly lessened by the Court of Appeal's recent ruling in *R v James and Karimi*⁶² that the decision of the Privy Council in *Holley* had had the effect of over-ruling the decision of the House of Lords in *Smith*. Yet, even without this ruling, we suspect that the latitude introduced by *Smith* was far less than it has been taken to be, since the function of the subjective test has been insufficiently appreciated.

From what we have said so far, it should be clear that the law restricts the scope of the provocation defence so that it cannot be extended to individuals who are habitually short-tempered or drunk. These are factors that themselves remove self-control, so it is simply not there to be lost as a result of provocation. We do not mean that such people will have no self-control whatsoever; but that they have insufficient self-control to inhibit them from homicide, and hence will not have something that the provocation can undermine. This is particularly important in the English context. Section 3 of the Homicide Act 1957 specifically requires that decisions on the objective test be left to the jury; the judge is not entitled to withdraw the issue. But no such statutory requirement covers the subjective test.⁶³

What does our account have to say about youth, the main factor that, since *DPP v Camplin*, has been recognized as relevant to the objective test,⁶⁴ or about depression, the factor that was central in *Smith*? We start with youth, where it might look as though our account says too much. For it might appear that, in so far as youth is relevant, it can only serve to reduce the amount of self-control that the defendant has. On this view, youth would be much like irascibility. But, then, like irascibility, it would be inconsistent with passing the subjective test.

⁶¹ See above n 57.

⁶² [2006] EWCA Crim 14; [2006] 1 All ER 759.

⁶³ Where there is no reasonable possibility that the defendant had been both provoked and, as a result, lost his self-control a judge is entitled to withdraw the issue of provocation from the jury (see above n 36).

⁶⁴ We leave aside discussion of the role of gender, since we know of no cases where this has been an issue.

However, that is not the only way to understand the effects of youth. There is an alternative that, we suggest, is better supported by both common observation and the scientific findings. The alternative holds that youth does not reduce the amount of self-control that the defendant has. Rather, it makes that self-control more vulnerable to undermining by factors like provocation. Youth is not like irascibility, whether this is understood as a tendency to react violently without losing self-control, or as the tendency to lose self-control habitually. Indeed, a youthful defendant who did exhibit general irascibility would be less able to avail himself of the provocation defence on precisely that ground. The effect of youth is on the defendant's behaviour in extreme circumstances. Extreme provocation will be more likely to result in a youthful defendant's self-control being undermined and hence tip him into killing. The empirical evidence, such as it is, bears out this picture. It looks as though the adolescent and young adult brain is simply less efficient at using the pre-frontal cortex for tasks like self-control: it is more likely to be overloaded.⁶⁵

What of the issue in *R v Smith*: depression? This will turn on the way that depression affects those who have it. Does it have the effect of removing the agent's self-control? If so, a depressed agent will not pass the subjective test, since that agent's self-control will already be absent at the time of the provocation. Or does it rather give a different kind of self-control, a kind that is more easily undermined? If so, the depressed agent will follow the pattern of the youth in passing the subjective test—although that does not, by itself, imply that youth and depression will fare the same when they confront the objective test. This is a difficult question, not least because a tendency to angry response is not even a typical characteristic of depression: it is not, for instance, listed in the criteria given in *DSM-IV*.⁶⁶ On the contrary, depression tends to flatten affect.

We incline to the view that, in so far as there is reason to go one way rather than the other, serious depression tends to remove self-control, rather than rendering it more readily undermined in extreme circumstances. Evidence for this comes from the observation noted above that depression, in so far as it affects self-control, tends to remove it across the board: it typically removes, for instance, the ability to stick to a diet. There is no evidence of a pattern of behaviour amongst the depressed that mirrors the pattern we have attributed to the young. Thus, we take issue with the implicit contention in *Smith* that depression, in so far as it is a relevant factor in a provocation defence, can be relevant in a way that is consistent with that defence. Indeed, so much

⁶⁵ See, for instance, B. Luna and J. Sweezy, 'The Emergence of Collaborative Brain Function: fMRI Studies of the Development of Response Inhibition' 1021 (2004) *Annals of the New York Academy of Science* 296; and, for a popular review, L. Sabbagh, 'The Teen Brain, Hard at Work', (2006) *Scientific American Mind* 20.

⁶⁶ *Diagnostic and Statistical Manual of Mental Disorders*, 4th Edn (Washington DC: American Psychiatric Association, 1994).

should be clear from Lord Hoffmann's own words. For, in summarizing the evidence presented in the initial trial, he said:

A psychiatrist called by the defence, who had seen Smith in prison less than a fortnight after the offence, said that he was suffering from an abnormality of mind, namely depression, which could reduce his "threshold for erupting with violence". Another said that he was suffering from clinical depression which made him "more disinhibited," i.e. less able to control his reactions.⁶⁷

The first of these comments is consistent with either interpretation. But, according to the second, it seems clear that, in so far as it had had an effect, the depression worked to remove Smith's self-control. According to this psychiatrist, it was the depression, and not the provocation, that had made Smith less able to control his reactions. If this is right, then depression, in so far as it is relevant at all to the provocation defence (which we rather doubt), works in a way that is inconsistent with that defence. In saying this, we do not contend that depression is irrelevant as a factor supplying the basis for a partial defence against murder charges. But these considerations indicate that its application should be by means of the diminished responsibility defence laid out in section 2 of the Homicide Act 1957, and not by means of the provocation defence as it is currently formulated.

Our main conclusion, then, is that the broadening of the objective test which the House of Lords endorsed in *Smith* was far less significant than many commentators, and perhaps the House itself, have thought. Most of the factors that were discussed—depression, drunkenness and chronic bad temper—act, in so far as they act at all, to remove adequate self-control in the first place. Thus, the broadening of the objective test in *Smith*, in reality, left it much where it was left by *Camplin*. The effects of youth, we have argued, are plausibly compatible with the defence; but we do not see any others that are.⁶⁸

A second implication is more positive. We have identified killing-as-a-result-of-loss-of-self-control-as-a-result-of-provocation as a very specific event. What would it be for it to happen to a reasonable person? Lord Millett, in his dissenting opinion in *Smith*, argued that it cannot be seen as a case in which the defendant acts reasonably:

[I]t can never be reasonable to react to provocation by killing the person responsible. Nor by pleading provocation does the accused claim to have acted reasonably. His case is that he acted unreasonably but only because he was provoked. But while this may not be reasonable it may be understandable, for even normally reasonable people may lose their self-control and react unreasonably if sufficiently provoked.⁶⁹

⁶⁷ Per Lord Hoffmann in *R v Smith*, above n 57 at 664.

⁶⁸ Again we leave aside the question of gender, having little idea of what is supposed to be at issue there.

⁶⁹ See above n 57 at 712.

In responding to this passage, Gardner and Macklem point out, quite rightly, that Lord Millett failed to make some relevant distinctions. Assimilating the idea of reasonableness to that of justification, they write:

But what is held out as justified, in the law of provocation, is not the killing, but rather the loss of temper which caused the defendant to kill. She was justified in “doing as she did”, as section 3 puts it, where this means “in getting so angry that she lost self-control to the point at which she killed”.⁷⁰

We contend that they have still not got to the bottom of the issue. It is one thing to get angry; it is another to lose one’s self-control. Yet it is the latter that is relevant to the objective limb of the provocation defence. (To talk of ‘losing one’s temper’, in the way that Gardner and Macklem do, but the Homicide Act does not, is to blur the distinction.⁷¹) Perhaps it is justifiable to become angry in the face of provocation. But is it justifiable to lose one’s self-control? We think not. Either one thinks of it as something that one does, in which case it is surely not justified. Or one thinks of it as something that just happens to one, in which case talk of it being either justified or unjustified is inappropriate (as one is neither justified, nor unjustified, in sneezing). And this surely entails that if the truth is somewhere between the two models—if losing one’s self-control is something that happens to one, but that one could, with sufficient effort, resist—then we should not think that it is ever justified.⁷²

This takes us back to something closer to Lord Millett’s position: that while the loss of self-control is not reasonable, it is something that can happen to the reasonable person. But, since loss of self-control is such a specific failing, to acknowledge that it has been lost is not to say that all of our expectations of reasonable behaviour should also be lost. As we have seen, people who lose their self-control will usually remain agents, acting on their beliefs and desires. Importantly, this means that we can expect them to keep some sense of proportion. This enables us to endorse Lord Diplock’s comments in *Phillips v R*. Giving the opinion of the Privy Council, he said:

Before their Lordships, counsel for the appellant contended, not as a matter of construction but as one of logic, that once a reasonable man had lost his self-control his actions ceased to be those of a reasonable man and that accordingly he was no longer fully responsible in law for them whatever he did. This argument

⁷⁰ J. Gardner and T. Macklem, ‘Compassion without Respect: Nine Fallacies in *R v Smith*’ [2001] *Crim LR* 623 at 628.

⁷¹ The importance of keeping this distinction clearly in mind has been given judicial recognition by the Court of Appeal in *Attorney General’s References Nos 74 of 2002, 95 of 2002, and 118 of 2002 (Darren Swatan, Leslie Humes, and Mark Wilkinson)* [2002] EWCA Crim 2982; [2003] 2 Cr App R (S) 42, where Mantell LJ said of the provocation defence (para. 24): ‘at the time of the killing, [the offender must have] lost his self-control. Mere loss of temper or jealous rage is not sufficient’.

⁷² Gardner and Macklem, above n 70, go on to say that ‘only if our beliefs, passions, attitudes (etc.) are justified are our actions on the strength of them excused’. We doubt that this is in general so. We often excuse actions performed on the basis of mistaken beliefs without thereby thinking that the beliefs were justified: it might simply have been unreasonable to expect the agent to arrive at a justified belief. But to pursue this issue would take us too far afield.

is based on the premise that loss of self-control is not a matter of degree but is absolute: there is no intermediate stage between icy detachment and going berserk. This premise, unless the argument is purely semantic, must be based upon human experience and is, in their Lordships' view, false. The average man reacts to provocation according to its degree with angry words, with a blow of the hand, possibly, if the provocation is gross and there is a dangerous weapon to hand, with that weapon.⁷³

We argued above that loss of self-control is not a matter of degree. But this is not to disagree with Lord Diplock's comments here. For what he is insisting is that the agent who has lost his self-control should retain a sense of proportion in what he does; and this is something with which our model of self-control can agree.

A final point should be stressed, one that we mentioned earlier. We do not have to think of loss of self-control as entirely outside the control of the agent. An agent who tries harder may be able to retain control; and even if he has lost it, may be able, by effort, to regain it. A useful parallel can be drawn with crying. Crying is a real phenomenon, and is, in an important sense, involuntary: most people cannot make themselves cry by simple decision. Yet, typically, people can resist crying, or can stop themselves once they have started; and how readily they cry, and in what way, will be greatly influenced by social convention. Resisting crying can take enormous effort though, and it is at least partly for this reason that we do not always blame people for crying, even in cases where it would be much better if they did not.⁷⁴

Plainly, losing self-control in such a way that one kills someone is a much more significant thing than bursting into tears. But, like crying, resisting it can take great effort.⁷⁵ The law thinks agents should make that effort, as is shown by the fact that the provocation defence does not lead to an acquittal. Equally though, it thinks that, if the effort was one that could not reasonably be required, the offence is less grievous than murder. So here is a second role for the objective limb: not just to ensure that defendants retain some sense of proportion, but to ensure that they cannot reasonably have been required to do more to resist the loss of control.

7. Further Implications of the Model

We conclude by outlining some further implications of the model that we have developed. In particular, the model makes clear three ways in which the

⁷³ (1968) 53 Cr App R 132 at 135.

⁷⁴ We were led to consider the case of crying by David Velleman's discussion in 'How We Get Along', ms, available at <http://homepages.nyu.edu/~dv26/>.

⁷⁵ We thus take a more nuanced position than Horder who takes the fact that an agent 'permits' him or herself to lose self-control as showing that the loss of self-control is not involuntary. (See 'Reshaping the Subjective Element in the Provocation Defence', above n 1 at 128). In one sense, this is right; but losing self-control is hardly a voluntary action either.

provocation defence is currently restricted; and correspondingly, three ways in which it might be broadened. In the first place, the defence might be broadened to include killings that result from loss of self-control in response to factors other than anger: factors like fear, or 'extreme emotional disturbance' more generally. Second, it might include killings that result from such factors in the *absence* of self-control, with no requirement that the self-control have been *lost*. Third, it might include killings that result from such factors even when self-control is still present.

The provocation defence in its current form has emerged from a long history;⁷⁶ it would not be surprising, therefore, if it contained features that were not applicable to modern times. Broadening the defence in something like the first way does seem to involve the removal of a feature that is largely arbitrary. As Horder has recently stressed, why should somebody who acts from fear rather than anger not be able to avail herself of the defence?⁷⁷ Fear, too, can undermine self-control, and, as Horder points out, is a more likely response to provocation in cases where the provoker is the more powerful. Legal opinion is turning to recognize this. In *Smith*, Lord Hoffmann indicated that there are cases where fear may now provide a sufficient basis for invoking the provocation defence: 'the law now recognizes that the emotions which may cause loss of self-control are not confined to anger but may include fear and despair'.⁷⁸

In contrast, we contend that broadening the provocation defence in either the second or third ways would involve a radical revision, something that would result in a totally new defence.⁷⁹ It is of the essence of the current defence to accommodate a particular human failing: the tendency to lose self-control in response to provocation. If this is broadened to include cases where self-control is lacking for some other reason, then we open the door to a huge number of possible claims. Perhaps some of these have a legitimate standing; others—where self-control is missing from irascibility or drunkenness, or from a frustrated sense of entitlement, or a festering resentment—do not. Opening the defence to include killings where self-control is not absent introduces even more. Perhaps an objective test can be used to rule out all of these cases, but there is no guarantee that it will. And even if it does, the invitation to spurious defences given by broadening the provocation defence, and the complexity and uncertainty this would introduce, are to be avoided. Better, surely, to retain the

⁷⁶ For an excellent review see Horder, *Provocation and Responsibility*, above n 9.

⁷⁷ See 'Reshaping the Subjective Element in the Provocation Defence', above n 1.

⁷⁸ See above n 57 at 673.

⁷⁹ The Law Commission has suggested taking what is, in effect, the third of these options, dropping a requirement of loss of self-control altogether; see above n 1.

provocation defence just for the job that it has evolved to do, and then to introduce further defences for further cases should these prove necessary.⁸⁰

⁸⁰ Horder aims to bring further cases under the scope of the provocation defence—a defendant who kills an assailant who indecently assaults her, fearing, unreasonably but as a result of the extreme emotional disturbance that the assault engenders, that he will go further; or a defendant who kills a burglar in his house, fearing, again unreasonably but as a result of the extreme emotional disturbance engendered, that that he might attack him. These cases would fall under the scope of the minimally extended provocation defence that we advocate. For in both cases there is a loss of self-control, though one manifested in fear rather than anger. What is distinctive in Horder's proposal, and not captured by ours, is a requirement that fear for one's own safety (or the safety of others) be a necessary condition for employing the provocation defence. See 'Reshaping the Subjective Element in the Provocation Defence', above n 1 at 134ff. We leave open the question whether this would be a good reform, merely noting that it is quite consistent with what we are arguing here, namely that the loss of self-control should remain necessary too.

