Culpability and Mental Disorder*

R. CUMMINS, University of Wisconsin-Milwaukee

I

It is widely held that the insane are not culpable for their offenses, or at least that they are not culpable for those offenses which are specific manifestations of mental disorder. Indeed, this is generally regarded by philosophers as a platitude. Yet it seems to me that what we have here is not a platitude, or even a truth, but an overly broad swipe with a dull blade. When we have brought the issues into sharper focus I think we will see that the mere fact that an offense was a manifestation of mental disorder never in itself exculpates, though it is sometimes relevant to determining whether the action was excusable in one of several standard ways. Moreover, I think it emerges that the issue is not nearly as significant as one might initially suppose it to be.

* This work was supported in part by a grant from the National Endowment for the Humanities.
Robert Cummins

II

Does mental disorder exculpate? There are two positions one can take. The first, which I call the Conservative Position, is that mentally disordered persons are culpable for their offenses unless they satisfy one of the "ordinary" exculpating conditions, i.e., an exculpating condition that could just as well be satisfied by someone who is psychologically normal. A mentally disordered person may satisfy an exculpating condition as the result of his mental disorder just as a color-blind person may satisfy an exculpating condition as the result of his color-blindness. But the fact that the offense was the characteristic result of mental disorder is never in itself sufficient to exculpate, according to the Conservative, any more than the fact that an offense is the characteristic result of color-blindness is sufficient to exculpate. Thus, I may adduce my color-blindness in support of my claim that I couldn't have known which button to push to stop the old elevator, i.e., in support of my claim to a "standard" excuse. But color-blindness does not constitute an independent excuse in itself, as is evident from the fact that its relevance evaporates if it emerges that the various buttons were clearly labeled. The Conservative holds an exactly parallel view concerning mental disorder: it may on occasion support a claim to satisfy an ordinary exculpating condition, but it is not itself an exculpating condition. This line, with which I have great sympathy, is evidently a highly conservative line in that it allows no exculpating conditions to the mentally disordered that are not equally satisfiable by the psychologically normal: hence the name "Conservative Position."

It might seem that there is one respect in which even the Conservative must allow that the mentally disordered stand differently from normal agents with respect to excuses. To harken back to the example supposed by the Chief Justices commenting on M'Naughten's case in 1843, suppose D believes, under the influence of insane delusion, that P is in the act of attempting to take away his life, and D kills P believing this to be self-defense. Here we have a mistake of fact, and hence an ordinary unqualified excuse for killing P. Now, under normal circumstances, we might very well turn our attention to the question of whether

---

1 This is the view put forward by Sir James Stephen in 1883 (A History of the Criminal Law of England, (London: MacMillan, 1883), 2:97, 2:125), and by Herbert Wechsler in 1955 ("The Criteria of Criminal Responsibility," University of Chicago Law Review, 22:367, p. 373). What appears to be an example of Conservative reasoning occurs in the answer the Chief Justices gave to the fourth question posed them by the House of Lords in connection with M'Naughton's case (10 Clark & Finnelly 200 (1843)).
the mistake was one that D should not have made; whether D was
negligent in supposing that P threatened his life. But in this case it may
seem that no such further question is admissible, since the crucial sup-
position on D's part is conceded to have been the result of insane
delusion.

For the sake of a clear opposition, and because it seems to me the
right line to take, I want the position I have called the Conservative
Position to be understood to admit no such exception. In the case
imagined, we have shifted attention to another offense, and this is to be
considered in the same way as the other: perhaps D has an excuse, and
perhaps not, but the mere fact of insanity — in this case the mere fact
that the supposition was the result of delusion — is not to be treated as
an excuse in itself. No doubt it will seldom happen that mistake due to
insane delusion is culpable. But the Conservative regards this as a con-
tingent matter to be established independently in each case.

Opposed to the Conservative Position is the Liberal Position, viz.,
that mental disorder, at least in some cases, precludes culpability in its-
self, quite independently of the excuses available to the sane. This
position has been most explicitly articulated and defended by Fingaret-
te, but it is, I believe, the majority position, at least outside of the legal
profession. Since no one wishes to deny that the Conservative Position
formulates a sufficient condition for exculpating the mentally disor-
dered, the difference between the two positions is simply over what is
necessary: where the Conservative Position holds that an exculpating
condition of the ordinary sort must be satisfied, the Liberal Position
holds that in some cases mental disorder precludes culpability in itself.

III

If we are to arbitrate this debate, it is essential to begin by getting
tolerably clear about what the "ordinary" exculpating conditions are
and why they exculpate. I divide these into defenses and excuses, and
take them up in turn.

2 Herbert Fingarette, The Meaning of Criminal Insanity (Berkeley: The University

3 It is obviously the Liberal Position that underlies the Durham decision (94 US
    App DC 228, 214 F2d 862, 45 ALR 2d 1430, (1954)) and the New Hampshire
    Rule (State v Pike, 49 NH 399 (1870)).
Defenses. If an offense is attributed to someone he may attempt to refute the allegation outright, either by denying authorship, or by offering a justification, i.e., an argument designed to show that the alleged offense was not actually an offense at all.

Excuses. When a defense is not available — i.e., if it is admitted that there was a genuine offense committed by the person in question — then the question arises as to whether that person has an excuse for committing that offense. Thus, the shift from defenses to excuses corresponds to a shift from a consideration of the deed — whose deed it was and whether it was actually wrong — to a consideration of the agent — whether his commission of this offense reflects ill upon him or not. Excuses, then enter the picture when we are attempting to morally assess the agent in the light of a particular offense he has committed.4

I shall need a term for the kind of thing that is at stake when the question of excuses does arise, i.e., a way of characterizing someone who has committed an offense and has no excuse. I propose to say that such a person is culpable on account of the offense in question. I prefer ‘culpable’ to ‘blameworthy’ or ‘to blame’ mainly because ‘culpable’ has no verb form. There is a temptation to assume that someone who is blameworthy, or to blame, ought to be blamed, or deserves or merits blame, or that blaming such a person wouldn’t be wrong. This is harmless if to blame someone is simply to judge, in the privacy of one’s own thoughts, that the person in question is culpable on account of his offense. But many writers, notable Brandt5 and Glover6 have assumed that blaming involves rather more than this, and it is a substantive moral question whether this something more is justified to any degree whatever by the mere fact of culpability. In the following discussion, therefore, I propose to distinguish sharply between factors bearing on

4 Some would hold that an agent with a valid excuse for A has done nothing wrong in doing A, and hence that A is not an offense in that case. If we say A is not an offense when the agent has an excuse for doing it, we obscure the difference between justifications and excuses. Those who think saying that x has an excuse on account of A precludes saying that A (the token) was wrong can avoid obscuring the distinction by thinking of an offense simply as a morally bad state of affairs which may or may not be due to wrong-doing.


culpability and factors bearing on how one should treat persons who have committed offenses. Failure to do this can only obscure the nature of culpability by encouraging a confusion between excuses on the one hand and reasons for lenient or compassionate treatment on the other.

Opinions differ concerning what excuses and what doesn’t. But given a valid excuse for committing a specified offense, it follows that commission of that offense does not reflect ill on or tend to morally discredit the agent. Excuses thus drive a wedge between agent and act where the act would otherwise morally taint the agent.

The approach to culpability and excuses I will take bears a close resemblance to something suggested by James Wallace.7 Wallace writes, “One way of faulting an agent for an action ... is to establish that the action is fully characteristic of some vice.”

An act is characteristic of a vice if it is the sort of act which standardly manifests it, as stuffing oneself standardly manifests gluttony, and lying standardly manifests mendacity. An act is fully characteristic of a vice if it is characteristic of that vice, and actually a manifestation of it. Thus bravado may be a manifestation of cowardice, but is not characteristic of it, and lying, though characteristic of mendacity may, on some occasions, manifest a virtue and hence not be fully characteristic of mendacity.

Using these notions we can formulate a helpful first pass at the point I am trying to make about the role of excuses in moral assessment by saying that D has an excuse for committing an offense 0 only if 0 was not fully characteristic of any vice of D’s. I have formulated this as a necessary condition only. Certainly D has no excuse if his offense is fully characteristic of some vice on his part. But there may be offenses which are inexcusable yet not fully characteristic of any vice: at least there are inexcusable offenses which manifest no vice with a name, and this is probably enough to doom a claim of sufficiency to objections which could only be removed by stipulation. We can repair this by replacing ‘vice’ by ‘moral defect’, but we pay a price: it is a good deal clearer what counts as a vice than what counts as a moral defect.

Still, the schema is suggestive: the idea is that we make rather finegrained assessments of agents in the light of particular offenses, and these assessments are regulated, channeled and corrected by admitting

excuses. When we ask whether D has an excuse, and specify the offense carefully, we are asking a rather sophisticated question: Is that offense fully characteristic of the sort of relatively permanent trait we call a vice or moral defect, i.e., a trait that is appropriately assessed morally? The problem we will encounter discussing agents with a mental disorder is not whether their offenses manifest relatively permanent and undesirable tendencies, but whether the tendencies manifested can count as moral defects, or as analogous to moral defects, for purposes of moral assessment. The concept I've been explaining is not the only possible or useful concept of culpability, but it is plainly the one wanted for present purposes, for the question is how and when (if ever) mental disorder precludes a negative moral assessment made on the basis of conduct. And it is surreally the central concept in any case. Moral assessment of persons is important because of the bearing it has on how to behave vis-à-vis the persons assessed. It is offenders, not conduct, we must incarcerate or treat or blame, and our behavior is rational only in so far as it fits with what we can discover about offenders from their conduct.

IV

Is Mental Disorder an Excuse?

Excuses mediate and regulate judgments concerning the bearing of an offense on the moral character of the agent, and we need now to inquire whether or not mental disease or psychological disorder ever plays this sort of role.

Let us begin by considering a "hard case" for the Conservative Position — a case in which the offense is clearly a manifestation of a serious psychological problem but which does not involve any ignorance of fact or ordinary incapacity of the sort which the Conservative could fall back on in explaining whatever intuitions we might have concerning a lack of culpability. Here is the case. Suppose a man walking along a deserted beach. He tops a dune and suddenly hears the cries and sees the waving arms of a child on a rubber raft being swept out to sea by the tide. He was a champion swimmer in college just six years ago, and he remains physically fit. He is thus able, in the ordinary sense, to rescue the child, and knowing this, he immediately starts toward the water. But as he nears the edge he hesitates, shows signs of tension, starts, then stops again, then starts again, and finally wades in woodenly. As the water reaches his knees, he stops and stands rigidly, face pale,
breathing shallowly. After about a minute in this state, he backs slowly out of the water and stands watching until the child is out of sight.

Here we certainly have a manifestation of a stubbornly permanent tendency, a tendency that is a crucial part of the man’s psychological make-up. The man has a neurotic fear of water (I don’t mean to suggest that he has some psychotic delusion, e.g., that he will dissolve); he is irrational in his actions, though not his beliefs, where water is concerned. Fingarette points out that an important feature of the sort of irrationality in action or belief which characterizes mental disorder is that we recognize it as part of the person’s nature. (Fingarette, p. 193) Thus we have this much analogy with the culpable conduct of a coward: our man is not a victim of an unfortunate mistake of fact, e.g., that the whole thing is a practical joke, nor did he lack the requisite skills, nor was he restrained by fences or incapacitated by a broken arm or exhaustion. Instead his conduct manifested a relatively permanent and undesirable tendency which is part of his psychological make-up or nature. In all this our man is indistinguishable from someone who manifests plain cowardice in not rescuing the child. Of course, one would resist the claim that the tendency manifested was a vice, or even that it was a trait of character, and cowardice is certainly both of these. And there are other important differences. Still, enough has been said to make it evident that mental disorder is not an excuse, i.e., that its relevance to culpability, if any, does not depend on its blocking inferences from act to the character or “nature” of the agent in the way that ignorance or inability blocks such inferences. An action fully characteristic of a mental disorder does show us something about the agent’s more or less permanent dispositions to act in a way in which unwitting action does not. When D blows up P’s house because X has secretly wired D’s light switch to a bomb, D does something which is not only not fully characteristic of any vice of D’s, it is not fully characteristic of any behavioral disposition of D’s. In contrast, mental disorder is relevant only when the action in question is a manifestation of some behavioral disposition of the agent. This, of course, does not show that mental disorder does not preclude culpability, but it does show that mental disorder does not preclude culpability in anything like the way the standard excuses do. Ignorance of what one is doing and mental disorder bear on culpability in such different ways that only confusion can be generated by lumping them together under one heading.
V

Granted that mental disorder isn’t an excuse, let us turn now to the claim that mental disorder exculpates in some other (perhaps unique) way. And let us begin by disarming two common sources of the feeling that mental disorder must exculpate somehow.

First, it is fairly common to represent mental disorder as a justification on the grounds that, in cases like the one rehearsed in the last section, a mentally disordered person is simply not able to do what wants doing. Since one cannot be obligated to do what one cannot do, there is no wrong-doing in the picture. But, in addition to the fact that it is notoriously difficult to specify a sense in which the neurotic is “not able” that does not apply equally to the coward (they can both swim, they were both unable to overcome their fear), I think this idea rests on a fundamental mistake. For, as Frankfurt has shown, the fact that an agent could not do otherwise does not exculpate him. Suppose A sees his enemy immobilized with cramps in the water along a remote and deserted beach and decides not to go in after him. Subsequent discovery that A cannot swim does not alter our judgment that A is culpable unless A realized at the time of the incident that he could not swim. This suggests that the unavailability of alternatives does not exculpate unless it was, in some sense, the agent’s reason (in part anyway) for doing what he did. Even if we grant that our illustrative phobist was unable to enter the water, it seems clear that his incapacity (if such it was) need not have been — typically would not be — any part of his reason for not entering the water.

In any case, this move plays directly into the Conservative’s hands. It is simply an attempt to argue that, because of his psychological condition, the neurotic satisfies an ordinary exculpating condition. This point is not undermined by the suggestion that only the mentally disordered are “unable” in the special way imagined in the example. Even if this were true (which seems unlikely), the point would still be that we have is a case of inability: special though it is, it must be represented as a species of the same generic condition which justifies generally — i.e., as a condition in which there is really no offense in the picture — or the whole move will lose its point. Similarly with attempts to argue that the mentally disordered cannot control themselves or their actions, that they are compelled, that they cannot “do otherwise”: however successful these arguments are in themselves, they cannot help the Liberal.

---

For they are surely attempts to show that, when an action manifests a mental disorder, the agent satisfies an exculpating condition that anyone might satisfy. I don’t think a lack of alternatives is, by itself, enough to exculpate an agent in any case, but even if it were, it would do the business regardless of whether the lack of alternative possibilities was due to mental disorder or to something else.

The second source of the feeling that mental disorder must exculpate somehow is the belief that it would surely be a mistake, or wrong, or a miscarriage of justice, to punish someone for an action which was fully characteristic of a (serious) mental disorder. This, of course, is quite right. But it is often coupled with the belief that it is not mistaken, wrong, or a miscarriage of justice, to punish someone for an action culpably done. One then concludes that an agent cannot be culpable on account of an action fully characteristic of a mental disorder. The second premise is sometimes only implicit: one notices that it would be wrong to punish a defendant for an offense which was fully characteristic of a mental disorder and concludes that such a defendant should be found not guilty. The missing premise is supplied by the legal context: a finding of guilty makes punishment permissible and sometimes mandatory.

Here the legal analogy is especially dangerous. A finding of guilty does make punishment legally permissible. That is, the legislature has passed a law permitting punishment (by proper authorities) of those found guilty of certain crimes. But there is clearly no analogous moral law which permits or mandates any special treatment whatever, leave alone punishment, for an action culpably done. And it is just as clear that the legislature could enact a law blocking punishment (and perhaps permitting or requiring other kinds of treatment) on a finding of ‘guilty but insane’. Once we recognize that there is no legitimate moral reasoning from the fact of culpability to any conclusion permitting aversive treatment of the agent, the obvious fact that the mentally disordered require special treatment becomes irrelevant to the question at hand, and the argument collapses.

This dismissal will seem high-handed to retributivists, for the thesis that we can argue from the fact of culpability to the permissibility of some form of aversive treatment (punishment, informal sanctions, or whatever) is the heart of moral retributivism. This is not the place to rehearse the difficulties with retributivism. I will content myself with two points. First, it is not enough to argue that the culpable deserve aversive treatment, for it won’t follow without more ado that anyone is justified in meting it out. Perhaps only the innocent (or properly authorized)
Robert Cummins

should throw stones. Second, it would be striking if the only or most plausible defense of the Liberal Position depended essentially on strong retributivist principles. Most Liberals, especially those who are psychiatrists, are unlikely to buy exculpation for the mentally ill at the price of adopting a retributivist line vis-à-vis the coward. Liberals have typically advanced their views in the name of humanitarian concerns quite foreign to retributivism.

VI

If the Liberal position is to have a hope of success, it must identify some morally relevant difference between cases in which conduct is fully characteristic of a moral defect and cases in which conduct is fully characteristic of a mental disorder, and this difference must not simply be one that justifies differential treatment, or one that, like an inability to walk or see, is relevant only because it is often evidence of a standard exculpating condition. Mental disorder is relevant only when the conduct in question is fully characteristic of the disorder and therefore does support a negative assessment of sorts. Thus the Liberal Position reduces to the assertion that mental disorders are not moral defects, i.e., not proper subjects of moral assessment.

This is certainly a very tempting position. The suggestion that it is morally bad to be agoraphobic in anything like the way it is morally bad to be cowardly strikes one as ignorant, perverse, or both. Still, though, we need to know what, if anything, underlies this intuition. If the Liberal position is to be more than merely tempting, it must be supported by some account of the moral assessment of agents in the light of their offenses — i.e., an account of culpability judgments — which explains why certain dispositions (cowardice, deceitfulness) are appropriate objects of moral assessment while others (phobias, color-blindness) are not.

Let us return, then, to the comparison of cowardice and phobias to see if we can uncover a difference which might support the distinction.

9 Even those with retributivist sympathies seem to recognize this. E.g., Davis argues that the guilty deserve to suffer but is careful to point out that this is consistent with the view that one ought never to impose suffering. (Lawrence Davis, “They Deserve to Suffer,” Analysis, March 1972).
Culpability and Mental Disorder

1. The coward will typically over-estimate the danger, under-estimate his capacities, or misjudge the relative importance of securing his own safety as opposed to accomplishing the deed, and do this in ways that, though perhaps self-deceptive, are nevertheless different from honest mistakes in being essentially self-serving and "typical". Or he may simply have an aversion to danger which is stronger than it should be. In contrast, the neurotic's fear is often largely independent of his assessment of the risk and his desire to avoid danger in a way which suggests a breakdown in the sort of rational processes that are supposed to underlie moral action.

2. Phobias, unlike cowardice, are in some sense "triggered off" in rather special circumstances, whereas manifestations of cowardice are often characterized by a kind of rationalization which is general in the sense that it could be applied to a variety of circumstances with "equal justice".

Both of these point in the same direction: the idea is that a manifestation of a phobia is not the outcome of rational processes in the way a manifestation of cowardice is. A coward, for instance, might be persuaded to cross a deep chasm on a rope bridge by adding safety lines and the like, whereas this would presumably not affect the agoraphobe's response at all. In general, it seems that altering the coward's assessment of the danger or of his own (relevant) capacities typically alters his behavior in the usual way. By contrast, phobic behavior seems to be largely independent of the agent's beliefs concerning the danger and his own capacities.10 This, I think, is what Fingarette is getting at when he describes insane conduct as a failure to respond relevantly to essential relevance. (Fingarette, pp. 186-9) The addition of safety lines is essentially relevant to the risk involved in crossing a chasm on a rope bridge, and relevant responses are, for example, giving it a try, arguing that it is still too dangerous, alleging some fault in the safety ropes, and the like. In contrast, there is a sense in which, however much an agoraphobe may reason about action involving crossing rope bridges

10 Of course, if we could somehow convince the agoraphobe that there is no chasm there at all, we could get him to cross. Still, there is a fairly sharp line between the sort of "information" that will influence the coward and the sort of "information" that will influence the phobist. A good rule of thumb is: nothing short of misinforming the phobist in such a way that he would have a standard ignorance excuse for the action contemplated will influence him significantly. In short, he must be deceived as to what is really going on.
and the like, his reasoning can never be practical reasoning: the usual bridge between conclusion and action is just not there in such cases.

Actually, I think serious questions might be raised as to whether phobic conduct ever really is "non-rational" in this way, but let us put that matter aside for now, and ask this instead: Supposing someone did satisfy this description, would that preclude culpability? 11

The thesis we want to consider, then, is this: a person is not culpable on account of conduct fully characteristic of an incapacity for rational action. An incapacity for rational conduct must not be thought analogous to an inability to swim or add. If the thesis is to bear on our question at all, we must think of an incapacity for rational conduct as a behavioral disposition characteristic manifestations of which cannot be explained by or made intelligible in the light of the agent's reasons. We will need a short means of referring to conduct which is fully characteristic of an incapacity for rational action: I will call this sort of conduct l-conduct ('I' for 'irrational'). Using this abbreviation we may express the thesis thus:

\[ \text{(F) An agent cannot be culpable on account of l-conduct.} \]

We can understand (F) in a relatively narrow or a relatively wide sense depending on what we are prepared to count as l-conduct. Fingarette seems to include both of the following: (i) Cases in which the conduct cannot be rationalized in the usual way, i.e., conduct which cannot be given the usual kind of explanation in terms of the agent's reasons (beliefs and desires). The phobist fits nicely. Indeed, on the assumption that his wants and beliefs (or whatever it is that standardly rationalizes normal action) can duplicate those of the normal agent (whether cowardly or courageous), conduct fully characteristic of a phobia gives us a paradigm of l-conduct. (ii) Cases in which the conduct can be rationalized in the usual way, but in which the agent's reasons are not

11 Fingarette thinks it would. He defines "insanity" as follows: "Insanity is failure to respond relevantly to what is essentially relevant by virtue of a grave defect in the capacity to do so inherent at least for the time in the person's mental make-up." (p. 203) He continues: "Since 1 and 2 [2 is an expanded version of 1] make it plain that where there is insanity there is inherent incapacity to respond relevantly, it is evident that: 3. So far as there is insanity there cannot be responsibility." Fingarette seems to regard his position as self-evident. The nearest he comes to supportive argument is what appears to be a play on two senses of "responsible" (reliable vs. culpable) together with a version of the retributivist move mentioned earlier. See pp. 201-2. M. Moore asserts a similar position without argument in B. Brody and T. Engelhardt (eds.), Mental Illness and Its Policy Implications (D. Reidel), forthcoming.
themselves rational. Here we have cases like M'Naughton and Hadfield in which some or all of those beliefs and desires which figure essentially in the agent’s reasons are themselves more or less independent of rational processes; inexplicable themselves by appeal to other beliefs and desires, and not responsive in the usual way to changes in the rest of the agent’s cognitive and conative structure.

I propose to leave aside cases of type (iii), for these are cases that yield (often enough) to the Conservative strategy, leaving only the separable question whether the failure to know or want better is itself culpable. Since the beliefs and desires in question are supposed not to be rationalizable, it seems best to regard this left-over question as falling in a special way under (i). So let us understand (F) as if it concerned only type (i) cases.12

So far as I know, the literature contains no substantial defense of (F). If we are to get on, I think we must risk the charge of constructing a straw man and attempt to formulate the argument ourselves. Debate about cases is not likely to help until we identify the thinking which underlies (F) and gives it life.

It seems to me that the central idea underlying (F) must be that attempts to assess an agent in the light of I-conduct are baffled for the same reason that inferences from offense to character are baffled by excuses: the offense bears no morally relevant relation to the agent’s reasons for doing what he did. In such cases we cannot explain why the agent did what he did by appeal to his reasons. We can explain why the offense was committed, but not, in the ordinary sense, why the agent committed it. This is what makes it seem right to characterize both I-conduct and excused offenses as “extraneous” to the perpetrator of the offense considered as a moral agent. For if the offense is not the outcome of practical reasoning or anything like it, then it is not and cannot be a reflection on the agent’s morality — on his moral principles.

Of course, I-conduct and excused offenses may be quite representative, quite characteristic or typical of the agent. A normal person might, on a particular occasion, have an excuse for doing what he usually does, and a phobist might be a coward as well. The point is rather that I-conduct is like excused conduct in that the fit, if any, between the offense and the agent’s morality is coincidental. One can no more cite a phobic response as evidence of cowardice than one can cite

12 Another reason for wanting to treat type (ii) cases separately is that there are cases of this type that certainly do not preclude culpability, e.g., the second case imagined by the Justices in their reply to the Lords’ fourth question in connection with M’Naughton’s case in which A kills B under the delusion that B has injured his reputation.
unwitting inaccuracy as evidence of mendacity. I-conduct may or may not be intended, desired, or believed right and proper, but none of this has anything like its usual relevance.13

As I see it, the Liberal and the Conservative share the general view that to judge someone culpable on account of a certain offense is to judge that the offense reveals a moral defect in the agent. The Conservative tends to see the agent as a bundle of dispositions, some good, some bad. He admits excuses because he realizes that the moral character of conduct does not always reflect the moral character of the dispositions manifested. The Liberal, on the other hand, tends to see the agent as a bundle of moral reasons, some good, some bad. He is prepared to judge an agent culpable on account of an offense only when the moral character of the offense reflects the moral character of the agent’s reasons. Thus, very crudely, the difference is that the Liberal thinks of a bad person as someone with morally bad reasons for what he does, whereas the Conservative thinks of a bad person as someone disposed to do bad things.14 And the Liberal’s argument is simply that it makes no sense to ask of just any behavioral disposition whether it is morally good or bad. What we can assess morally are an agent’s moral principles, perhaps some of his desires, and, if these are independent matters, his intentions and aims.

The argument for (F), then, is this: conduct which does not reflect the agent’s reasons does not support a moral assessment of the agent, for the only sources of conduct it makes sense to assess morally are certain sorts of reasons, e.g., moral principles, desires, aims. (F) provides a ground for the Liberal position in turn on the assumption that conduct fully characteristic of a mental disorder is, at least sometimes, I-conduct. Conduct fully characteristic of a phobia is conduct fully characteristic of

---

13 Again, it is worth emphasizing that it is questionable whether conduct ever can really run against intention, desire, and belief in this way. I propose to let this pass, however, my purpose being to examine the Liberal position in its natural and most advantageous habitat.

14 This makes the Conservative sound like a utilitarian. There is certainly a kind of intellectual affinity. But the Conservative needn’t endorse utilitarian (or any other teleological) views concerning what constitutes “doing bad things”. Imagine someone who gets up each night, puts one bullet in a revolver, spins the chamber puts the puzzle to the forehead of his sleeping wife, and pulls the trigger. Even supposing nothing ever happens, and no one finds out, the Conservative is free to say of this case that such behavior reveals that the fellow is a villain of the worst sort. The teleologist, however, will have some difficulty explaining where the offense lies.
a defect in the agent, but the defect is not a moral defect, because it is not a defect in the agent's morality — i.e., his (moral) reasons.

A striking feature of the line of argument just summarized is that it supports something considerably stronger than (F):

(F*) An agent is not culpable on account of an offense that does not have its source in the agent's reasons (hereafter, "heedless offenses," or "heedless conduct").

For it seems clear that conduct that does not have its source in the agent’s reasons can no more reflect reasons than conduct that cannot have its source in the agent’s reasons. This is an unwelcome fact, for surely (F*) is too strong. Suppose D is engrossed in the Stanley Cup Finals. Without heeding what he does, he drops a cigarette butt over the railing of the upper deck into the buffant coiffure of P in the mezzanine below. Don’t we have an offense here that supports a negative moral assessment of D? It may even be fully characteristic of a trait of D’s character: "Just the sort of thing D would do," say those who know him. But does it have its source in morally bad reasons?

The defender of (F*) must argue either that D is not culpable or that his conduct does reflect a moral defect in his reasons. If he does the former, he plays into Conservative hands, so let us concentrate on the latter alternative. The Liberal might get his foot in the door as follows. What, exactly, is D’s offense? Not injuring a fan, surely, for though there is no doubt that D should have taken heed of what he was doing, the fact is that he did not. Hence he did not realize he was injuring a fan. (If he did realize this, but didn’t care, then his offense is injuring a fan, and it reflects a morally undesirable evaluation of injury to another. The reason he did what he did was just that he didn’t care much whether someone got hurt.) So D’s offense is not injuring a fan, but something like unjustifiably risking injury to a fan. If we ask why D did this, the answer must be that he did it because he undervalues risking injury to others, or something similar. This is exactly the sort of thing that must obtain if D is not to be exculpedated under (F*). Of course it might be that D did not realize that he was risking any such thing, or that D is simply incapable of monitoring his own conduct in the way we call heeding what one is doing. But in these cases we needn’t explain what moral defect in D is reflected by his heedlessness: none is reflected.

This line of defense has some plausibility. It does seem that conduct of the sort lawyers call recklessness or culpable negligence does argue a moral defect in the agent’s reasons in this sense: if we ask why the agent comported himself as he did, the right answer is generally that he doesn’t take seriously enough some likely consequence of his conduct or the risk of such a consequence. The trouble is that, in order to pursue
Robert Cummins

this line, we must construe 'r is x's reason for y' so broadly - so liberally - that the force of the underlying argument for (F*) is dissipated. If we stretch the concept of a reason for conduct far enough to cover the case of D lately rehearsed, we make it impossible to identify an agent's reasons for his conduct independently of identifying the moral defect manifested in that conduct. To see this, we need only suppose that D is like millions of others who heedlessly drop lighted cigarette butts as a matter of habit. On this supposition, we cannot assume that D under-values risk of injury to others. That's not why he did it; he did it simply because he has this bad (although extremely common) habit. This much is sufficient to settle the issue of culpability: we've identified a moral defect in Jones which was manifested in a characteristic way by his conduct. To insist that the bad habit is Jones' reason trivializes (F*); it stretches the concept of a reason for conduct so far that (F*) no longer constitutes an independent criterion of exculpation. Counting habits as reasons for conduct leaves us unclear as to why we should not count the agoraphobe's unreasonable fear as his reason for not crossing the bridge.

The Liberal's use of (F*) is likely to play into Conservative hands in any case, for it is simply not plausible to suppose that only the mentally disordered perpetuate offenses 'for no reason': Perhaps it is quite right to insist that the important point is whether the agent had the capacity to act rationally in the sort of situation in question, not whether he actually did act rationally. Since Liberals are not likely to thank me for an argument which undermines a distinction central to their position, it seems we must seek a different justification for (F). I will now try to show, however, that any attempt to motivate (F) independently of the unacceptable (F*) is bound to founder: there is no cogent reason for holding that an incapacity for rational control exculpates while a simple failure to exercise such control does not.

Why should capacity for rational control be a necessary condition of culpability given that actual exercise of such control is not? The only way I can think of to exploit the distinction between an incapacity and a simple failure is this: we cannot say of those lacking the capacity to exercise control that they should have exercised it. But it is difficult to see how this could help: what needs to be shown is that merely heedless conduct systematically reflects ill on the reasons of its authors while I-conduct does not. But the fact that someone should have taken heed doesn't allow us to treat him as if he had taken heed, hence does not reforge the lost connection between reasons and conduct. It does, of course, reinstate the basis for negative assessment, for a failure to take heed can be fully characteristic of a moral defect. But that defect need not be a reason. Once again, it might simply be a bad habit: some
people habitually pay no attention to what they are doing in certain circumstances.

I think we must conclude that (F) cannot be supported in a way that keeps it significantly different from the overly strong (F*). There is no denying, however, that (F) encapsulates a tempting and historically influential thought about culpability and how it is affected by mental disorder. Surely, one thinks, to act as a moral agent is paradigmatically to engage in practical reasoning, reasoning that begins with moral judgments and terminates in conduct. Someone who is incapable of practical reasoning in a certain sphere of conduct cannot act as a moral agent in that sphere, and so cannot be assessed morally on the basis of conduct in that sphere. We have just seen that this thought will not stand scrutiny. But it is one thing to see that a certain idea is mistaken, quite another to see why and how it goes wrong. Perhaps, then, it is worth taking a moment to try to defuse this idea in addition to refuting it.

(F) was offered in support of the doctrine that mental disorders are not moral defects. The strategy was to identify mental disorders (or some of them) with incapacities to rationalize conduct in a certain sphere. We thus go to the heart of the matter if we ask why an incapacity to rationalize conduct in a given sphere should not in itself count as a moral defect. The answer is supposed to be that only those who can rationalize their conduct in a sphere can express their moral qualities in that sphere. This is false. A person who manifests a morally bad habit expresses a morally bad quality in conduct that is not rationalized. What is true is that l-conduct, like habitual conduct, does not express moral reasons. The picture of moral action as the outcome of a practical syllogism leads us to overlook morally assessible qualities that are not reasons.

There is another reason why mental disorders are often held distinct from moral defects, viz., that people can acquire and alter such things as habits and character traits as a matter of deliberate policy, but not their phobias and the like. This view is also false. One can deliberately nurture a phobia. And while one typically needs help to get over a phobia, this is also true of many bad habits and most character traits.5

15 Exclusive attention to reasons as moral qualities tends to reinforce this mistake, for reasons are typically alterable by deliberation. Perhaps this also explains why delusions, which are reasons but are not typically alterable by deliberation, seem to many to contrast so sharply with what is paradigmatically morally assessible.
Robert Cummins

But the view is not only false, it is irrelevant. The fact that a bad habit or trait is not culpably acquired or maintained has no tendency to show that it is not a bad habit or trait. Vicious habits are like vicious dogs in that their viciousness does not depend on how they were acquired. I can excuse myself for having such a dog (though not for keeping it) by pointing out that I got it from a reputable trainer, but I cannot excuse its bites in that way. The idea that one can excuse conduct fully characteristic of a bad trait by offering an excuse for having the trait forces one to the absurd conclusion that an agent can be morally assessed only qua trait acquirer. No doubt this sort of assessment looms large in judging an agent over his lifetime or over a period of years. We engage in this sort of (diachronic) assessment at the end of someone’s career, or life, or childhood. But such assessments are relatively rare. Most often, we want to know what someone is like now; whether they are good or bad now. Here it is plainly which traits are reflected in conduct that counts, not how those traits were acquired. We may say of someone at the end of a trial, or at his funeral, ‘‘He couldn’t help being what he was.’’ But this would surely be a pointless remark if it weren’t fairly clear that what he was was pretty bad.

We began our investigation of the thesis that mental disorder can exculpate in its own right by looking for a systematic difference between cases in which conduct is fully characteristic of a mental disorder and cases in which conduct is fully characteristic of a vice (or any other uncontroversial moral defect). The suggestion we have been considering is the thesis that conduct fully characteristic of a mental disorder is not culpable when and because it is conduct fully characteristic of an incapacity for rational action in a certain sphere. Defeating this proposal does not, of course, refute the Liberal position. But we are now in a position to discern a general weakness: to establish his view, the Liberal must show that at least some mental disorders whose manifestations are (sometimes) offenses are not moral defects. One is inclined to feel it is backward and unenlightened to deny this, but I think this feeling is mainly due to a confusion. What is backward and unenlightened (perhaps inhumane) is treating the mentally disordered offender in the same way we treat the ordinary offender (which may be backward and unenlightened as well). Once we see that a policy of differential treatment is not prejudiced by allowing that some mental disorders may be moral defects, the feeling that such an allowance is backward and unenlightened largely evaporates.

Once this feeling does evaporate, it is surprising how open the question appears to be. It seems to me that the only substantial line we have on the concept of a moral defect in an agent is this: a moral defect is a tendency characteristically manifested in wrong conduct. Of course, conduct fully characteristic of a moral defect is not always a case of
wrong doing. But there would plainly be no point in calling a habit, say, a morally bad habit if characteristic manifestations were never or only atypically morally wrong.17 If this is right, then the Liberal begins at a disadvantage in this debate: phobias are no better off than cowardice in point of potential for undesirable conduct. This, indeed, is where we began in this section.

VII

I doubt that any further analysis will settle this question. Perhaps what we have here is rather one of those peculiar cases that arise from time to time in meta-ethics when analysis must be supplemented by a substantive moral decision if we are to have an answer to our query. In particular, we might still wish to say that the neurotic is not culpable on account of his failure to rescue the child, on the grounds that we should not - and this is a moral 'should' now - judge a person culpable on account of an offense which is fully characteristic of a neurotic fear. So while mental disorder is not an excuse in itself, perhaps it is wrong on moral grounds to regard it as, or act as if it were, compatible with culpability.

I can easily imagine reasons (though not very compelling ones) that might be given in support of the view that the courts should not judge such a person culpable. A court must determine what to do with the persons it tries, and neurotics plainly call for different treatment than cowards. A policy of radically different treatment might be difficult to sustain, and might be less effective in particular cases, if it is not supported by judgments attributing corresponding differences in culpability. The alternative would be to put the decision for psychiatric treatment on a par with the decision for probation or suspended sentence. Widespread public reaction against probation and suspended sentences — modes of treatment not supported by corresponding judgments of culpability, though generally well-supported on other grounds — indicates that the first alternative might be preferable.

One might argue on similar grounds that we, in our personal moral judgments, should not regard a person as culpable on account of an offense which manifests a mental disorder. For our judgments will affect

16 For more on this point see my "Could have done otherwise," forthcoming in The Personalist.
our attitudes, and our attitudes will be reflected in our actions. It seems clear that our actions vis-à-vis the neurotic should reflect something different in the way of attitudes from our actions vis-à-vis the coward. It may be useful to exhort a person whose actions do not reflect this difference to think of the coward only as culpable. On the other hand, to thus exhort someone who already responds appropriately strikes me as misguided: the only substantive moral issue here is one of appropriate response. To insist in addition to a yes-or-no answer to the question about culpability can only muddy the water. The neurotic is like the coward in not having an excuse, but unlike the coward in requiring special treatment. (And isn't this last really a matter of degree at most?) To force the issue of culpability here is bound to result in slighting one of these facts.

**VIII**

*The Legal Problem*

In the foregoing, I have been concerned to argue that the issue of moral culpability is moot in the case of the mentally disordered offender who has no "ordinary" excuse. If this is correct, then attempts to make progress on the issue of the legal liability of such offenders by insisting that moral culpability is a prerequisite of legal liability are doomed. This, I think, is a welcome result: as long as we see the issue as one of culpability, we are bound to focus on the trial as the exclusive locus for possible reform. This, I shall argue shortly, is a mistake.

Two factors underlie our inability to get a definitive result concerning the moral culpability of the mentally disordered offender who lacks a standard excuse. The first and most significant is the realization that a judgment of moral culpability does not license or require any special treatment of (action directed toward) the offender. This defuses the intuition that the mentally disordered require special treatment: however sound this intuition may be, we cannot argue from it to the Liberal conclusion on the grounds that the culpable offender may be treated in ways precluded by the presence of mental disorder. The second important factor is that our intuitions concerning what does and does not count as a moral defect come unstuck in the case of the mentally disordered offender. We can list things that will count as moral defects in normal agents and pass on suggested candidates with tolerable assurance. But our only substantial abstract line on the notion — hence the only thing that has a chance of bridging the gap between the normal agent
and the mentally disordered agent — is that moral defects are characteristicly manifested in wrong-doing. This line, important though it is, is evidently inadequate to settle the question whether serious mental disorders may be counted moral defects for purposes of assessing agents in the light of their offenses.

Shifting our attention to the legal liability of the mentally disordered, we can expect substantial relief from both these problems. To begin with, legislatures can and do enact laws licensing a variety of court actions vis-à-vis mentally disordered agents. Since the court must do something with the defendants who come under its purview — release, fine, incarcerate, commit, execute, etc. — the problem resolves itself into what the court should be permitted and/or required to do with mentally disordered defendants. I do not mean to suggest that this is an easy matter: I only want to point out that the legal context puts us back in business by reinstating the relevance of our intuitions concerning appropriate treatment of the mentally disordered. So our first problem drops out.

Our other problem — what, if anything, to count as moral defect in the mentally disordered agent — doesn't exactly drop out, but it undergoes radical alteration. While it is moot whether someone can be morally culpable on account of an offense fully characteristic of a mental disorder, no one doubts that mentally disordered offenders are and should be legally liable at least in this minimal sense: mental disorder does not and should not suffice, by itself, to force the court to direct unconditional release.17 If a mental disorder is not a moral defect, then one cannot be culpable on account of an offense fully characteristic of it. One evidently can be legally liable on account of such an offense,18 however, so the question is not "whether" but "when". If we think briefly of the factors underlying the legal liability of the mentally disordered — dangerousness to self or society are the typical ones — it emerges clearly that the legal context typically presents no question remotely comparable to the question whether mental disorders are or can be moral defects. Mental disorder frequently does — and sometimes should — license courts to abridge the freedom of defendan-

---

17 Henceforth, I shall use 'legally liable' in this minimal sense: a defendant is legally liable if the court is (should be) empowered to impose sanctions, or sanction-substitutes involving loss of freedom.

18 In the case of a mentally disordered defendant, the bearing of the offense on legal liability appears to differ significantly from the bearing of an offense on moral culpability and from the bearing of an offense on the legal liability of a normal agent.
ts, so there is no question that mental disorder does, sometimes, constitute a "legal defect", i.e., a condition sufficient to make a defendant legally liable.

Let us, then, cheerfully abandon the moot issue of the mentally disordered defendant's moral culpability and turn to the potentially more tractable issue of his legal liability. At first blush, this may seem naive optimism: surely no debate has generated so much heat and so little light as the debate over proper criteria of legal liability for the mentally disordered. My hope, however, is to short circuit this unproductive debate by arguing that the parties to it have been seeking criteria for the wrong thing. Once we appreciate the fact that the issue is not one of moral culpability, but rather one of liability to a variety of court ordered prescriptions, the traditional debate collapses for want of an issue.

There is no question that the traditional debate is predicated on the assumption that a defendant who is not morally culpable on account of the offense charged should not be legally liable. This assumption seems to have arisen mainly from the mistaken view that moral culpability and legal liability have analogous and comparable implications, either because moral culpability is supposed to render one liable to informal censure and sanctions analogous to punishment, or because the legal process is thought of primarily as a means of determining guilt or innocence rather than as a means of determining liability to sanctions and their substitutes. Neither of these views is defensible. The first has already been scouted, and the second evaporates when we remember that the courts are inevitably concerned primarily with determining what they should do. Questions of culpability are taken up because of, and in the light of, the bearing they are supposed to have on how the court should act. As Goldstein has emphasized,19 when the insanity defense is raised the issue is typically not whether to free the defendant, but whether to commit rather than incarcerate or execute. If we think of the legal process as primarily concerned with the question of culpability, the insanity defense will seem anomalous, for it will seem that the law allows for the enforced detention and treatment of exculpated defendants. The appearance of anomaly disappears once it is recognized that the legal process is designed to determine how the court should act. Many factors other than culpability, or even the outcome of a trial, are and should be brought to bear on this problem. So the question is this:

---

Culpability and Mental Disorder

how should the court act given a mentally disordered defendant, and how can these actions be implemented without undermining necessary legal safeguards? The very complexity of this question is welcome: at least we shall avoid the familiar small circles in favor of larger and more interesting ones.

IX

What Should the Court Do?

Let us begin with the question of what the court should do when confronted with a case involving mental disorder, leaving aside, for the moment, how it could or should be done so as to preserve legal safeguards.

This question is surprisingly simple to answer, at least in outline. Whether the court should direct special treatment does not depend on the mental condition of the defendant at the time of the alleged offense, but on his current condition. Doubtless defendant's condition at the time of the alleged offense, and even the fact (if established or stipulated) that he committed the offense in question, is evidence relevant to a determination of present condition. But it is only current condition that is relevant to current treatment. If it is established that the defendant's current condition is such as to require (or justify) commitment, the question of responsibility for the offense may simply be dropped as irrelevant to how the court should act. (Of course, finding out whether he really did it might be relevant to determining someone else's responsibility for the same or related offenses.)

It should be clear from this that the courts have no business directing indefinite commitments in the absence of powerful evidence of current disorder. This rather obvious fact, often flouted, has some rather powerful consequences.

1. If the current condition of the defendant is such as to justify commitment, holding a trial to determine responsibility or culpability is simply a waste of effort.

2. If the current condition of the defendant is not such as to justify commitment, then the only options are those available in the case of a normal defendant.
Robert Cummins

Thus it emerges that the presence or absence of mental disorder at the time of the offense has at best only a very indirect bearing on whether the court should commit: it comes in only as psychological evidence of current condition. If the current condition of the defendant does not justify commitment, then, and only then, does the question of culpability become relevant. Hence the bearing of defendant's mental condition on culpability is relevant only in a context in which the option of commitment has already been ruled out. If it is successfully argued that the defendant was culpable, once again, the court has none but the usual options (punishment, probation, etc.). If this is not successfully argued, the court has no option but unconditional release. A successful insanity defense should never lead to commitment, for it should never arise at all except in cases in which commitment has already been ruled out.

Once we accept (1) and (2) above, a very substantial simplification of the issues results (which is not to say that the remaining issues are simple, of course). We must still face the problem of how mental disorder bears on culpability, but some of the difficulty can be cleared away simply by setting the problem in its proper context. In this case, setting the issue in its proper context means knowing what options the court has available when the problem arises. The options the court should have available when the problem arises are just the options it has when there is no hint of mental disorder: release or normal sanctions.

Realizing this might seem to favor Conservatism: since the options are the usual ones, the excusing conditions should be the usual ones. But this would be a mistake; a mistake akin to defending Liberalism on the grounds that punishment of the mentally disordered is inappropriate. If the Liberal is right, then there are defendants who are not culpable and who, at the time of trial, are not candidates for commitment. If there are such defendants, they should surely be released. So it seems that the Law must face the Liberal/Conservative issue after all, and we found that issue to be moot. In the moral sphere, the right response to a moot issue is to withhold judgment, but the courts cannot do this, for they must do something, and they must make an effort to do it on principled grounds.

How shall this issue be decided? If we take the Conservative line, we run the risk of punishing victims we do not know to be guilty. Of course we do not know them to be innocent either. A familiar traditional principle might seem to apply here: when in doubt, release. Better risk releasing guilty parties than punishing innocent ones. But application of this principle to the case at hand assumes that there is a fact of the matter to be known; it assumes that either the Liberal is right or he isn't, though we can't tell which. This is at least debatable. I am inclined to think that what we have here is a class of possible defendants who are
neither culpable nor innocent. After all, what reason have we to suppose that everyone must be one or the other? The concept of culpability, we have seen, is tied to the concept of a moral defect. But that concept we found to have fuzzy edges: it just isn’t sharp enough to allow us to cut the class of possible defendants neatly in two with no remainder.

If this is right, then what we have is a case for legislation in the old utilitarian spirit. Since there is no justice or injustice to be done in this area, and since we have to do something, we may as well do some good if we can. Viewed in this light, it seems to me we should be Conservatives for the very simple and practical reason that we know how. We know how and why mental disorder is supposed to work exculpation according to the Conservative. We do not know anything comparable about cases not covered by the Conservative position. In practice, this means that we have no idea how to be Liberals; no idea, that is, how to put the Liberal position into effect.

Another way to put this same point is that we don’t know what class of defendants we are talking about. The best we can do is an open-ended negative characterization: we are talking about someone who has (i) committed an offense; (ii) was mentally disordered at the time of the offense but not presently committable; has no “standard” excuse. This is obviously not enough. Mental disorder can sometimes be irrelevant. No one holds that everyone satisfying (i) — (iii) is exculpated. But in the absence of a defensible account of how, lacking standard excuses, mental disorder can exculpate — in the absence, in short, of a viable defense of the Liberal position — all we can add is the useless and trivial (iv): the agent’s disorder was such as to exculpate him. If, as I have argued, there is no viable defense of the Liberal position (and no viable refutation either), then (iv) is the best we can do.

Recommendations

1 Commitment. The question of whether or not to commit depends on the current state of the person in question, and the issue of culpability is therefore irrelevant. The matter would therefore seem to be one for the civil courts, and should be referred thence by the criminal courts or other authorities whenever the facts warrant such an action. There is no point in pursuing the issue of a defendant’s culpability if commitment is a real option.20

2 Exculpation. If there is no question of commitment, the question of culpability arises and is to be settled along Conservative lines: mental disorder is relevant only as evidence of
Robert Cummins

a standard excusing condition (i.e., only as evidence of the sort of condition that excuses normal defendants, e.g., ignorance of fact, lack of requisite skills, etc.).

Unlike most reforms, these would be very easy to carry out, for nothing new is required. Their most attractive feature is that, if adopted, the intractible problem of “defining legal insanity” drops out of sight. It is replaced by two much more tractible problems: (1) whether defendant’s current condition justifies commitment, and (2) if not, whether defendant, as the result of mental disorder (or anything else), has an ordinary excuse for the alleged offense. The trick is accomplished by separating the question of culpability from the question of commitment, and leaving the latter to the civil courts where it belongs.

20 Actually, I believe that commitment should never be an option: law-breaking should be a necessary condition of any deprivation of liberty. Those so grossly disabled as to be unfit for trial cannot simply be abandoned, of course. But neither should such persons be committed unless we are also prepared to commit, e.g., severe burn victims.