

# THE GUILTY MIND

WILLIAM E. MANN

*University of Vermont*

**Abstract.** The doctrine of mens rea can be expressed in this way: MRP: If A is culpable for performing  $\varphi$ , then A performs  $\varphi$  intentionally in circumstances in which it is impermissible to perform  $\varphi$ . The Sermon on the Mount suggests the following principle: SMP: If A intends to perform  $\varphi$  in circumstances in which it would be impermissible for A to perform  $\varphi$ , then A's intending to perform  $\varphi$  makes A as culpable as A would be were A to perform  $\varphi$ . MRP and SMP are principles representative of intentionalism, a family of views that emphasizes the importance of intention to judgments about culpability. This essay examines an intentionalist's defense of MRP with respect to lying, strict criminal liability, and the distinction between intention and foreseeability, along with a defense of SMP with respect to failed attempts, and self-defense.

Consider two passages that are the focus of this paper.

When Jesus, in the Sermon on the Mount, turns his attention to the commandment against adultery, he amplifies it by saying that “everyone who looks at a woman with lust has already committed adultery with her in his heart” (Matt. 5:28).

It is a time-honored principle in the criminal law that *actus non facit reum nisi mens sit rea*, an act does not make [its agent] guilty unless the mind be guilty. Call this the *mens rea* principle.

Suppose one interprets “with lust” in Jesus’ pronouncement as signaling the presence of an intention to commit adultery. Then, even if the lustful agent were never to act on the intention (for lack of opportunity, say), he would not escape a charge of guilt; he would have a guilty mind. Suppose further that Jesus’ pronouncement generalizes, that for any kind of wrongdoing,  $\varphi$ , to intend to  $\varphi$  is already to do something for which one is culpable. Intention can be sufficient for culpability.

According to the *mens rea* principle, a person is not legally liable for her action if she did not intend to do what she did. (Cases of criminal negligence and reckless behavior provide counterexamples to the principle.)

No harm will be done, however, if we set them aside.) The *mens rea* principle has its roots in common-sense morality. Your believing that I trod clumsily but accidentally on your sore foot would provoke one set of reactions; your believing that I did it intentionally, quite another. Intention can be necessary for culpability.

Together, then, the Sermon on the Mount and the *mens rea* principle make a strong case for the importance of intention to judgments of culpability. But importance comes in degrees, as do notions of intention. As a result there can be different versions of *intentionalism*, a family of views that lays stress on the significance of intention in assessments of culpability. In what follows I shall present a robust version of intentionalism championed by a philosopher, let us call him Aurel, who takes both the *mens rea* principle and the Sermon on the Mount seriously. A few preliminary remarks are in order to describe the general contours of Aurel's position.

Sometimes to say that a person acted intentionally is to say nothing more than that she did not act accidentally.<sup>1</sup> Aurel's notion of acting intentionally is more robust than that. To act intentionally is to act as the result of a deliberative exercise that takes beliefs and desires as inputs and yields a decision as output. Thus Aurel's conception of acting intentionally supports the following conditional concerning intending to  $\phi$ : If A intends to  $\phi$  then A will  $\phi$  if the opportunity arises.<sup>2</sup> Aurel is willing to concede that phenomenologically the decision-making exercise can seem to the agent to take place instantaneously. What is important to him is a kind of logical, not temporal, priority, namely, that the analysis of acting intentionally requires specification of beliefs and desires. For this reason, if we suppose that to act willingly or voluntarily is simply to do what one wants to be doing, Aurel distinguishes acting willingly from acting intentionally. Many actions are performed both willingly and intentionally, but not all are. I surrender my wallet to the mugger intentionally but not willingly. The nicotine-depleted smoker might light up willingly but nonetheless "absent-mindedly."

---

<sup>1</sup> Cf. T. M. Scanlon, "Intention and Permissibility," *The Aristotelian Society*, Supplementary Volume 74 (2000): 306.

<sup>2</sup> Further refinement of the conditional would build in a requirement that A believe that the opportunity is not immensely unlikely. I cannot seriously intend to win the Powerball lottery.

With this understanding of intention as resolving-to-do-should-opportunity-arise, Aurel understands the *mens rea* principle and Matthew 5:28 (generalized) in the following ways, respectively,

**MRP** If A is culpable for performing  $\phi$ , then A performs  $\phi$  intentionally in circumstances in which it is impermissible to perform  $\phi$ .

**SMP** If A intends to perform  $\phi$  in circumstances in which it would be impermissible for A to perform  $\phi$ , then A's intending to perform  $\phi$  makes A as culpable as A would be were A to perform  $\phi$ .

Aurel takes **MRP** and **SMP** to be principles concerning the appraisal of an agent's performance, not a principle about the rightness or wrongness of actions performed. The two principles are silent about what makes it impermissible for A to perform  $\phi$ . If it helps, you can imagine that acts are impermissible if they fail to maximize happiness, or are not validated by the Categorical Imperative. Aurel is apt to favor the view that acts are impermissible if they contravene a divine command, but that is a topic for another paper.

In order to get a full sense of the robustness of Aurel's intentionalism, we need to see how he uses it to criticize moral and legal practices that deviate from adherence to **MRP** and **SMP**. In what follows I present three kinds of case that test Aurel's allegiance to **MRP**, namely, Lying, Strict Criminal Liability, and Intention and Foreseeability. I will then present two kinds of case to which **SMP** has application.

## LYING

The first test for Aurel's intentionalism is provided by the phenomenon of lying. When asked to say what a lie is, most people will agree that a lie is a false statement that the speaker believes to be false. Both conditions seem necessary. In telling you the truth that St. Paul is the capital of Minnesota I have not lied to you even if I believe that Minnesota's capital is Minneapolis. And if I sincerely tell you that Brasilia is the capital of Brazil, when in fact Brazil's capital is São Paulo, although I might misinform you, I have not lied. Further reflection will lead many people to think that the two conditions are not sufficient. That is, there

are cases in which A tells B something that is false and that A believes to be false, but in which A has not thereby lied to B. Examples: A tells B a joke—"So, space aliens abducted Dick Cheney last week . . ."; A recites the line, "O, what a rogue and peasant slave am I!" while auditioning for a production of *Hamlet*. What these cases suggest is a third condition, something to the effect that a lie has to be told in a context in which truth-telling is the norm (and clearly the telling of jokes and the reciting of scripts are not such contexts).

Aurel is happy enough to accept these three conditions, but, given his intentionalism, he does not regard them as jointly sufficient. Consider the following cases:

**Devotion:** Francesca protests Paolo's innocence to Paolo's brother and her husband, Gianciotto. In fact, Francesca has been sleeping with Paolo. Moreover, Francesca knows that Gianciotto knows about the couple's infidelity; thus that her protestation does not fool Gianciotto as to the facts. The context of the announcement is a context in which truth-telling is the norm. The point of Francesca's protestation is to give Paolo enough time to escape.

**Not Guilty:** Grimesby pleads Not Guilty at his arraignment before the judge, even though Grimesby knows that he committed the crime. Because his participation in the crime was recorded on videotape that has been aired repeatedly on the local television stations, Grimesby does not imagine that the judge or anyone else in the court believes that he is not guilty. He simply takes his plea to be a step necessary to obtaining a trial.

By Aurel's lights neither Francesca nor Grimesby has lied. Why not? Because neither had an intention to deceive anyone. According to Aurel, a fourth condition necessary to pin down the notion of lying is that the speaker must intend to deceive at least some members of the audience by means of what the speaker has said.<sup>3</sup> Now Aurel's addition will not be allowed to pass muster without an inspection. Some might try to handle **Devotion** and **Not Guilty** without the addition, either by claiming that a lie has been made, irrespective of the agent's intention, or by claiming that the context of utterance is not a context governed by a norm of

---

<sup>3</sup> For a recent endorsement of this condition, see Bernard Williams, *Truth and Truthfulness: An Essay in Genealogy* (Princeton: Princeton University Press, 2002), 96.

truth-telling. The latter tactic might be deployed with regard to **Not Guilty**. “A judicial arraignment,” a critic might allege, “is just not a setting in which one expects a defendant to speak the truth. Grimesby’s plea is not a declaration having a truth value. It is more like a password that he must utter to proceed to the next stage of the legal process.” Notice that Aurel *can* agree that arraignments are exceptions to the assumption of truth-telling. But he *need* not agree. Is the critic making an empirical observation about people’s attitudes towards arraignments? Or is the critic’s claim rather that there is something constitutive about arraignments that exempts them from the norm? And what will the critic say of the case of the defendant who pleads Guilty? These are questions for the critic to answer, not Aurel. For all that the critic has said, Aurel can and does maintain that *whether or not* arraignments presuppose truth-telling, Grimesby has not lied, because he lacked an intention to deceive.

The critic’s tactic is less plausible in **Devotion**. There is a norm of truth-telling when the audience is one’s spouse. So let us consider a critic who says that contrary to Aurel’s claim, Francesca lied. This critic says: “I grant you that Francesca’s primary aim was to buy time for Paolo, but her means of doing that consisted in telling a lie, directly about Paolo and indirectly about her own involvement with Paolo. The context presupposed truthfulness, and Francesca exploited that presupposition in her utterances.” Aurel can acquiesce in much of what the critic says, except, of course, for the assertion that Francesca lied. To be sure, depending on how we embellish the story, it might be that Francesca’s protestation is in service of a grander plot to assassinate Gianciotto. There is nothing amiss about inquiring skeptically into Francesca’s motivation. But, Aurel will insist, it is one thing to impute a base motive to an action, and another thing to classify an action as a lie. It is tempting to make the inference from “basely motivated linguistic performance” to “lie.” We should resist the temptation.

Aurel supposes, then, that **Devotion** and **Not Guilty**, properly understood, count in favor of requiring an intent to deceive clause in the characterization of a lie. But might there not be other cases of verbal actions in which it is undeniable that the agent lied even though the agent had no intention to deceive? Thomas Carson has designed two such cases to show that the intent to deceive condition is not essential to lying. Let us call the first one **Craven Witness**:

Suppose that I witness a crime and clearly see that a particular individual committed the crime. Later, the same person is accused of the crime and, as a witness in court, I am asked whether or not I saw the defendant commit the crime. I make the false statement that I did not see the defendant commit the crime, for fear of being harmed or killed by him. It does not necessarily follow that I intend that my false statements deceive anyone. (I might hope that no one believes my testimony and that he is convicted in spite of it.) Deceiving the jury is not a means to preserving my life. Giving false testimony is necessary to save my life, but deceiving others is not; the deception is merely an unintended “side effect.” I do not intend to deceive the jury in this case, but it seems clear that my false testimony would constitute a lie.<sup>4</sup>

The second one we can call **Cheating Student**:

Suppose that a college Dean is cowed whenever he fears that someone *might* threaten a law suit and has a firm, but unofficial, policy of never upholding a professor’s charge that a student cheated on an exam unless the student confesses in writing to having cheated. The Dean is very cynical about this and believes that students are guilty *whenever* they are charged. A student is caught in the act of cheating on an exam by copying from a crib sheet. The professor fails the student for the course and the student appeals the professor’s decision to the Dean who has the ultimate authority to assign the grade. The student is privy to information about the Dean’s *de facto* policy and, when called before the Dean, he (the student) affirms that he didn’t cheat on the exam. . . . The student says this on the record in an official proceeding and thereby warrants the truth of statements he knows to be false. He intends to avoid punishment by doing this. He may have no intention of deceiving the Dean that he didn’t cheat.<sup>5</sup>

Consider first **Craven Witness**. Carson claims that the lying witness does not intend to deceive *the jury*. That is not enough to show that intent to deceive is inessential to lying. Suppose that Riff and Bernardo are on their way to the local numbers parlor to lay down a few illegal bets. They are accosted by Officer Krupke, who wants to know what they are doing. Riff tells Krupke that he and Bernardo are going to church. Riff lies, intending to deceive Krupke, not Bernardo. It is no part of a plausible intent to deceive condition that one must intend to deceive

<sup>4</sup> Thomas L. Carson, “The Definition of Lying,” *Noûs* 40 (2006): 289.

<sup>5</sup> Carson, “The Definition of Lying”, 290.

*every* member of one's audience. From the way in which **Craven Witness** is described, it appears that there is at least one person who the witness intends to deceive about what he saw, namely, the defendant. Let us make it explicit, then, on Carson's behalf, that the witness knows that the defendant knows that the witness saw the defendant commit the crime. Aurel can still demur from the claim that the witness does not intend to deceive the jury, by invoking the distinction between acting intentionally and acting willingly. Aurel can claim that it is more accurate to say that the witness does not willingly deceive the jury, but does so, nonetheless, intentionally.

Now on to **Cheating Student**. Carson claims about this case both that the student lies and that the student has no intention to deceive anyone. But notice the context in which the case is embedded. The student began a project of deception when he cheated on the exam, a project in service of the goal of getting a higher grade. He could have terminated the project by not appealing his professor's decision to the Dean. Aurel agrees that the student's warranted statement is a lie, but insists that it *is* intended to deceive, by furthering the deceptive project that began with the cheating. We can call the principle to which Aurel appeals the Furtherance of Deception Principle:

**FDP** If a statement is made with the intention to further a project of deception, then the statement is made with the intention to deceive.

Note that "the intention to deceive" in the consequent of **FDP** need not be an intention regarding the content of the statement picked out in **FDP**'s antecedent. A *true* statement can satisfy **FDP**: deception need not confine itself to the false. It might be that A's long-range intention is to lie to B about some sordid episode in A's past. In order to gain B's trust, and thus to set B up for accepting the falsehood, A reveals to B a whole series of *truths* about A's past. The true revelations that A thus makes qualify under **FDP** as statements made with the intention to deceive. Of course the intention to deceive in the case of the true revelations is not directed at the content of A's statements. Here it is instructive to compare Aurel with Bernard Williams. When Williams defines the notion of a lie, he says that it is "an assertion, the content of which the speaker believes to be false, which is made

with the intention to deceive the hearer with regard to that content.<sup>6</sup> Let us adopt Williams's idiom, but add the requirements that the speaker's assertion must *be* false and be made in a context in which truth-telling is the norm. Finally, let us blend in the upshot of **FDP**. Aurel can then say that

**LIE** A lie is an assertion made in a context in which truth-telling is the norm, whose content is false, which the speaker believes to be false, and which the speaker makes with the intention either to deceive at least one of the hearers with regard to that content or to further a project of deception undertaken by the speaker.

According to Carson, **Cheating Student** is a case of lying without intent to deceive. By Aurel's lights, intentional deception is still implicated essentially. I suggest that, at a minimum, **LIE** is a serious contender for an adequate definition of lying. Given the prominent role that it assigns to intention, **LIE** allows for the creation of a special case of **MRP**:

**MRP: Lie** If A is culpable for lying, then A makes an assertion in a context in which truth-telling is the norm, whose content is false, which A believes to be false, and which A makes with the intention either to deceive at least one of the hearers with regard to that content or to further a project of deception undertaken by A, in circumstances in which it is impermissible for A to make the assertion.

Finally, we should note that **LIE** is an attempt to depict what a lie *is*. It is aimed at determining whether a person in a particular situation has lied. **LIE** is silent on the normative question raised by the last clause of **MRP: Lie**, namely, under what circumstances lying is impermissible. It is open to Aurel, for example, to take the hard line that there are no circumstances in which lying is permissible. But for now let us leave him content to argue for the correctness of **LIE** and **MRP: Lie**.

---

<sup>6</sup> Williams, *Truth and Truthfulness*, 96.



## STRICT CRIMINAL LIABILITY

Imagine the following scenario: Perry the lawyer has been approached by Cheatley, who wants Perry to defend him against a charge of murder. “Who is the victim?” Perry asks.

“Somebody named Vicky. I forget her last name; I never met her.”

“Then how did you kill—let me rephrase that—how is it that Vicky met her demise?”

“It began when she opened a mailing I sent...”

“Oh, dear, a letter bomb. Courts take a dim view of people who send letter bombs. We’d better go for an insanity defense.”

“No, that’s not what happened at all. Vicky was an agent for the Internal Revenue Service. When she opened my tax return, she realized quickly that it was fraudulent. That realization gave her a fatal heart attack. So I’ve been charged with her murder.”

The scenario is not as farfetched as one might imagine. First, filing a fraudulent tax return is a felony. So Cheatley was engaged in felonious activity when he filed his return. Second, Cheatley’s filing the return brought about Vicky’s death: had he not filed the return, she would not have had her heart attack.<sup>7</sup> Were Perry to take the case, he would argue that there is a huge gap between intending to defraud the government and intending to kill an IRS agent, claiming truthfully that Cheatley did harbor the first intention but not the second, and further, that had Cheatley known that his fraud would result in someone’s death, he would have abandoned his intention to defraud.

Depending on the venue in which the case is to be tried, Perry will find his argument facing greater or less resistance. Many states in the United States have *felony-murder laws*, laws which maintain that if someone is killed in the course of an attempt to commit a felony or to flee from a felony, all accomplices to the felony can be charged with murder. Felony-murder laws are the most dramatic members of a class of *strict criminal liability laws*, items of legislation that criminalize certain kinds of behavior in certain circumstances irrespective of the agent’s intentions. In both wording and precedent, different states apply the notion

---

<sup>7</sup> I do not believe that the truth of the counterfactual conditional is sufficient to impute causal agency. Holding all other factors constant, Vicky might have had her heart attack whether or not Cheatley filed his fraudulent return.

of a felony-murder differently. But would any state sustain a conviction of Cheatley on charges of first-degree homicide?

Probably not. But there are actual cases that come surprisingly—some would say disturbingly—close to Cheatley’s. In *People v. Hickman*<sup>8</sup> the Illinois Supreme Court upheld a lower court’s finding that the defendant was guilty of murder as a result of an unsuccessful burglary attempt at a liquor warehouse. The police had the warehouse under surveillance at night. As Hickman and two accomplices exited the warehouse, the police closed in. The three burglars fled. In the ensuing confusion one police officer mistook an armed police detective for one of the burglars. The officer shot the detective, killing him. There was no evidence that Hickman was armed before, during, or after the burglary. It does not take excessive charity to suppose that had Hickman believed that his burglary would lead to someone’s death, he would have abandoned the burglary. The relevant part of Illinois’s felony-murder statute states that “(a) A person who kills an individual without lawful justification commits murder if, in performing the acts which cause the death ... (3) [h]e is attempting or committing a forcible felony other than voluntary manslaughter.”<sup>9</sup> One might be surprised to learn that Illinois regards burglary as a “forcible felony,” especially when compared to burglary’s more violent sibling, armed robbery. But surely even under Illinois’s expansive conception of a forcible felony, fraudulent tax filing does not qualify. Nonetheless, *People v. Hickman* suggests that *People v. Cheatley* might be closer to legal reality than we had hitherto expected.

It is easy to anticipate some of Aurel’s reaction to strict criminal liability laws in general and felony-murder laws in particular. Strict liability laws flout the *mens rea* principle by dispensing with the intentionality requirement. Felony-murder legislation abuses the well-entrenched notion of murder by promoting some palpably unintentional homicides to the rank of first-degree homicides. Aurel campaigns for a conception of murder that respects the *mens rea* principle:

---

<sup>8</sup> 59 Ill. 2d 89 (1974). Another salient case is *People v. Fuller*, 89 Cal. App. 3d 618 (1978).

<sup>9</sup> Ill. Rev. Stat. 1971, ch. 38, sec. 9-1(a)(3).

**MRP: Murder** If A is culpable for murder, then A kills some person intentionally in circumstances in which it is impermissible for A to kill that person.

There is more to be said than this. For an intentionalist like Aurel, felony-murder laws cry out for justification. So let us imagine a defender of their propriety offering the following rationale:

“The defense I offer is a defense that is relative to the basic legal structure of a society. Different civil societies exhibit different patterns of response to behavior they regard as criminal. These patterns typically embody any number of historical accidents. Once embodied, the patterns take on a life of their own. Entrenchment can set in, in the name of evenhandedness and predictability; think of the importance attached to *stare decisis*. Working against universal entrenchment are social pressures to dislodge legally ensconced practices now believed to be wrong. The result is that any human society’s system of criminal law is apt to be a medley of vagariously assembled doctrines. Otto von Bismarck is supposed to have said: ‘Laws are like sausages. It is better not to see them being made.’ The point is that like sausages, human laws *are* made by human beings with all their foibles.

“Yet all these different patterns are recognizable attempts to provide security for citizens by punishing behavior that threatens their safety. As such, much of the criminal law is directed against behavior that we all regard as seriously morally wrong. But sometimes legislators find themselves having to achieve a balance among competing goals, each of which has warrant. For example, American jurisdictions place a high premium on the presumption of innocence in criminal proceedings. The presumption places the burden of proof on the prosecution and, in criminal cases, the standard of proof is high, ‘beyond a reasonable doubt.’ Adherence to the presumption of innocence results in an asymmetrical, adversarial trial dynamic, in which the prosecution’s efforts are aimed at proving guilt while the defense’s task is to induce reasonable doubt, chiefly by rebutting parts of the prosecution’s evidence. Subscription to the presumption of innocence is optional: some other countries that have criminal judicial systems equally as just as the American system do not follow the principle or its attendant adversarial trial structure so assiduously. Although the presumption is optional, to

say this is not to say that it is capricious, misguided, or wrong. At tension with the presumption of innocence is the goal of public safety: Americans understandably want to keep the number of violent crimes at a minimum. An ideal system of criminal justice would be one that infallibly identified the intentions and motives of every defendant brought before it. But for obvious reasons that is humanly impossible. Humans don't have the time, the resources, or the expertise that would be required to search the souls of people. Yet that is what American prosecutors would seem to have to do, following the presumption of innocence, namely, prove beyond a reasonable doubt that the defendant had the requisite *mens rea*.

“What courts can do is assess behavior and other kinds of publicly accessible evidence. Some types of behavior are dangerous to others. In order to strike some sort of balance between public safety and the presumption of innocence, a legislature can adopt the policy that if a kind of behavior is felonious, can be shown to have been undertaken intentionally, and results in the death of another, then it is permissible to tax the agent of the behavior with a penalty that is the penalty he would have incurred had he unambiguously intended the other's death. The legislature need not be construed as adopting the patent fiction that the agent intended the death.<sup>10</sup> It can be viewed as declaring instead that the consequences of the agent's actions have exposed him to the same penalty he would have faced had he directly killed the other person.

“Felony-murder laws thus have two deterrent virtues. First, they discourage the commission of certain kinds of felonies by raising the stakes. They put would-be felons on notice that if things go

---

<sup>10</sup> Here is a sampling of some dubious legal fictions.

(F<sub>1</sub>) Whoever intends  $\phi$  also intends all the consequences of  $\phi$ .

(F<sub>2</sub>) Whoever intends  $\phi$ , knowing that  $\phi$  is wrong (or illegal, take your pick), intends all the bad consequences of  $\phi$ .

(F<sub>3</sub>) Whoever intends  $\phi$ , knowing that  $\phi$  is wrong (etc.), intends all the bad foreseeable consequences of  $\phi$ .

(F<sub>1</sub>) and (F<sub>2</sub>) are enough to impale Hickman; maybe (F<sub>3</sub>), too—it all depends on how much is packed into the notion of foreseeability. (Some consequences are more foreseeable than others. Cheatley cannot have been expected to have foreseen that his fraudulent tax return would bring about a heart attack, but maybe Hickman should have realized that his burglary would run the risk of someone's dying.)

awry with their felonious activity, they may find themselves facing the most serious penalties one can face. Second, they undercut the advantage the would-be felon might have thought he would have with the presumption of innocence. In cases of homicide to which felony-murder laws apply, prosecution no longer has the burden of proving the defendant's *mens rea*.

"I wish to emphasize that this defense of the legitimacy of felony-murder laws is provisional and contingent. Provisional, because it depends on empirical claims to the effect that felony-murder laws further the goal of public security by deterring people from certain kinds of crime. If those claims are false then the defense is punctured. Contingent, because there are different legal systems that get along fine without felony-murder laws. One cannot claim that they are necessarily woven into the fabric of an ideal legal structure. (The United Kingdom abolished them in 1957 and hasn't slid into chaos as a result.) Do they violate **MRP: Murder**? It is clear that they allow for the possibility that people can be held *legally* culpable for murder even when they are not morally culpable. But the boundaries of the province of legal culpability are drawn by a society's legislature: they are no more grounded in the nature of things than is the boundary between Minnesota and Ontario."

By now one might be excused for wondering how Aurel will react to this defense. The answer is simple. The defense *is* Aurel's. All that Aurel needs to do is to add two sentences to it: "In contrast, **MRP: Murder** is not a matter of convention. Its conditions on culpability for murder are not subject to human legislation and thus cannot vary from one jurisdiction to another." Note that, as was the case with **MRP: Lie**, **MRP: Murder** does not pronounce on what kinds of intentional homicides, if any, are permissible.

### INTENTION AND FORESEEABILITY

Aurel's intentionalism would be flawed if its emphasis on what the agent intends were to result in an unsupportable exaggeration of the difference between intended consequences and merely foreseen consequences. Discussions of the role of intention in ethics frequently lead to discussions of the Doctrine of Double Effect, which in turn

frequently lead to a consideration of the distinction between intended and merely foreseen consequences. The doctrine is supposed to apply to choice situations in which the consequences are unavoidably mixed, some being good, some bad. The doctrine has received insightful and provocative attention recently in its application to choices made in medical ethics, in the conduct of combatants in war, and in how one should conduct oneself in an area plagued by distressingly many wayward trolleys.<sup>11</sup> There are several ways of formulating the doctrine, not all of them obviously equivalent, but I shall assume that the following version is representative:

**DDE** If A's performing  $\phi$  would unavoidably have both good consequences, **GC**, and bad consequences, **BC**, then it is permissible for A to perform  $\phi$  if and only if (1)  $\phi$  itself is not morally forbidden, (2) A does not intend to bring about **BC**, (3) **BC** is not causally necessary for **GC**, and (4) the badness of **BC** does not outweigh the goodness of **GC**.

Illustration: In the course of prosecuting a formally declared war against Bulgaria, the Commander in Chief of the armed forces of Ulceria has to decide whether to bomb Bulgaria's major munitions factory, thereby bringing the war to a speedier conclusion. The **DDE** tells the Commander that his action is permissible, even if the bombing results in loss of life of some noncombatant civilians, as long as the bombing itself is not a forbidden act, the Commander does not intend the death of the civilians, their death is not a causal means necessary for bringing about the speedier conclusion, and the good consequences of the speedier conclusion (say, in Bulgarian and Ulcerian lives spared) is not outweighed by the bad consequences of the Bulgarian civilian deaths.

Let us put pressure on clause (2) of **DDE**. Suppose that the Commander *knows* full well that bombing the munitions factory will kill several innocent noncombatants. (He knows this, we might suppose, because it is common knowledge that the Bulgians force kidnaped

---

<sup>11</sup> See Philippa Foot, "The Problem of Abortion and the Doctrine of Double Effect," *Oxford Review* 5 (1967): 5–15; Judith Jarvis Thomson, "The Trolley Problem," *Yale Law Journal* 94 (1985): 1395–1415; F. M. Kamm, *Morality, Mortality: Volume II: Rights, Duties, and Status* (New York: Oxford University Press, 1996); and Scanlon, "Intention and Permissibility."

Ulcerian citizens to work in the factory.) Is it not facetious in this circumstance to maintain that although he foresees that they will be killed if he bombs the factory, he does not intend their death? How, in this circumstance, could foreseeing their death not just be intending their death, accompanied, perhaps, by some self-deception? And if it is not the same thing, how could the difference be significant enough so that intending their death would make the bombing impermissible whereas merely foreseeing their death would not?<sup>12</sup> Put the question in the idiom of Aurel's robust conception of intention: how can the decision to bomb, based in part on the true, justified belief that the noncombatants will be killed, not be a decision to kill the noncombatants?

It may be that the pressure on clause (2) is generated by the thought that known consequences are intended consequences:

**KCIC** If A intends to perform  $\phi$  and knows that  $\psi$  is a consequence of A's performing  $\phi$ , then A also intends  $\psi$ .

If true, **KCIC** would deal a serious blow to **DDE** by collapsing the distinction between foreseeing a consequence and intending it. But **KCIC** does not appear to be true. Suppose that A can save the life of either B or C, but not both. Suppose further that A knows all the facts pertinent to her predicament. A lets a randomizing device, for example, a toss of a coin, determine whom she will save: heads, it is B; tails, C. Suppose the coin lands heads up. A now intends to save B, knowing that it follows that C must die. A does not intend that C die, even though she surely foresees that C's death is a consequence of her intentionally saving B.

I do not claim to have vindicated clause (2), much less **DDE**, for there may be other considerations independent of **KCIC** that allow pressure to continue to be applied to (2), and there are other quarrels one might have with the other clauses. But I do want to point out that an intentionalist of Aurel's stripe need not swear fealty to **DDE**, at least not in the way in which it is typically understood. Keeping with the spirit of **MRP**, Aurel can put forward something that looks like the Doctrine of Double Effect, but which nonetheless differs from it in significant ways:

---

<sup>12</sup> Scanlon reports that a similar case was suggested to him by Judith Jarvis Thomson; see "Intention and Permissibility," 304–305.

**MRP: Double Effect** If A's performing  $\varphi$  unavoidably had both good consequences, *GC*, and bad consequences, *BC*, then A is culpable for performing  $\varphi$  if and only if *either* (1)  $\varphi$  itself was morally forbidden, *or* (2) A intended to bring about *BC*, *or* (3) *BC* was causally necessary for *GC*, *or* (4) the badness of *BC* outweighs the goodness of *GC*.

Clause (2) shows that **MRP: Double Effect** does not dissolve the problem of distinguishing between what is intended and what is merely foreseen. But **MRP: Double Effect** is not simply a De Morgan transformation of **DDE**, specifying a logically equivalent principle in terms of forbiddenness instead of permissibility. The notion highlighted by **MRP: Double Effect** is culpability, not forbiddenness. Recall that the original **MRP** was put forward as a principle of agent performance appraisal, not as a principle about what is forbidden, obligatory, or permissible. The latter sort of principle is insensitive to the time at which it is applied to a particular case. **DDE**, for example, can be invoked prospectively, as a guide to action, or retrospectively, as a vehicle of moral criticism of the action committed. **MRP: Double Effect**, in contrast, has only a retrospective function—hence its formulation is in the past tense—namely, the assessment of the agent's performance after the fact. An omniscient judge could, one presumes, determine whether A merely foresaw or intended that *BC* would come about, but short of omniscience, **MRP: Double Effect** is no better off than **DDE** concerning the foreseen-intended distinction. Moreover, the fact that **MRP: Double Effect** is exclusively retrospective and exclusively agent-judgmental may seem to rob it of much of its interest.

In a while Aurel will have a chance to respond to this attempt at dismissal. For now, however, let us turn to cases relevant to principle **SMP**. So far we have looked at cases that attempt to belittle the importance of intention as a necessary condition for imputations of culpability. The next two cases, Failed Attempts, and Intentions and Motives, cast doubt on the sufficiency of a presence of intention for ascriptions of culpability, thus challenging principle **SMP**.



### FAILED ATTEMPTS

Here are two alternative hypothetical cases, one a “success,” the other a “failure.”

**Premeditation:** Alpha has coolly planned in advance to kill Beta, believing with good reason that Beta would soon blow the whistle on Alpha’s embezzling activities. Alpha raises his rifle, takes careful aim, and dispatches Beta with one shot.

**Mosquito:** Same scenario, two different actors. As Gamma peers down the barrel of his rifle aimed at Delta and begins to squeeze the trigger, a mosquito lands on his nose, causing his shot to go astray.

In **Premeditation** Alpha is guilty of first-degree homicide, something for which Alpha would presently stand to face the death penalty in thirty-eight out of fifty states in the United States. In **Mosquito** Gamma is guilty of attempted murder, a felony, to be sure, but one that exposes Gamma to a much less severe penalty. Why the disparity? Alpha and Gamma have exactly the same evil intentions and motives. Gamma did everything he could to further his quest, just as Alpha did. Why should an adventitious mosquito make such a big legal difference?

Skeptics about the importance of intentions will say that the disparity indicates that we care more about results. How much more? As a statistical exercise, one might try to compute the average sentence meted out to defendants found guilty of first-degree homicide (arbitrarily assigning sixty years, say, to life sentences and to the death penalty), along with the average sentence imposed on defendants found guilty only of attempted homicide. The first average divided by the second would then give us the *failed attempt discount index*, or FADI, which could serve as a measure of the extent to which a particular jurisdiction discounts failed attempts compared to successful attempts. Thus if the jurisdiction hands out, on average, fifty-year sentences for first-degree homicides and five-year sentences for attempted homicides, its FADI is ten. One could then compare the FADIs among different jurisdictions to see whether interesting patterns emerge, and to gauge the strength of the skeptics’ case for demoting intentions.

As sociologically interesting as the construction of FADIs might be, they would merely quantify the practice; they would not justify it. The normative question remains: why *should* we follow a schedule

of legal punishment in which Gamma's deed is treated more lightly than Alpha's? One might naturally seek to give a justification for the disparity by appealing to consequences. But while it is true that Alpha's act had worse consequences than Gamma's, it is hard to find a plausible systematic rationale for punishing Gamma less severely among the familiar consequentialistic dimensions of general deterrence, particular deterrence, incapacitation (or social quarantine), rehabilitation, and education.<sup>13</sup> Perhaps in recognition of the lack of justification, Section 5.05 of the American Law Institute's Model Penal Code recommends that

Except as otherwise provided in this Section, attempt, solicitation and conspiracy are crimes of the same grade and degree as the most serious offense which is attempted or solicited or is an object of the conspiracy. An attempt, solicitation or conspiracy to commit a [capital crime or a] felony of the first degree is a felony of the second degree.<sup>14</sup>

It appears that the recommendation is to treat attempted burglary as if it were burglary, attempted arson as if it were arson, and so on, with the exception of a capital crime or a felony of the first degree, which would surely encompass premeditated, intentional homicide. In this case, the failed attempt would be demoted to second-degree homicide, which would carry with it a lesser penalty.

By Aurel's lights the Model Penal Code has almost got it right. If the various consequentialistic justifications for downgrading failed attempts are themselves unpersuasive, then there appears to be no provisional or contingent argument, akin to the rationale Aurel offered for felony-murder laws, that justifies treating failed attempts less severely than successful attempts. In short, the FADI for any jurisdiction should be one, even for failed attempts at homicide. We can imagine Aurel saying, with rhetorical zeal, that Gamma is as guilty of murder as Alpha is, thereby giving a forceful rendition of the phrase "guilty for all intents

---

<sup>13</sup> For criticism of these sorts of attempts at justification, see David Lewis, "The Punishment That Leaves Something to Chance," in *Papers in Ethics and Social Philosophy* (Cambridge: Cambridge University Press, 2000), 227–243.

<sup>14</sup> American Law Institute, Model Penal Code (1962), as reprinted in Wayne R. LaFare, *Modern Criminal Law: Cases, Comments and Questions* (St. Paul: West Publishing Co., 1978), 741 (brackets in original).

and purposes.”The fact that Delta did not die because of Gamma’s effort is immaterial to the assessment of Gamma’s culpability. Aurel subscribes to a close relative of **SMP**:

**SMP: Failed Evil Attempts** If A attempts but fails to perform  $\varphi$  in circumstances in which it is impermissible for A to perform  $\varphi$ , then A’s attempting to perform  $\varphi$  makes A as culpable as A would have been had A succeeded in performing  $\varphi$ .<sup>15</sup>

It may have occurred to you by now that failure comes in at least two flavors. One can fail in the attempt to do something evil, but one can also fail in the attempt to do something good. Suppose that Alberta and Alberto individually intend to donate substantial amounts of money to a charity. For that purpose it happens that both of them entrust their money in accounts managed by Bezzle, the banker. Bezzle decides to drain one of the accounts and abscond with the funds. Bezzle flips a coin to determine which account to drain. It happens to be Alberto’s. As a consequence, Alberta’s money goes to the charity while Alberto’s money does not. If Alberta is praiseworthy then Alberto should be equally as praiseworthy. But Alberta’s existence in this case is, at bottom, irrelevant. It is Alberto who should be equally as praiseworthy whether or not he succeeds. The principle at work here is

**SMP: Failed Good Attempts** If A attempts but fails to perform  $\varphi$  in circumstances in which it is permissible for A to perform  $\varphi$ , then A’s attempting to perform  $\varphi$  makes A as praiseworthy as A would have been had A succeeded in performing  $\varphi$ .

### INTENTIONS AND MOTIVES: THE CASE OF SELF-DEFENSE

Intentions are one thing, motives another. **SMP** ignores the importance of motives to the assessment of character (and thus may do an injustice to the interpretation of Matt. 5:28). We can examine the difference between intentions and motives by looking at the legal plea of self-defense.

---

<sup>15</sup> Here and below let us understand “A attempts to  $\varphi$ ” to rule out “half-hearted” attempts.

Criminal law relies on a set of purely external, formalistic criteria by which it gauges putative cases eligible for a plea of self-defense. Chief among these criteria are the following:

*Imminence:* The harm threatened by the attacker must be temporally immediate and unavoidable. B's threat to attack A a week from next Tuesday does not justify A in attacking B today (unless, perhaps, A is already unjustly in B's clutches). In the eyes of the law, a preemptive strike is not self-defense.

*Proportionality:* The harm inflicted by the intended victim must be just sufficient to thwart the attacker and must not exceed in any case the harm the attacker would have inflicted. If B's proposed act of aggression is simply to mush a custard pie in A's face, A cannot plead self-defense for gunning down B.

*Reasonable Person Standard:* The intended victim is expected to have assessed the situation and responded as a reasonable person would in the victim's circumstances. A's prospects for a successful plea of self-defense evaporate if it turns out that A believed B to be a space alien. (In this case, perhaps an insanity plea would be more successful.)

Consider now two hypothetical cases. In both of them the agent, fully aware of the eligibility criteria for a successful plea of self-defense, contrives to be the intended victim of an attack, using the attack as the occasion for killing the attacker.

**Greed:** Avaritia stands to inherit Luger Shortfuse's considerable fortune. But Luger is young and fit, with prospects of living a long life. Obsessed with the prospects of wealth and knowing of Luger's penchant for violence, Avaritia intentionally provokes him into attacking her with a poker. Avaritia puts to use Luger's loaded pistol, which she had planted nearby, thereby quickly becoming a wealthy heiress.

**Tender-heartedness:** Mauser Shortfuse, Luger's father, is suffering horribly from a terminal disease for which palliation is ineffective. Mauser wants to be put out of his misery but does not want his death to be counted as a suicide. Nor does he want his kin to be put at risk of a charge of murder, were they to exercise conventional modes of euthanasia. His daughter, Mercy, who knows all this and is moved by nothing other than compassion for her father, ruefully but determinedly exploits Mauser's one family foible by goading Mauser into attacking her so that she can kill Mauser quickly.

One can refine these cases so that Avaritia and Mercy both pass the tests for a successful plea of self-defense. And one can easily imagine someone who reasons in this fashion: “I would not care to live next door to the Shortfuse family. Even so, I would harbor less cold feelings for Mercy than for Avaritia. My differential response does not appear to be based on a difference in their intentions, for they both act from the same sort of intention—to kill a family member while making it look like self-defense. Nor is it based on the facts that Luger had no desire to die while Mauser did, and that Avaritia deprived Luger of a longer, happier future than the future denied to Mauser by Mercy. Perhaps not all intentional killings are forbidden—let’s set aside for another day a discussion about killing in warfare and legally sanctioned executions—but these two killings are forbidden. Nor can I detect more evil remote intentions in Avaritia. Her remote intention is to inherit, to which her intention to kill is a means, while Mercy’s remote intention is to alleviate suffering. There is nothing wrong with an intention to inherit *per se*. So I’m left to conclude that what explains my antipathy towards Avaritia is the presence of greed in her, which is not present in Mercy. Greed is not an intention; back in the old days it was numbered among the seven deadly sins, an acquired vice that disposes its possessor to seek after more material goods than is necessary or fitting. So when it comes to agent performance appraisal, **SMP** can’t be the whole story.”

Aurel can acknowledge that whether it is acquired voluntarily or not, greed is the sort of character trait that ought to be resisted or stifled, precisely because it disposes its possessor to form evil intentions. But whatever culpability one has for the possession of a vice, it pales in comparison to the culpability one has for intentionally exercising it. A vice is a bit like Luger’s loaded and unattended pistol—merely dangerous in itself but calamitous when used to evil purposes. The best way, perhaps the only way, to resist a vice like greed is the same way one tries to break any bad habit—by avoiding situations that encourage the exercise of the vice and by resolving not to acquiesce in the vice when the situations are unavoidable. One may not succeed, but there seems to be no other way of even trying. The principle that applies here, by Aurel’s lights, is this:

**SMP:Vice** If A attempts but fails to avoid performing  $\varphi$  in circumstances in which it is impermissible for A to perform  $\varphi$ , because A's vice-driven desire to perform  $\varphi$  overpowers his attempt to avoid performing  $\varphi$ , then A is less culpable than A would have been had A not (even) attempted to avoid performing  $\varphi$ .

Aurel also has a bone to pick with the formalistic criteria that help to define a successful plea of self-defense. That they can be manipulated, as in **Greed** and **Tender-heartedness**, is a symptom that they set the standard too low. More specifically, they are insensitive to the role that the agent's intention plays in these cases. Even more specifically, they function as surrogates to intention: instead of attempting to assay what the agent intended, courts can simply apply the criteria, thus allowing Avaritia and Mercy to slip through the cracks undetected.

Aurel's complaint here may strike one as unrealistically high-minded. We cannot read people's minds. Courts must make decisions by applying objective tests to publicly observable behavior, resisting the temptation to construct invidious and unverifiable hypotheses about the agent's intentions. Even if some sort of technology were developed that would allow authorities to eavesdrop on people's thoughts, we would regard that development with horror. Aurel himself is deeply suspicious about the wisdom, justice, and continued stability of any political authority. He is thus inclined to agree that there are many instances of wrongdoing for which government intervention would be worse than allowing the wrongdoing to go undetected, unapprehended, and unpunished. What else can one hope for?

## CONFESSIONS

When it comes to appraising the hidden springs behind a person's publicly observable performance, one can hope for an authority who is omniscient, infallible and incorruptible, perfectly just and perfectly merciful, who will sort out and set right the injustices in the world. That is the direction in which Aurel has been steering us.<sup>16</sup> And Aurel

---

<sup>16</sup> Note how **MRP: Double Effect**, even though exclusively retrospective, takes on importance from the point of view of a divine judge.

is not a fictional character. I take myself to have been channeling, in contemporary idiom, the thought of Aurelius Augustinus, known more familiarly as St. Augustine, supplemented in some cases by insights from Peter Abelard, who, I am prepared to argue, knew Augustine's moral thought very well.<sup>17</sup> A champion of human-made law would do well to attend to their arguments, but need not fear usurpation of power: Aurel's judge presides over a different jurisdiction.

---

<sup>17</sup> For my more historical excursions in these fields, see "Inner-Life Ethics," in *The Augustinian Tradition*, ed. Gareth B. Matthews (Berkeley: University of California Press, 1999), 140–165; "To Catch a Heretic: Augustine on Lying," *Faith and Philosophy* 20 (2003), 479–495; and "Ethics," in *The Cambridge Companion to Abelard*, ed. Jeffrey Brower and Kevin Guilfooy (Cambridge: Cambridge University Press, 2004), 279–304.