

GIDEON YAFFE

TRYING, ACTING AND ATTEMPTED CRIMES

I. INTRODUCTION

When the legislature defines a crime—when it specifies a punishment for a person shown beyond a reasonable doubt to have acted in a certain way in certain circumstances and with certain results—it succeeds in defining a second crime as well, namely the crime of attempting to commit the crime defined. The legislature could choose specifically not to proscribe the attempt to commit a crime defined, but in the absence of an explicit statement to that effect, by defining the crime the legislature grants the state the power to punish also for the attempt. Why do we have this practice? What problem are we solving by having a system in which it is automatically a crime to attempt almost every crime explicitly defined? It is natural to answer that we are solving a problem of inequality: since there is no difference in desert between the person who completes the crime and the person who tries but fails, and since there is often no other good reason to treat them differently, there ought not to be a difference in treatment by the state. A system that does not punish attempts to commit crimes treats equally deserving citizens differently and for no good reason.

To accept this justification is to make what many see as a salutary commitment to the denial of moral luck. Those who take there to be moral luck, in the sense in which the term is to be used here, hold that it is possible for two actions to differ morally in virtue of the fact that they differ in some respect that was entirely out of the control of the agent of

either action.¹ In fact, the intuitive thought that there can be a moral difference between an action, on the one hand, and a failed attempt to perform the very same action, on the other, is often cited as support for the contention that some moral intuitions point towards the existence of moral luck.² After all, often the difference between succeeding and merely trying is nothing but a difference in the air currents wafting through different parts of the casino. Yet, it seems intuitively that this difference can make a moral difference. So, it is no surprise that those who take the intuition to be flawed also think that the point of criminalizing attempts is, precisely, to correct for the influence of this flawed intuition in our legal system. People who take this line are also disturbed by the sentencing practices that we find in many jurisdictions, practices which seem to involve a commitment to moral luck. Many jurisdictions provide lesser penalties for attempted crimes than for completed; under the common law, for instance, an attempted felony is a misdemeanor.³ In response, a number of theorists,⁴ and the drafters of the *Model Penal Code*,⁵

¹ Thus, for the purposes of this paper, my concern is with moral luck as it applies to the assessment of actions, and not to the assessment of agents, except in so far as assessment of actions is a form of assessment of agents. My concern then is not with so called “constitutive” luck: luck that influences what sort of person an agent is. My concern is, however, with “consequential” luck and, to some degree, with “opportunity” luck. What one does often depends both on what consequences flow from what one does and on what opportunities one had to act. If either can make an action morally worse or better, even when the consequence or opportunity is out of the agent’s control, then there is moral luck.

² Cf. Thomas Nagel, ‘Moral Luck’ in *Mortal Questions*, Cambridge: Cambridge University Press, 1979, pp. 24–38.

³ States use various methods for calculating the sentence for an attempt, but most offer a lesser penalty for the attempt than for the completed crime. California, for instance, gives half the penalty of the completed crime for the attempted (see Cal. Penal Code §664). This remains the most common formula.

⁴ See, for instance, Sanford H. Kadish, ‘The Criminal Law and the Luck of the Draw’ in *Journal of Criminal Law and Criminology*, 84 (4), 1994, pp. 679–702; Stephen Morse, ‘Reasons, Results and Criminal Responsibility’ in *U. Ill. L. Rev.* 363, 2004, pp. 364–435.

⁵ *Model Penal Code*, Philadelphia: The American Law Institute, 1985, Part 1, commentary on §5.05, p. 490.

have claimed there to be no adequate justification for a discrepancy between sentences for attempted and completed crimes, and they have done so on the grounds that there is no moral luck and so often no difference in moral culpability between those who try and fail, on the one hand, and those who succeed in committing a crime, on the other. Such theorists accept the justification above for the proscription of attempts: such a proscription corrects for morally irrelevant luck and so moves us a step closer to a system in which equally deserving citizens are treated the same. This justification, they think, should also be reflected in our sentencing practices.

If we accept this justification for our practice of criminalizing attempts, then we ought not to allow the way in which we criminalize attempts to simply reintroduce the problem of inequality that criminalizing attempts was meant to solve. That is, we ought not to allow that there are those who do less than attempt a crime and who do not differ morally from those who attempt it, and who are for no good reason nonetheless treated differently by the state from those who attempt. If attempt law allows for this, then we would find that proscribing attempts solves a problem only by reintroducing the very same problem. And, in fact, it appears (although, as will be argued here, *merely* appears) that attempt law does allow for this in an objectionable way. The crime of attempt requires an act; there is no attempted crime if the defendant has not *done* something to further his criminal intention. But what act an agent performs is, often, at least partially a matter of luck. This point is often, and naturally, countered with the following platitude: an act on the part of the defendant is an essential element of attempt for *evidential reasons*. The thought is that acts, and nothing other than acts, constitute evidence of the thing that we really care to criminalize when we criminalize attempts, such as a criminal intention, disposition, or a resolute determination to act criminally. In what follows, it will be argued that it is very difficult to flesh out this response satisfactorily. It is very difficult, that is, to do both of the following two things: (1) deny moral luck in the way required to support the justification just described for the practice of criminalizing attempts, and (2) adequately

justify, along the lines suggested by the appeal to the evidential role of acts, the fact that the crime of attempt requires an act. Despite the difficulties, however, it will be argued that it can be done. By reflecting on the notion of trying to act, we learn that trying to commit a crime is something that can only be adequately evidenced when the defendant has performed some actions, actions that he may not have performed were it not for various events entirely out of his control.

II. THE PROBLEM

In virtually every jurisdiction in the United States, and under the *Model Penal Code*,⁶ to be guilty of an attempt to commit crime C a defendant must (1) intend to C^{7,8} and (2) take steps towards C-ing that (a) go beyond merely preparing to C and (b) go some distance towards completion of the crime (although how far one needs to go varies by jurisdiction). In the second of these two requirements, the law governing attempts reflects a pair of intuitive distinctions, both of which are inordinately difficult to analyze satisfactorily. To see them, imagine that I

⁶ See *Model Penal Code* §5.01(1).

⁷ I discuss this requirement in a companion to this paper, 'Trying, Intending and Attempted Crimes' in *Philosophical Topics*, 32 (1–2), 2004, pp. 505–532.

⁸ It is a non-trivial matter to specify the content of the intention required for attempt. It is false to suggest, as I seem to be in the main text in calling the relevant intention an "intention to C", that the relevant intention must be an intention to *fulfill all of the elements of the completed crime*. It is not obvious, to identify just one problem with such a view, that the circumstantial elements of the completed crime need to be in the content of the intention required for attempt. To mention another problem, it is obvious that the defendant need not intend *to commit a crime*, that is, to violate a law; just as the completed crime might require no awareness on the defendant's part that he is violating a law, the attempted crime might not require any such awareness, much less any such intent. In any event, since determining the content of the needed intention is not my task here, I will use a variety of possibly misleading phrases, such as "the intention to C", "the intention to act criminally", and the "intention to complete the crime", to refer to the required intention. All such phrases should be understood to be referring to that intention, whatever its content, that is properly an element of attempt.

intend to climb a staircase of 100 steps. I tie my shoes. I ascend the first stair, then the second, and so on. There is an intuitive difference between tying my shoes and ascending the first stair. The former act is merely preparatory while the latter is a *step* towards doing as I intend. There is also an intuitive distinction between ascending the first stair and ascending stairs one through fifty. The latter act, the act of ascending the first fifty stairs, is more of a step towards doing as I intend than is the former. So, there is a difference (possibly of kind, possibly of degree) between mere preparations and steps, on the one hand, and a difference of degree among steps, on the other. Steps are all distinct from preparations, but some are more substantial than others. In addition to an intention, the crime of attempt requires that a defendant took steps and that her steps were sufficiently substantial; preparations won't do and neither will small steps. A guilty defendant took steps towards doing as he intended and he went far enough to warrant criminal liability.

Criminal lawyers and judges use these two intuitive distinctions, and it is quite possible that they mark genuine joints in nature between types of actions. Perhaps, that is, it would be possible to provide criteria for distinguishing between mere preparations and steps, and provide criteria for determining when one step is more substantial than another. However, even if we had such criteria, the distinctions are not, in and of themselves, normatively useful. To see this, notice, first, that the distinction between mere preparations and steps is normatively irrelevant: whatever distinguishes steps that are far enough along to warrant criminal liability from those that are not will also serve to distinguish steps warranting criminal liability from mere preparations. If we have, that is, a way of explaining why taking 50 steps is enough for liability for the ascent of the stairs, while taking one step is not, then we will also have a way of explaining why tying my shoes is not enough for criminal liability; we won't need criteria for distinguishing the tying of the shoes from the taking of one step. And, further, notice that if we are able to explain why taking 50 steps warrants criminal liability, while taking 49 does not, then we would be able to explain why taking one or two steps does not, even if we cannot

provide entirely satisfactory criteria for distinguishing the intuitive difference of degree between taking one step and taking two. That is, the distinction in degree between taking one step and taking 50, if it is normatively relevant, will be only derivatively so; what will really matter normatively is the question of where to place the threshold. What these points suggest is that our purposes here will be best served by dividing the territory a bit differently from the way in which these intuitive distinctions divide it. I explain.

Consider a broad class of actions that includes both those actions that we think of intuitively as mere preparations and those we think of as steps: actions that are believed by the agent to be necessary means to doing as he intends and are performed in part at least because of that belief; call that class of actions “the class of means”.⁹ Many actions that an agent performs, even if required in order to do as he intends, do not belong in this class. Much that is necessary to successfully climbing the staircase is not performed by me in part because I believe it to be so; I can’t climb the staircase if I don’t breathe, but (without telling some special story) it doesn’t appear that I breathe *because* I believe it to be a necessary means to climbing the staircase, and so breathing is not in the class of means. Similarly, if I stand at the bottom and announce, “Now I shall ascend!”, my report of my intention is not ordinarily made because I believe it to be a necessary means to climbing the staircase; reporting my intention is not ordinarily in the class of means. I do, however, both tie my shoes and ascend the first stair in part because I intend to ascend the staircase and believe those acts to be necessary to doing as I intend. Now, the normatively significant distinction—the distinction we would like to understand better and which is reflected in the law governing attempts in every jurisdiction—is between actions *in the class of*

⁹ By “necessary means” I include acts that are believed by the agent to be a necessary part of some sufficient means, but are not believed to be strictly necessary. So, for instance, I might think that loading the gun is not strictly necessary for killing someone, since I could strangle him instead, I might also think that it is necessary if I’m going to kill him by shooting him. In such a case, the loading of the gun is a member of the class of means.

means; in particular, some are thought sufficient for the act element of attempt, others are not.

How should the law divide the class of means? This question has been answered in various ways at various times in the history of the criminal law and it is answered differently in different jurisdictions through various well-known “tests” used to define the act element of attempt. To give just one example, the “Last Act” test specifies that any act that is such that no additional acts are needed for completion of the crime is enough for the act element of attempt. The “Last Act” test, then, provides a criterion for dividing the class of means into those that are enough for the act element of attempt and those that are not. For convenience, let’s define the term “substantial steps” to mean “acts performed as means to doing as the agent intends and that are sufficient for the act element of attempt”, leaving it open, for now, what test is best for determining when an act in the class of means is a substantial step, in this sense. In this sense of the term, throughout the history of the criminal law, and in every jurisdiction, there are two elements of an attempt to commit crime C: (1) an intention to C, and (2) substantial steps. We’ll use the term “insubstantial steps” to refer to acts that are members of the class of means but are insufficient for the act element of attempt. (In the fourth section of this paper, the question of how to draw the substantial/insubstantial distinction will be discussed.) By dividing the territory this way, it is possible to make some headway in understanding the justification for the act element of attempt without appeal to any distinction, intuitive or otherwise, between mere preparations and steps.

Now, if we were constructing a law that involved a wholehearted rejection of moral luck and a wholehearted commitment to a moral justification for punishment, we would want to define “substantial step” in such a way as to ensure two things. First, we would want to define the notion in such a way that the difference between the agent who attempts the crime and the agent who completes it is solely a matter of luck; we would want the gap between substantial step and completion to be filled only by luck. After all, given the rejection of moral luck, the thought that the attempt and the completed crime are to be

punished equivalently, together with the thought that punishment is to align with moral culpability, requires that the difference between them—the difference, that is, between a substantial step towards success and success itself—is merely lucky. And, second, we would want to ensure that the difference between any agent intending to commit the crime who does not take substantial steps, on the one hand, and an agent who attempts the crime, on the other, *is not* merely lucky. After all, the former is not guilty of criminal attempt, while the latter has attempted a crime and so is liable to be punished; if we don't believe in moral luck, and the only difference between these two agents is lucky, we are morally required not to treat them differently, contrary to what the law requires.

In fact, it is very difficult (although, as this paper will establish, not impossible) for a substantial step to be defined in a way that meets these two strictures. How far past one's body one's actions progress is at least partly a matter of luck, for what actions one has performed in moving one's body is often a matter of what events one's bodily movements cause, or fail to cause, and this is subject to a whole variety of well-known vicissitudes. Whether, to use Donald Davidson's famous example, my intentionally moving my finger is the act of illuminating the room depends on whether my bodily movement causes the switch to flip, and whether its flipping causes electricity to flow to the light, and whether that flow actually results in the light's being illuminated, and whether the light's illumination spreads to the room itself.¹⁰ Thus it seems that if a particular act counts as a substantial step, so should any step that differs from it only in virtue of what happens to be caused by the agent's bodily movement. If shooting someone is a substantial step towards killing him, then so should pulling the trigger be, for whether the pulling of the trigger is *a shooting* depends on much that is out of the agent's control, such as the gun's being in working order, and the absence of any fortuitously placed birds to intercept the bullet. But, by the same reasoning, if pulling a trigger is a substantial step towards killing someone, then so should be moving

¹⁰ Donald Davidson, 'Actions, Reasons and Causes' in *Essays on Actions and Events*, Oxford: Oxford University Press, 1980, pp. 3–19.

one's finger (in just the way one does when one pulls a trigger), for whether the moving of one's finger is a pulling of a trigger, rather than a mere moving of one's finger, turns on such chance factors as the presence of the trigger; if the gun hadn't had one, one's finger movement would not have been a pulling of a trigger. Thus, there's no principled reason to define the shooting to be a substantial step and not the moving of the finger, if one's aim is for the difference between attempt and completion to be a merely lucky difference.

Further, and worse, whether one's intent makes its way into even simple physical action is always, in part, a matter of luck—we need well-wired bodies and curare-free bloodstreams to accomplish the trick. Thus it seems that the gap between merely intending to act and taking substantial steps will involve at least some luck. If there's no moral luck it would follow that any definition of "substantial step" that sees a simple bodily action as necessary to have taken a substantial step is drawing a moral difference where there is only a lucky difference, for the person who performs any physical action still differs only by luck from someone who has done less, or, perhaps, has merely intended. Why is the twitch of the finger a substantial step towards shooting someone, and not the mere tensing of the muscle in the man whose fingers aren't responding to the tensing of his muscles? And if the tensing of the muscle is sufficient, why isn't the sending of the neural impulse to the muscle sufficient in the man whose muscles aren't responding to neural impulses? And so on.

In short, the fact that the route from bodily movement to completion is subject to luck places pressure on those who deny moral luck to see substantial steps as no more than basic physical actions. And the fact that the route from intention to bodily movement is subject to luck places pressure on those who deny moral luck to see a substantial step as something less, even, than any bodily movement, including a basic physical action. These are the results we reach by repeatedly subtracting luck and repeatedly refusing to allow there to be either a moral difference or a difference in punishment between the person whom we started with and the person who results when luck is subtracted. Thus, it seems that a legal regime that takes some

acts performed from the belief that they are means to doing as the agent intends to be necessary for an attempted crime, and which holds firmly to a moral justification of punishment, is committed to some form of moral luck, even if it also gives the same sentences to those who merely attempt crimes and to those who complete them.

I say that this *seems* to be so. In fact, under the view to be argued for in this paper, this is merely an appearance: one can consistently accept the following three propositions:

- (i) There is no moral luck.
- (ii) An attempted crime consists of an intention and substantial steps.
- (iii) An event is a substantial step only if it is an action performed by the agent as a means to doing as he intends.

For simplicity, let's call the conjunction of (ii) and (iii) "the requirement of substantial steps". The challenge is to justify the requirement of substantial steps while consistently denying that there is moral luck.

In fact, it can seem very easy to meet the challenge. A first approach sees the requirement of substantial steps as arising from what is known as "the act requirement", the legal stricture resting on the idea that criminal punishment is never justifiably applied simply in virtue of the thoughts or dispositions of its recipient; *some* kind of action is needed for criminal responsibility. This requirement imposes a restriction on the definitions of crimes: no crime can be defined in such a way that all of its elements can be present despite the defendant having performed no voluntary action.¹¹ The easiest way to meet this demand, in crafting a definition of a crime, is to include an act among the elements of the crime.¹² The act

¹¹ This is a bit weaker than the claim that every crime must have an act as an element. It leaves open the possibility that a definition of an offense meets the act requirement if some result or circumstantial element requires that the defendant performed a voluntary act. However, for our purposes this distinction is of little importance.

¹² Another way would be to include a result or circumstance among the elements that can only be present if the agent performs some voluntary act. See the previous footnote.

requirement is undoubtedly justified.¹³ Further, it seems that it should be justifiable independently of any view about moral luck; whether one accepts or denies moral luck, that is, one would still have sufficient reason to accept the act requirement. Thus, the requirement of substantial steps, on this view, is purely formal: it's required for the definition of attempt to meet the act requirement, but it is not intended to sort the morally deserving of punishment from the morally undeserving, nor is it intended to lump defendants who differ only with respect to luck. Since the act requirement is not intended to do either of these things, on this view, neither is the requirement of substantial steps. The seeming inconsistency among (i), (ii) and (iii) arises, in part, from the assumption that the requirement of substantial steps is to serve one or both of these functions. The suggested view simply denies this assumption by identifying a distinct purpose of the requirement of substantial steps, namely, to meet the act requirement.

There is something right about this first approach, but it is not sufficient to justify the requirement of substantial steps, and so is insufficient to relieve the tension among (i), (ii) and (iii). The problem is that a requirement much weaker than the requirement of substantial steps could serve to satisfy the demands of the act requirement. A law of attempts, for instance, that took action expressive of intent, but not taken by the defendant to be a means to the execution of his intention—such as, for instance, a verbal report by the defendant of his intention—would also satisfy the act requirement's demand.¹⁴ Indicating why the inconsistency

¹³ For discussion see, for instance, Herbert Morris, 'Punishment for Thoughts' in *Monist*, v. 49, 1965, pp. 342–376; Gerald Dworkin and David Blumenfeld, 'Punishment for Intentions' in *Mind*, v. 75, 1966, pp. 396–404; Michael Moore, *Act and Crime*, Oxford: Clarendon Press, 1993, esp. chs. 2 & 3.

¹⁴ There is logical space to deny this claim. One who does so, however, would need to provide a justification for the act requirement, consistent with the denial of moral luck, which had the implication that only acts taken as means to doing as one intends could satisfy the act requirement. To provide such a justification is just as hard as it is to provide a justification of the requirement of substantial steps; in fact, the two tasks are the same. Hence someone who chooses to occupy this bit of logical space doesn't evade the problem that this paper is trying to solve.

between (i), (ii) and (iii) is merely apparent requires more than can be provided by an appeal to the need not to criminalize mere thoughts and dispositions.

A second approach tries to make (i), (ii) and (iii) consistent by appealing to special purposes of legal punishment that are furthered by sometimes punishing only one of two equally deserving defendants. Under such a view, although there is no moral difference, but only a lucky difference, between the person who has the intention and takes steps that fall short of being substantial, on the one hand, and the person who takes substantial steps, on the other, there is some other difference, admittedly due only to luck, that makes it appropriate to punish only the latter. For instance, if our aim in punishing is to threaten just enough unpleasantness, and no more, as is needed to lower the incidence of the peculiar forms of damage that crimes cause, then we might define substantial steps, with Oliver Wendell Holmes,¹⁵ to be those acts that impose a severe risk of performance, are in “dangerous proximity”, of the completed crime. Thus, although the person who intends to act and takes less-than-substantial steps is just as morally culpable as the person who takes substantial steps, the latter, but not the former, is worthy of criminal punishment. It makes sense to send one’s few firetrucks to the fire that’s close to the city, neglecting the one far away, and to do so isn’t to imply that the fire one chooses to fight is any more of a conflagration than the fire burning harmlessly far away; it’s simply more dangerous. Similarly, we should invest what resources we have for punishment, according to this line, in punishing those types of act that are more likely to cause the harms associated with completed crimes, which isn’t to say that they are morally worse than those that fall farther short of completion. This approach to justifying the requirement of substantial steps, while accepting that there is no moral luck, is exactly as defensible or indefensible as the view of the justification of criminal punishment that it presupposes, a view that has

¹⁵ Oliver Wendell Holmes, *The Common Law*, Boston: Little Brown, 1881, pp. 68–69.

been extensively explored.¹⁶ This paper says nothing to suggest that this approach is flawed, although I believe it to be.¹⁷ For our purposes here it is enough, perhaps, to lay bare one of the assumptions of the ensuing discussion: if the requirement of substantial steps can be justified in a way that does not require us to cleave criminal punishment away from moral desert, such a justification is to be preferred to any alternative. Such a justification is provided here, and so there is no need to huddle behind the need for social order as justification for giving unequal treatment to equally deserving defendants, as this second approach does.

A third approach appeals to the nature of trying to act and insists that trying to act involves two components: a mental state such as an intention, and some further act performed by the agent in furtherance of the intention. The idea is that in proscribing trying to act criminally, we are, necessarily and because of the very nature of trying to act, proscribing more than just a mental state; the mental state must be manifested in the form of mental or bodily action. While there may be intention, there will have been no *effort*, on this view, unless and until there has been some step. Hence, the requirement of substantial steps is justified for roughly the same reasons that our law is justified in holding that there has been no murder if there has been no death; the very nature of murder requires a death and the very nature of attempt requires action.

As we'll see, the approach to justifying the requirement of substantial steps to be presented here bears affinities to this approach and we will return to compare them later. However, it is important to see that as so far formulated the approach fails to meet the stricture on a justification imposed by those who think there is no moral luck. After all, just as the question of whether or not a person commits a murder is often in part a

¹⁶ For an entrance into the vast literature on this topic, see Joel Feinberg, 'The Classic Debate' in *Philosophy of Law*, Feinberg and Coleman, eds., Belmont: Wadsworth-Thomson, pp. 799–804.

¹⁷ I say a word or two about such an approach as applied in another context in 'The Government Beguiled Me': The Entrapment Defense and the Problem of Private Entrapment' in *The Journal of Ethics and Social Philosophy*, v. 1, n. 1, 2005. See especially pp. 18–19.

matter of luck—if the paramedics had arrived just moments earlier, there would have been no death and so no murder—under the view that takes a component of trying to act to be a substantial step, trying to act will, itself, be partly a matter of luck. Compare two agents who have the relevant mental state but only one of whom is lucky enough to have that mental state cause the action component of trying to act; nothing but luck distinguishes them, but if trying to act is legally proscribed, then this lucky difference makes a moral difference. Perhaps it will turn out that trying to act includes action, as proposed. But even if it is true it cannot be this fact that justifies the requirement of substantial steps without that justification relying on the existence of moral luck as a premise. The mere fact that trying to act includes action doesn't justify punishing those who try and not those with the same intentions who, because of luck, fail to try, unless there is moral luck.

A fourth approach, which is the centerpiece of the discussion here, starts where the second approach does, by asserting that the requirement of substantial steps serves a distinct legal interest from the interest in punishing the culpable. However, the third approach departs from the second by suggesting that the requirement of substantial steps serves the interest of punishing the culpable *whose culpability is in the right way evident to us*. Further, there might be reasons for thinking that only when a defendant has taken substantial steps do we have the right kind of *evidence* of his culpability. So, although we recognize that there are equally culpable agents who differ only in luck from those whom we convict of attempted crimes, the lucky difference does not speak to a difference in culpability, but, instead, to a difference *in our evidence for* culpability. So, just as we don't see an inconsistency in a legal regime that convicts only the thief whose theft is proven by the admissible evidence, and not the thief who, through a lucky break could not be proven to be one, perhaps because crucial pieces of evidence were inadmissible, there is no inconsistency in (i), (ii) and (iii). As attractive as this easy solution to the problem is, the next section of this paper will argue that it is incomplete as it stands and completing the solution is no easy task. Roughly, the

problem is that there doesn't appear to be anything particularly special about substantial steps when only their evidential import is considered; other forms of evidence would do just as well. But if so, then why not simply require that the defendant have the features that are evidenced by substantial steps and that he or she be proven to have them beyond a reasonable doubt? Why make substantial steps a distinct element of the crime of attempt?

The third section of the paper attempts to answer these questions. The aim of that section is to provide a criterion that an agent must meet if he is to be rightly said to be trying to do something. From this it emerges that it is no simple thing to have evidence that a person is really trying to do something he failed to do. To have such evidence we need to have evidence that he is guaranteed with necessity to complete the act under certain specified circumstances. It emerges that actions taken as a means to reaching an intended end could provide such evidence, and it seems extremely unlikely that anything else could provide such evidence. Thus, substantial steps are important for evidential reasons, although not for the evidential reasons adduced in Section 2. Substantial steps, it emerges, are an *indispensable* form of evidence; they are indispensable evidence that the agent is really trying to act criminally; it is almost certain that nothing else would do. This we learn from reflecting on the very idea of trying. Section 4 indicates what the implications might be of the result of Section 3 to the question of how to distinguish between substantial and insubstantial steps.

III. THE OBSTACLE TO AN EVIDENTIALIST SOLUTION

Consider an agent who has committed a crime, has no justification or excuse, and is morally deserving of punishment for doing so. Many of the properties of the agent are irrelevant to the justification of the claim that he deserves punishment for the crime. Ordinarily included in the class of irrelevant properties, for instance, are his hair color, his birthday, his parents' criminal history, and so on. Doubtless we can imagine extraordinary cases in which all of these properties are relevant for some reason or other; we can imagine cases, that is, in which

appeal to them is part of what justifies the claim that the agent is deserving of punishment for his crime. But, still, there are always, in every case, many properties of the agent that play no part in that justification. Let's call that set of features of an agent, his actions or his circumstances that *do* play a part in the justification of the claim that he deserves punishment for a crime, the agent's "culpability-constituting features". We can imagine a variety of theories about what is and is not included in the class of culpability-constituting features of agents generally, or of a particular agent. Is the agent's degree of remorse of relevance? Is the fact that the agent suffered a childhood of deprivation and abuse of relevance? Is the agent's character of relevance? These and related questions are difficult, but the concept of culpability-constituting features as such is neutral with respect to various answers to them.

On an extreme view, the culpability-constituting feature of someone who commits a crime without justification or excuse¹⁸ is just his mental state, such as his intention to act in the ways necessary for the completed crime, or his reckless indifference to the harms that his conduct would be likely to cause. According to such a position, agents with the same intention, who act differently, might differ in many ways, but they do not differ in the degree to which they *deserve* punishment; if the law mandates that one is punished and not the other, or that they are to be punished differently, on this view, it is not because of a difference in the degree to which they are morally deserving of punishment, but because of some other difference between them. At the

¹⁸ Throughout, I am assuming that the defendant accused of attempt has neither a justification nor an excuse for his conduct. Therefore, limiting ourselves to consideration of the culpability-constituting features of those who complete crimes *and lack justification or excuse* is harmless. There are, however, difficult questions that I am setting aside about how to understand the relationship between justification and excuse and those features that justify the claim that the defendant is morally deserving of punishment. Do excuses, for instance, always indicate that the agent lacks one of those features? Do justifications show that the person is not to be punished despite possessing features that make him morally deserving of punishment, or would under other circumstances?

opposite extreme is a view that takes the culpability-constituting features of an agent to include the acts and results, and perhaps even the circumstances necessary for the crime he completed. On this view, there is a difference in the culpability-constituting features of the defendant who completes the crime and the defendant who attempts it and fails. Let's assume that one of the class of views falling short of this last extreme is correct; let's assume, that is, that something less than completion of the crime is needed to realize the relevant culpability-constituting features. This is not to say that we need go all the way to the other extreme and assert that the defendant's mental state determines, all by itself, the degree to which he is deserving of punishment; we could go this far, but we needn't. All that we are assuming for now is that if we do take the culpability-constituting features of agents to include more than mental state, we don't thereby take them to include all that is involved in completing the crime, and we might very well take mental state to be sufficient.

The view to be considered in this section can be stated like so: Substantial steps are required for an attempted crime because they serve as evidence of the same (or worse¹⁹) culpability-constituting features as found in those who complete the crime without justification or excuse. The idea is simple enough: If precisely that which justifies taking the person who completes the crime to be morally deserving of punishment is present also in the person who attempts the crime, then the person who makes the attempt is just as deserving of punishment as the person who completes the crime. But it is also important that we have adequate evidence of this similarity between the defendant being tried for attempt and those who complete the

¹⁹ This parenthetical appears here to capture cases in which the completed crime does not require intent (such as purposive homicide which requires only that the defendant foresaw with certainty that a death would result from the act he intended) but the attempted crime does. In such cases, the acts the defendant undertook as means provide evidence of his intent, not his recklessness, and so they provide evidence of a different set of features from the culpability-constituting features of the person who completes the crime. However, since intent is thought to make one that much more deserving of punishment than recklessness, the requirement of substantial steps is still justified because of the evidence that substantial steps are thought to provide.

crime. Under the view, to require substantial steps is to constrain the evidentiary route through which the culpability-constituting features are shown to be present: the prosecution must provide the sort of evidence of those features that is provided by showing that the defendant adopted means to doing as directed by his criminal intention.

The *Model Penal Code* adopts this approach in its test for determining if a defendant's acts do, indeed, constitute substantial steps. According to the *MPC*,

Conduct shall not be held to constitute a substantial step...unless it is strongly corroborative of the actor's criminal purpose. (*Model Penal Code* §5.01(2))²⁰

There are two elements to the *MPC*'s approach that distinguish it from other evidentialist views of the sort under consideration in this section. First, the *MPC*'s approach assumes that the defendant's mental state is all that is included among the culpability-constituting features of the person who completes the crime; that is why the act need only be evidence of the mental state of the person who attempts the crime (namely his intention), and nothing other than that, for it to count as a substantial step. Second, the *MPC*'s approach assumes that an act needn't be sufficient evidence to establish, all by itself and beyond a reasonable doubt, that the defendant had the relevant mental state. We can imagine views that vary from the *MPC*'s in both of these dimensions, either taking something other than mere mental state to be required for desert of punishment for the completed crime (such as a trait of character or a disposition), or by either strengthening, or even weakening, the evidential relation between the act and the relevant culpability-constituting features, if the act is to count as a substantial step. Under

²⁰ The *MPC*'s formulation does not make explicit that the conduct in question must be believed by the defendant to be a means to doing as he intends. For all the text says, a verbal report that was "strongly corroborative" of the defendant's intention would satisfy the act requirement. However, despite what the text says, it is clear that the *MPC* requires that the conduct in question be believed by the defendant to be a means to doing as he intends.

discussion in this section is the full family of views of which the *MPC*'s is just one member.

It's important to emphasize the distinctive feature of the *MPC*'s approach which is shared by all other approaches in the family under consideration here and which can seem somewhat puzzling. All of these views make two demands: first, that the prosecution show beyond a reasonable doubt that the defendant had the intention required for the crime of attempt, and, second, that the acts which the defendant took constitute evidence of his intention (and possibly more than just his intention). To meet the first demand, the prosecution can appeal to the defendant's steps but they can appeal to much more than just those; they can appeal, for instance, to his verbal reports of his intentions, to other acts that are expressive of intent but are not performed as means to doing as he intends, and even to things other than his actions that might serve to evidence his intent, such as the judgments of psychiatrists and other experts. But to meet the second demand, the prosecution must argue, even if it has already met the first demand, that the acts which the defendant took as means to doing as he intended constitute, themselves, non-trivial evidence of his intent (and of whatever else is included among the relevant set of culpability-constituting features). So, the prosecution has to provide adequate evidence of intent *and* a large part of the prosecution's case for the claim that the defendant had the requisite intent must derive from the acts the defendant performed in furtherance of his intention.

By analogy, but also by contrast, consider an approach for justifying the exclusion of coerced confessions from evidence. Although coerced confessions are, generally, of less probative value than uncoerced confessions—they are more likely to be made by the innocent, that is, than uncoerced confessions—there are, still, coerced confessions that constitute evidence of guilt; some people confess under coercion *because they are guilty*, even if they also do so in order to stop the beating. So, in excluding coerced confessions, we exclude from the jury's consideration some facts that could help them to reach a more accurate verdict. This practice is justified by appeal, first, to the obvious and pressing interests that are protected through this blanket

exclusion, such as the interest that guilty and innocent alike have in a police force that doesn't do this sort of thing, and the interest that we all have in keeping from the jury evidence the epistemic import of which they are likely to exaggerate; and, second, through a claim to the effect that these protected interests outweigh the resulting loss in protection from crime. What we have here is a common form of justificatory explanation of legal doctrines specifying the exclusion of some type of evidence: in justifying the doctrine in question, we note both the probative value of the excluded evidence and the losses of other sorts associated with allowing it to reach the jury; we then claim that the losses resulting from exclusion are outweighed by the gains.

The particular evils that we exclude coerced confessions to avoid are either not at risk or not avoidable through imposing the requirement of substantial steps. It is true that under the evidentialist justification of the requirement of substantial steps, that requirement, like the exclusion of coerced confessions, constrains the evidentiary route through which the prosecution can meet its burden. But the requirement differs from the exclusion of coerced confessions in (at least) one important respect: where the exclusion of coerced confessions *bars* the prosecution from producing a certain sort of evidence, the requirement of substantial steps, under the class of views on offer, instead *requires* the prosecution to produce a certain sort of evidence. It's not enough merely to show beyond a reasonable doubt that the defendant had the culpability-constituting features of those who complete the crime; this needs to be shown in a particular way, through appeal to the acts that he performed as means to doing as he intended. What this implies is that there is an important difference between the justification for the exclusion of coerced confessions and the justification along evidentialist lines for the requirement of substantial steps. A justification of the requirement of substantial steps cannot appeal to the dangers of allowing any particular sort of evidence, for the requirement does not exclude any kind of evidence. It must rather appeal to either the dangers of allowing conviction when there are *only* other forms of evidence of the relevant culpability-constituting features, or, correlatively, the goods achieved through requiring

the kind of evidence that is supplied by substantial steps. What this implies is that in requiring substantial steps, in contrast to excluding coerced confessions, we do not incentivize the police not to act badly, for there is no kind of evidence, no matter how brutally acquired, that *the requirement itself* tells our government it is not allowed to use. Similarly, the requirement of substantial steps does not remove from the jury's consideration any evidence that the jury is likely to weigh more heavily than it warrants; the jury will still be asked to consider, for instance, the defendant's verbal reports of his intentions. Thus, it seems, the view on offer needs to identify some other interests, distinct from the interests served by excluding coerced confessions, which is served by constraining the way the prosecution must meet its burden in the way the requirement of substantial steps does. What could those interests be?

By excluding coerced confessions we lower the incidence of a form of involuntary speech, namely confessions prompted by coercion. Perhaps by imposing the requirement of substantial steps we make room for certain forms of valuable voluntary speech that would otherwise be quelled, namely speech expressive of intent.²¹ The point can be illustrated with an example. Consider the case of *State v. Duke* (700 So. 2d 580 (Fla. App. 1998)). The defendant had a series of conversations in an internet chat room with someone going by the name of "Niki" and claiming to be a 12-year old girl. "Niki" and Duke discussed having sex and then arranged to meet for that purpose in a parking lot on a particular night. They agreed that Duke would signal by flashing his lights and then take "Niki" back to his home. Duke went to the parking lot at the appointed time and flashed his lights, at which point he was arrested by the detective who had posed as "Niki". He was charged with attempted sexual battery.

Although the intention to commit sexual battery of the person who commits that crime without justification or excuse may not be the only culpability-constituting feature of that person, it is certainly among them. Thus, two crucial questions to ask in determining if Duke has taken substantial steps, under

²¹ Thanks to Seana Shiffrin for pressing this point in conversation.

the evidentialist justification of the requirement under discussion here, are these: What is the nature of our evidence of Duke's intention? And do the actions that Duke took as means to further his intention serve as particularly strong evidence of his intent—in the language of the *Model Penal Code*, do those actions “strongly corroborate” his having the relevant intention? The relevant evidence consists in two different sorts of conduct on Duke's part: conduct that is merely expressive of various mental states such as desires and intentions but is not performed in order to further his intentions, and conduct that Duke takes to actually further his intention to commit sexual battery. By definition, none of the conduct in the first category can constitute, even partially, substantial steps on Duke's part; those actions aren't performed by him as means to further his intention, and so aren't substantial steps. So, if Duke took substantial steps, he did so only in virtue of having performed the acts in the second category.

What conduct of Duke's belongs in the second category, the category of acts performed as means to doing as he intends? It seems clear that in this category belong Duke's proposal to meet, his suggestions about how to go about it, his driving to the parking lot at the agreed time, and his flashing his lights. In performing these actions, he is acting on his intention to commit a crime; he's adopting means to doing as he intends. What conduct belongs in the first category, the category of acts expressive of intent, but not performed in order to further the intention? Consider, for instance, the various things that Duke says in the chat room prior to proposing a meeting. In saying these things, is Duke taking means to doing as he intends—is he “warming up” his intended prey, say—or is he, rather, simply acting in a way which is explained by his intention, but is not performed by him in order to further it? By analogy, imagine that I intend to leave my house and, on my way out, I turn on the porch light. My turning on the porch light is expressive of my intention to leave, but it is not performed in order to further that intention as, say, my turning the front door knob is; I get no closer, and don't believe myself to get any closer, to leaving my house by turning on the porch light. Still, the fact that I did this is

expressive of my intention to leave. Are the things Duke says in the chat room prior to proposing the meeting more like the turning on of the porch light, or more like the turning of the door knob? The fact is that we don't know enough about Duke to know the answer to this question. What this suggests is that there is reasonable doubt as to whether or not the conduct in question is performed by Duke as a means to doing as he intends, and so if that were all he had done there would have been reasonable doubt that he had taken substantial steps. Let's assume, then, that the conduct in question is merely expressive of intention, and not performed in order to further his intention. It makes sense to construe Duke's conduct in this way, within the strictures of reasonable doubt.

Now, the suggestion which launched discussion of this example was this: perhaps the requirement of substantial steps serves to protect certain forms of speech, speech that would be disincentivized were there no such requirement. In the context of our example, there appear to be two forms of speech that would be disincentivized were there no requirement of substantial steps: (1) speech expressive of the intention to commit sexual battery made by those who have that intention, but not made in furtherance of that intention, and (2) speech that would constitute evidence of intent to commit sexual battery made by those who don't actually have that intention but who are speaking as a way of acting out a fantasy or pretending to have the relevant intention. After all, people offering both kinds of speech in chat rooms would find themselves at risk of prosecution for attempted sexual battery were there no requirement of substantial steps, for there would be evidence, in both cases, of one of the culpability-constituting features of those who commit the crime, namely the intention to do so.

There are a wide variety of reasons why we might not want our legal system to have a chilling effect on these kinds of speech. But it's important to see that many of these reasons cannot coherently be appealed to by the advocate of the evidentialist justification for the requirement of substantial steps under discussion. For instance, we might hold the view that speech of the sort being considered serves as a harmless "release

valve” for those with the intention to commit crimes; the suggestion is that the requirement of substantial steps then lowers the incidence of completed crimes by providing those who might commit them with another way to release their criminal impulses. The problem with this suggestion is that laws against attempted crimes generally serve to lower the incidence of completed crimes and so the harder it is for people to be found guilty of attempt, the less of an impact such laws will have. It is then an empirical question whether we will have more or less completed crimes by incentivizing or disincentivizing the sort of speech described. But it doesn’t seem that the justification of the requirement of substantial steps should need to await the outcome of the relevant empirical studies, studies that would need to be done to answer that question.

Another reason we might offer in favor of laws that don’t disincentivize language expressive or naturally taken to be expressive of criminal intentions is this: to speak in such a way as to express intention simply doesn’t provide any evidence that the speaker will act on such an intention, but since it is action in furtherance of criminal intention that we want to discourage, nothing is gained by discouraging the relevant kinds of speech. This reason, however, cannot be accepted by someone who holds the evidentialist justification of the requirement of substantial steps. After all, someone who offers this reason, and who also accepts that criminal punishment is justified only when deserved, holds that the person who has a criminal intention but does not take a means to doing as he intends is less deserving of punishment than someone who does adopt means. Thus, for such a person, taking a means is not of importance because it is *evidence* of some set of culpability constituting feature or features of the person who commits the crime; it is, rather, of importance because it is *part of* what makes such a person deserving of punishment.

Of course, one can take (certain forms of) speech, and the freedom to engage in (those forms of) it, to be of intrinsic value, and see there as always being some kind of loss from any sort of social or legal policy that results in less of it. Someone who takes this approach can justify the requirement of substantial

steps on the evidentialist line under discussion, but only if he is also willing to accept that the loss in probatively valuable evidence brought about by the requirement, and all that is entailed by that loss, is outweighed by the value of the speech and the freedom to speak that is thereby not disincentivized. It would be extraordinary to discover that all those who take the requirement of substantial steps to be justified are committed to such a strong position about the value of speech and the freedom to engage in it.

However, imagine that it were agreed that the balance sheet of the requirement of substantial steps is positive. Say it were agreed, that is, that the loss in probatively valuable evidence brought on by the requirement is outweighed by the gain in fewer disincentives to speech. Would this serve to justify the requirement of substantial steps? It would not serve to those who accept the following (roughly formulated) restriction on justifications of criminal law doctrines: A justification of a criminal law doctrine is acceptable only if it could lead a defendant who is punished as a result of the doctrine, while other equally deserving defendants are not, to recognize the justice of his treatment. Doubtless this restriction would need to be formulated with greater care if it is to be fully defensible. But the rough idea has great appeal. Imagine a defendant who is arrested after saying precisely the things that Duke said prior to proposing a meeting and imagine that it is shown that that defendant had the intention to commit sexual battery; this man is acquitted of attempt on the grounds that his conduct did not constitute substantial steps. Now imagine that Duke is convicted and he asks why he, who had the same intention as the other man and who caused no more harm than the other man, is to be punished while the other man goes free. Can we expect Duke to recognize the justice of this when it is explained to him the effect that the alternative would have on speech? This seems to be the wrong kind of reason to give to Duke in such a circumstance. It is not a reason linked to the deeply moral functions of the practice of criminal punishment.

The discussion so far has focused on goods and evils gained and avoided through the requirement of substantial steps that

are, themselves, independent of the probative value of the evidence provided by action performed as a means of doing as one intends. The analogy with the exclusion of coerced confessions pushed us in that direction since that exclusion is to be justified by appeal to such goods and evils. But perhaps there is some positive probative evidential value to action taken in furtherance of intention. Perhaps such behavior serves as evidence of some set of culpability-constituting features that simply are not evidenced any other way. If this were so, that would serve the evidentialist justification of the requirement of substantial steps under discussion. What are the prospects of such an approach?

In answering this question, assume for a moment not just that an intention to commit sexual battery is *among* the culpability-constituting features of the person who completes that crime, but that it is the only such feature. Under this assumption, it is hard to see what special probative significance there could be to action performed by Duke in furtherance of his intention. To be sure, in Duke's case, conduct expressive of intent, but not performed in order to further his intention, was insufficient to establish intent. Notice, however, that the acts he performed in order to further his intention were also insufficient: without knowing about his chat room conversations we can imagine a variety of reasonable explanations for the act, say, of driving to the parking lot and flashing his lights that don't involve ascribing him with any criminal intention, much less the intention to commit acts that would constitute sexual battery. Imagine replacing each act that he performed in furtherance of his intention with an act merely expressive of intention, but with the same probative significance. If, say, his driving to the parking lot raises the probability by x that he has the intention, then imagine that the record shows, instead, some act expressive of intention, but not performed in furtherance of it, but which raises that probability by precisely x . Perhaps we imagine that he's had conversations with a friend in which he's mentioned his intention to have sex with a child; "No," he said, in response to questioning, "I really do intend to do it. I'm not just talking here." Or perhaps we imagine that he's purchased pornography depicting the very acts that he has described himself as having an intention to commit.

So long as the only culpability-constituting feature is the intention, the evidentialist approach fails to identify any special evidential significance to action taken in furtherance of intention, and so fails to explain why substantial steps are essential for the crime; they provide good evidence of intent, granted, but evidence of equal quality could be provided through an alternative route. But what if the culpability-constituting features of the completed crime include more than just intent? In that case, do actions performed as means to doing as one intends provide some special kind of evidence? Here we encounter two types of account of the culpability-constituting features. According to the first, the acts that are taken to be substantial steps are, themselves, among the culpability-constituting features of the completed crime. In short, on such a view, the person who completes the crime is deserving of punishment in part because of the steps that he took in acting on his intention, steps that the person who merely attempts the crime also took. According to the second, there is something more than intention that is required for desert of punishment, but it is something distinct from those acts that are the defendant's substantial steps. For instance, we might say that what makes the person who completes the crime deserving of punishment is not just that he intends to do something criminal, but also that he is *resolved*, or determined, to act on his intention. So, something more than intention is among the relevant culpability-constituting features: determination, or resolve (whatever that is) is also among them.

The first of these two approaches amounts to giving up on the evidentialist approach for reasons similar to those encountered by one effort to support the requirement of substantial steps through appeal to considerations of free speech. If substantial steps are, themselves, part of what makes a criminal deserving of punishment, then they are hardly *evidence* of something else of significance; they *are* the significant things, on such a view. To take the evidentialist line here would be like saying that the victim's being dead is an essential element of murder because it's good evidence that the victim is dead. More importantly for our purposes, such a view involves a commitment to moral luck in a way encountered earlier. After all, whether the agent has

performed the acts that are taken to be among the culpability-constituting features will be, in part, a matter of luck and so there will be agents who differ in the degree to which they are morally deserving of punishment despite the fact that the difference between them is only due to luck. From the point of view of one who wishes to justify the requirement of substantial steps without appeal to moral luck, this is to give up the game.

The second of these two approaches, exemplified by the view that it is not just intent, but intent-plus-resolve that is required for desert of punishment, has the same problem as that illustrated by our discussion of the view that there is no more to the relevant culpability-constituting features than intent. Without saying more about the nature of the additional culpability-constituting features, there is no reason to think that something other than substantial steps could not serve just as well as evidence of those features. Obvious candidates for the culpability-constituting feature, such as resolute intent, do not provide the view with what it needs. True, we might not take first-person reports of the defendant's resolute intention to be very good evidence of it. We don't necessarily trust the person who says, confidently, "I'm resolved to jump from the plane when the hatch opens." "Talk is cheap," we say, implying that the real test of such resolution is action; no mere report of one's own degree of resolve will serve as evidence of resolute intention. However, there are other forms of evidence of resolute intention besides either first-person reports or action performed in furtherance of it. It is common enough for applicants to various jobs that require resolve to be given psychological tests in order to determine if, for instance, they would really, when it came to it, leap in front of the bullet shot at the President. It seems quite likely that people who pass such tests really are more likely, maybe even *much* more likely, to have not just the intention to do the thing the job might require, but the *resolute* intention. To pass such a test, however, is not to take steps towards furthering the intention and so can't be a substantial step. Still, the results of such tests would serve as useful evidence of resolute intent. Further, the fact that a person has been habituated in certain ways can serve as strong evidence of the resoluteness of his intention. Someone who has undergone a

certain kind of military training is more likely to be resolute in his intention to kill the enemy than someone who has not. Thus it seems that evidence of such training, or, more generally, such habituation, can serve as quite good evidence of resolute intent, and not just intent. Relatedly, we often think that a person's emotional reactions in a variety of contexts speak to the resoluteness of certain intentions on which he has not acted. The fury felt by a person on reading about new legislation permitting more abortions, might, itself, serve as evidence—not decisive evidence, but evidence still—of the resoluteness of his intention to bomb an abortion clinic, an intention that he may have taken no steps towards executing.

Even if we shift away from resolute intention directly, the problem persists. For instance, imagine a view according to which one of the culpability-constituting features of the person who completes the crime is his failure to voluntarily abandon the plan at some late stage.²² The claim would then be that for the person who intended to commit the crime but who failed to complete it to be guilty of attempt, it must be true that he would not have abandoned the plan at the same late stage had he reached it.²³ The evidentialist justification of the requirement of substantial steps would then require asserting that action taken by the agent as a means to doing as he intends is an essential form of evidence of the fact that the agent would not abandon his plan.

The problem with this is that if the adoption of means is to count as evidence that the defendant would not reconsider, then it must be true that people are more likely to abandon their plans if they don't adopt means than if they do. It is true that to adopt means is to sink costs into an enterprise that one hopes to

²² Such a view would fit nicely with the defense of voluntary abandonment under which a defendant can show himself to be not guilty by showing that he actually already abandoned his plan to act criminally prior to arrest.

²³ An approach that required that, in fact the defendant reached this late stage without abandoning, and not just that *he would not* have abandoned had he reached that stage, would be allowing moral luck: after all, whether or not one reaches that late stage is often a matter of luck of just the sort that those who advocate the same sentences for attempts as completed crimes deny to be of moral significance.

recoup by succeeding in the enterprise. To abandon a plan after sinking costs, then, is to assure a loss. Perhaps people are more likely to follow through, then, if they've already adopted means because they want to avoid wasting their efforts. However, it seems unlikely that reasoning of this sort will extend smoothly to the kinds of cases that concern the criminal law. Consider just one example that would seem, at first glance, to be amenable to this form of analysis. In a British case, *Regina v. Gullefer* (1 WLR 1063 (1990)), the defendant had bet on a dog race. In the middle of the race, seeing that his dog would lose, he ran onto the track in the hope that the race would be voided and he could reclaim his money from his bookmaker. He was charged with attempted theft of the bookmaker's money.²⁴ There were, obviously, losses associated with his running onto the track: public embarrassment and permanent expulsion from the dog track, for instance. Now compare Gullefer to someone who, like Gullefer, has the intention to steal from the bookmaker, but who has yet to adopt any means. Is there any reason to think this person less likely to abandon his plan than Gullefer is? It doesn't seem that the question turns on the investment in stealing that Gullefer, in contrast to this person, has already made, and so it doesn't seem likely that the adoption of means really serves as evidence of the truth of the relevant counterfactual. It just seems silly to imagine that Gullefer would reason like so: "I don't want to waste the effort of having run onto the tracks, so I'll ask the bookmaker for my money back." He might or might not ask the bookmaker for the money back, but the issue doesn't seem to have anything to do with his having run onto the tracks, but depends, instead, on other aspects of his temperament that he might or might not share with the person who intends to steal from the bookmaker, but who has yet to act on that intention. In fact, Gullefer's conduct is sufficiently bird-brained that we might think him *less likely*, rather than more likely, to follow-through with the plan than the person who intends to steal but has yet to take any steps.

²⁴ The case is discussed in R. A. Duff, *Criminal Attempts*, Oxford: Oxford University Press, 1996, pp. 58–59.

So, we still don't have an adequate explanation for the requirement of substantial steps. Under various proposals as to what is included among the culpability-constituting features of the completed crime, the requirement of substantial steps condemns a defendant who took substantial steps and not one who did not, even if we have equally good evidence in each case that the defendant possesses the proposed culpability-constituting features of the person who completes the crime. This cannot be justified under a view that takes substantial steps to be of significance only because of their evidential value.

The best way to think of the discussion thus far is like so: Recall that the aim is to find a way of consistently saying that (i) there is no moral luck, that (ii) substantial steps are required for criminal attempt, and that (iii) only acts performed in order to further the defendant's intention can be considered substantial steps. A successful evidentialist approach to making (i), (ii) and (iii) consistent must take one of two approaches. A first approach, which involves a close analogy to the justification of the exclusion of coerced confessions, argues that although there is nothing of distinctive probative value gained through the requirement of substantial steps, there are other things that would either be lost without the requirement or gained with it—there would be, for instance, a limitation placed on free speech were there no requirement of substantial steps that would be absent under a system that imposed the requirement. As we've seen, plausible efforts to take this approach encounter serious difficulties. A second approach argues, instead, that the requirement of substantial steps is justified precisely because of the great probative value of actions taken as means to doing as one intends; such an approach involves claiming, that is, that action taken as a means to doing as one intends is of such high probative value that the prosecution ought to be required to produce it. This second approach requires two things: First, the culpability-constituting features evidenced by action taken as a means to doing as one intends must be distinct from those actions themselves. Otherwise, acts taken as means are thought of as important for some other reason besides the evidence they provide of the culpability-constituting features and, more importantly, such a view would involve a commitment to moral

luck. Second, a successful evidentialist approach of this variety must identify at least one culpability-constituting feature which *can be evidenced only* by actions performed to further an intention. Otherwise, the requirement of substantial steps places a gratuitous constraint, a constraint not worth the costs, on the evidential route through which it is established that the defendant possesses the relevant culpability-constituting features. As we will see in the next section, these two constraints can indeed be met. It will be argued that our ordinary notion of trying to act requires that an agent who is trying engages in an act that may be distinct from the adoption of means but which cannot be evidenced by anything other than acts performed as means to doing as one intends.

IV. THE SOLUTION: ACTION AS EVIDENCE OF TRYING

The evidentialist strategy examined in the last section involved an assumption. It involved the assumption that the requirement of substantial steps must serve as evidence for something shared between the defendant who completes the crime and the defendant who makes a failed attempt. The tag “culpability-constituting features” was used to refer to the thing that must be shared, that which provides the justification for taking the person who completes the crime without justification or excuse to be deserving of punishment. The strategy was to suggest that conduct undertaken as a means to acting criminally was evidence of just those culpability-constituting features. However, it’s possible to justify the requirement of substantial steps on evidentialist lines without assuming that the defendant’s actions must be evidence of something *shared* with the agent who completes the crime. Perhaps, instead, actions taken as means to doing as the defendant intends are evidence of the very thing that we take ourselves to be criminalizing when we have laws against attempted crimes; maybe, that is, such actions are evidence that the defendant was *trying to fulfill his criminal intention*. This section argues that this is so. Further, and importantly, it argues that such actions are an indispensable form of evidence. Evidence must meet stringent requirements if it is to serve as evidence that an agent is trying to do something; actions taken as means to doing as one intends meet these

requirements and it seems close to impossible for any other evidence to do so.

In order to understand the requirements that a fact must meet if it is to serve as evidence that a defendant is trying to act, start with the following claim about the relationship between our ordinary concepts of ability, opportunity and trying:

(*) If ((S has the ability to A) & (S has the opportunity to A) & (S tries to A)), then S A's.

A variety of ordinary inferences that we make are explained by appeal to our tacit adherence to (*). If we're told that S tried to A but failed, we infer either that S lacked what it takes to A (that is, that S lacked the ability to A) or that S's environment threw up obstacles to S's A-ing (that is, that S lacked the opportunity to A). If we're told that S has the ability to A and is trying to A, we predict whether or not S will succeed in A-ing through examination of his circumstances; we try to determine, that is, if S has the opportunity to A. Similarly, if we are told S has the ability and opportunity to A, we predict whether S will A through consideration of the question of whether S is trying to A. All of these tendencies to make inferences, and to identify the crucial pieces of information needed to reach conclusions, indicate a commitment, in our ordinary thinking, to (*).

Despite its attractions, J. L. Austin famously denied (*):

Consider the case where I miss a very short putt and kick myself because I could have holed it. It is not that I should have holed it if I had tried: I did try, and missed. It is not that I should have holed it had conditions been different: that might of course be so, but I am talking about conditions as they precisely were, and asserting that I could have holed it. There is the rub.²⁵

In Austin's example it would seem that the antecedent of (*) is true, while the consequent false. However, this appearance can be misleading. There is a sense in which the golfer who misses has the ability and the opportunity and tries, but it is not the sense invoked in the antecedent of (*). The golfer does not differ *in any way we can discern* from the man who had the ability to sink the putt; the golfer's circumstances do not differ *in any way we can discern* from the man who had the oppor-

²⁵ J. L. Austin, 'Ifs and Cans' in *Philosophical Papers*, New York: Oxford University Press, 1979, pp. 205–232. See p. 218n.

tunity to sink the putt; and *nothing that we can discern* distinguishes the golfer's actions and efforts from those of the man who tried to sink the putt.²⁶ This is the truth expressed by someone who says that the golfer had the ability and opportunity and tried to sink the putt. And it is perfectly compatible with (*), which implies that *despite what we can discern*, one of those three conditions wasn't satisfied. Austin may have been right that (*) is not embedded in the grammar of the term "can", a term which, itself, seems to straddle both ability and opportunity; it seems to straddle, that is, both the properties of the agent and his circumstances that make room for action. But the truth expressed by (*) is not a truth about grammar, but a truth about the metaphysics of agency.²⁷

²⁶ Compare Thomas Reid: "[A] being may have a power at one time which it has not at another. It may commonly have a power, which, at a particular time, it has not. Thus, a man may commonly have power to walk or to run; but he has not this power when asleep, or when he is confined by superior force. In common language, he may be said to have a power which he cannot then exert. But this popular expression means only that he commonly has this power, and will have it when the cause is removed which at present deprives him of it." (Thomas Reid, *Essays on the Active Powers of Man*, Cambridge: MIT Press, 1969, p. 269.). Although Reid is discussing a principle different from (*) (he's discussing the principle that if a person has the power to do something then he also has the power to exert that power) his point would apply just as well to Austin's example. Ordinary speech might appear to allow for counterexamples where there are none in fact.

²⁷ Austin adds the following remark which might be thought friendly to the response to him that I am offering here: "But if I tried my hardest, say, and missed, surely there *must* have been *something* that caused me to fail, that made me unable to succeed? So that I *could not* have holed it. Well, a modern belief in science, in there being an explanation of everything, may make us assent to this argument. But such a belief is not in line with traditional beliefs enshrined in the word *can*: according to *them*, a human ability or power or capacity is inherently liable not to produce success, on occasion, and that for no reason (or are bad luck and bad form sometimes reasons?)." (Austin, 'Ifs and Cans', p. 218n) Austin could be taken to be suggesting, here, that (*) has the same status as the principle of sufficient reason. To be sure, a person can use the word "can", or the words "ability", "opportunity" and "try", without betraying a commitment to the principle of sufficient reason. But, still, the principle of sufficient reason is true, and so is (*).

Although (*) can survive Austin's criticism, it needs to be interpreted carefully. In particular, the fact that acts take time to perform has implications for how (*) should be interpreted. The paradigm cases in which (*) seems to be true are those in which the time for action is now and the action can be completed through nothing more than effort on the agent's part to complete the act now—raising one's hand or averting one's eyes; in such cases it seems that effort on the agent's part is guaranteed to issue in action. But more complicated examples, involving actions that require multiple steps to complete, seem to invalidate (*). Say that I have the ability to ascend the 100-stair staircase, and I have the opportunity to do so, and I try to do so; I ascend 50 stairs, at which point I simply change my mind, turn around and return to the bottom. It seems that there was a time, namely the time at which I took the first step, at which I had the ability to ascend to the top of the staircase, the opportunity to do it, and tried to do it; and yet I did not do it, neither then nor later. As described, the case is one in which I am trying to do something at one moment in time that I stop trying to do before the act is completed. We can also imagine cases in which the agent has the ability to do something at one moment and loses that ability before he has done that thing: at the fiftieth stair my leg breaks. And we can imagine cases in which it is opportunity which is lost before the act is completed: at the fiftieth stair the entire staircase comes crashing down. Do cases of these kinds show (*) to be false? No. They rather show that (*) should be understood like so: say that A is an act that, if it is to be performed, would begin at time t_1 and be completed at time t_2 . What is true is that an agent who has the ability and the opportunity over the t_1 – t_2 interval and tries to A over that entire interval, A's.²⁸

When an agent tries to do something and then does it as a result of the fact that he tried and in the appropriate way, the completed act begins at precisely the moment that the agent began to try to engage in the act. The act of ascending the

²⁸ Put even more carefully: Where A is a type of action that can be tokened over the t_1 – t_2 time interval, (*) says the following: If (from t_1 to t_2 ((S has the ability to A) & (S has the opportunity to A) & (S tries to A))), then S will engage in a token of A over the t_1 – t_2 interval.

staircase begins at just the moment that the agent tries to ascend the staircase, a trying that prompts him to raise his foot to ascend the first stair.²⁹ Of course, we can try, fail and try again and succeed. In such cases, the completed act begins when the second trying—the trying that blossomed into success—begins. What this implies is that it is not possible for an agent to have the ability and opportunity to do something, and try to do it, before the particular act in question begins.³⁰ If the agent is trying, then the act is beginning. What (*) says is that if an agent with ability and opportunity is trying to do something, then only one of three things can possibly prevent him from doing it: he can lose the ability, he can lose the opportunity, or he can change his mind and stop trying.

The concepts of ability, opportunity and trying are interlocking place-holder concepts. Ability is whatever it is (intrinsically) about an agent that ensures that she will act when she has the opportunity and tries; opportunity is whatever it is (intrinsically) about an agent's circumstances that ensures that she will act when she has the ability and tries; and trying is whatever it is that an agent does that ensures that she will act when she has the opportunity and ability. This implies that (*) is not merely true, it is a necessary truth: every possible world in which an agent with ability and opportunity tries to act is one in which the agent acts. In making this claim, I am asserting that there are only three kinds of things that can account for the fact that an agent fails to act a certain way on a certain occasion: it could be something about the agent, something about

²⁹ A problematic example: Say I try to ascend the first fifty stairs, planning to decide when I get to step fifty whether to continue to the top. When I get to the fiftieth stair I decide to go ahead and thereby try to ascend the staircase; I succeed in ascending the staircase. If my ascent of the staircase began when I ascended the first step, then it began before I tried to ascend the staircase, for I didn't try to do that until I'd reached the fiftieth stair. However, it seems to me that this case should be described like so. The term "the ascent of the staircase" does not refer to a single act but instead to two distinct acts that I performed: the act of ascending the first fifty stairs, and the act of ascending the second fifty. Each of those acts was begun at precisely the moment I tried to perform it.

³⁰ This is possible with respect to a type of act tokened later by an act that begins later as a result of a later trying.

his environment, or he could have failed to try. However, there are difficult cases. What should we say, for instance, about failures that are explicable, at least in part, through appeal to laws of nature? In fact, such failures are much more the rule than the exception. A person who fails to leave a locked room, after struggling with the door, fails to leave, in part, because of the laws of physics: were the laws such that less force needed to be applied to a steel lock to break it, the person would have left. Should we characterize the laws of nature as part of the agent's circumstances, and thus eliminating his opportunity to act, or should we, instead, say that whether or not the agent's intrinsic qualities (in our example, the strength of his muscles) qualify as *an ability* to act depends on the laws? For our purposes, it doesn't matter how this question is answered. The point is, simply, that the laws must either eliminate the agent's ability or opportunity, or else must account for the fact that the agent does not try to act. However we classify the laws—as ability-undermining, opportunity-undermining or trying-undermining (as it were)—our classification presupposes that (*) is both true and necessary.

Without saying what trying is, we can nonetheless say that, when interpreted correctly, (*) is both true and necessary. This provides us with a tool for determining if a particular agent S, at a particular time t1, is trying to perform some act C that, if completed, would be completed at time t2. To know if she is, we need to know that the following counterfactual is true:

(**) If (1) from t1 to t2 S has the ability and the opportunity to C and (2) S does not (at least until after t2) change his mind about doing any of the things that he is doing at t1, then S would C.

This counterfactual must be true if C is among the things that the agent is actually trying to do. In assessing this counterfactual, we follow a multi-step process: First, we consider possible worlds which are different from the actual in that whatever it was about the agent or his circumstances that accounts for his failure to act is different. In supplementing the circumstances so as to make them such as to provide opportunity, we imagine, for instance, that Duke (the man who met the undercover detective in a chat room, believing the detective

to be a 12-year old girl) was not apprehended by police after flashing his lights, and we imagine that the person whom he contacted in the chat room really was a 12-year old girl.³¹ We also supplement the defendant's intrinsic properties so that they amount to an ability to commit the crime. And we imagine that these changes are stable from the time that concerns us (namely t_1) to the time at which the act would be completed (namely t_2). Second, taking the set of all the worlds that are so supplemented, we narrow our consideration to that subset in which the defendant does not give up doing anything that he's started at or before t_1 . In Duke's case, we consider all those worlds in which he has ability and opportunity and says all the things in the chat room that he actually said, drives to the parking lot and flashes his lights, as he actually did, and, most importantly, continues to try to do all the things that he actually tried to do. Third, we ask if, in that remaining set of worlds, the defendant performs the crime. Since (*) is a necessary truth, failure to perform the crime in any of those worlds is definitive of a failure to actually try to act; and, conversely, performance of the crime in all of those worlds is definitive of trying to commit the crime.

That evidence of trying to act criminally is evidence of some sense in which *necessarily* the agent would have done so had she had ability and opportunity is the truth groped for, but not grasped, by those who take the culpability-constituting feature to which acts taken as means are supposed to attest to be *resolute* intention. The thought is that to be trying to act criminally it must not merely be the case that the defendant would have acted had he not encountered obstacles, or not lacked the ability; that might be true of the person who merely intends to act that way. The thought is that we need more than just this, we need more of a *guarantee* that the agent would really have acted criminally had he had the ability and opportunity than we have of the person with the criminal intention. The notion of

³¹ If Duke's intention was to molest an underage girl, then we supplement his circumstances accordingly; it is not important that the girl be 12, but only that she be underage. The point is that the content of the intention dictates how we are to supplement the defendant's circumstances and intrinsic properties in assessing (**).

“resolute” intention is supposed to provide that greater guarantee. But it provides it ineptly, for what is required is not merely that the defendant who was really trying is *that much more likely* to have acted criminally, had he had the ability and opportunity; what is required is that *necessarily* he would have acted criminally, under that condition.

The evidentialist justification of the requirement of substantial steps would be aided by an argument for the following general claim: If an agent took means to doing as she intended, then there is evidence for thinking that (***) is true. If this is so, then when an agent adopted means to doing as directed by her criminal intention, we have reason to believe that she’s trying to act criminally. After all, what else besides trying to act could combine with ability and opportunity so as to lead to action with necessity? Or, to put the point a slightly different way, if we have evidence that (***) is true, then we have evidence for thinking that something the agent did was such as to link with ability and opportunity in just the way in which trying to act so links; something the agent did plays the same distinctive functional role as trying to act. While this isn’t decisive evidence that the agent was trying to act criminally, it is very good evidence for that, for it seems highly unlikely that anything other than trying links with ability and opportunity in this way.

But what are acts performed as means to doing as one intends evidence of? They are, at least, evidence that the defendant was trying to perform *those acts* and was trying to perform them *as means*; they are evidence, that is, that the defendant tried to adopt means to furthering his criminal intention. The defendant didn’t just perform an act that was, in fact, a means to C; he performed the act *under the description*, “a means to C”, or for the reason that the act was a means to C. Buying a screwdriver might be a necessary means, on a certain occasion, to killing someone with it. Of course, one can buy a screwdriver for various innocent purposes, as well. However, if there is evidence that a particular defendant bought the screwdriver *for the purpose of killing someone with it*, then there is evidence that the defendant was trying to undertake means to killing someone. This might not be the same as trying to kill someone, but,

still, it is a first step: when a person *has* performed a means, there is evidence that he was *trying* to perform an act as a means to furthering a criminal intention. So, what's required is a link between trying to perform a means to C, on the one hand, and trying to C, on the other. If we could show that it's necessarily the case that an agent who is trying to perform a means is trying to perform the crime, then we would have shown that acts taken as means are evidence of attempt. And if we can then show that nothing other than trying to perform a means could be necessarily connected with trying to act, then we will have shown that nothing other than acts taken as means could provide adequate evidence of attempt; and then we will have justified the requirement of substantial steps on evidentialist grounds.

Now consider the following argument: It is irrational for a person who intends to perform an act, and who tries to take some means to the action, to fail to try to perform the other acts that he believes to be necessary means to it; such a person invests himself in a course of conduct, expending energy and resources to do so, but doesn't invest himself to the degree necessary to reach his intended end. Thus, a rational person, intending to reach an end, who tries to perform one means to it tries to perform all necessary means to that end. But a person who tries to perform all the necessary means to an end and also has the ability and the opportunity to reach that end, will reach it. After all, if he has the ability and opportunity to reach the end, he must have the ability and opportunity to take all necessary means. If he also tries to take those means, then he does all that's necessary for reaching his end. If he nonetheless fails to do so, it must be that he actually lacks the ability or opportunity, contrary to hypothesis. What this shows is that if we have evidence suggesting that a person who intends to act criminally is trying to take a means to the performance of the intended crime, we thereby, *ipso facto*, have evidence that he would have completed the crime had he had the ability and opportunity to do so. But then, given (*), we have evidence for thinking that he's trying to act criminally. What this shows is that actions taken as means to

satisfying a criminal intention provide a very special form of evidence that a rational defendant is trying to act criminally.

More formally, the argument just offered can be put like this, where t_1 is some time at or preceding the moment at which crime C would be begun and t_2 is the time at which C would be completed:

1. If (at t_1 ((S intends to C) & (S tries to M) & (S believes M to be a means to C and performs M for that reason) & (S is rational))), then at t_1 S tries to perform all the acts that are necessary means to C .
 2. If ((at t_1 S tries to perform all the acts that are necessary means to C) & (S has the ability and the opportunity to C from t_1 until t_2) & (S doesn't change his mind before t_2 and stop doing any of the things that he is doing at t_1)), then S C 's.
 3. [from 1 and 2] If ((S intends to C) & (at t_1 S tries to M) & (S believes M to be a means to C and performs M for that reason) & (S is rational) & (S has the ability and the opportunity to C) & (S doesn't change his mind before t_2 and stop doing any of the things that he is doing at t_1)), then S C 's.
 4. If (**) is true, then S is trying to C
- ∴ [from 3 and 4] If (at t_1 ((S intends to C) & (S tries to M) & (S believes M to be a means to C and performs M for that reason) & (S is rational))), then at t_1 S is trying to C .

If this argument succeeds, it brings us much of the way towards supporting an evidentialist justification of the requirement of substantial steps. The conclusion says that a rational defendant who has a criminal intention and was trying to adopt means to doing as he intended, was indeed trying to commit the crime. Further, if the premises of the argument are necessary truths (which is still to be shown), and the conclusion follows from the premises, then the conclusion is also a necessary truth. What this shows is that at the very least satisfaction of the requirement of substantial steps by a rational agent with a criminal intention is sufficient for trying to act criminally, even though acts taken as means to doing as one intends are nothing more than evidence of something else, namely of one's efforts to take such steps, efforts which are guaranteed with necessity (in a rational agent) to be found with an effort to complete the crime.

As noted, the argument does not show, full stop, that the agent who performs some means to doing as he intends is trying to commit the crime. Rather, it shows that this is true of *rational* defendants. Even assuming the argument is valid, does this additional requirement undermine its import? So long as the prosecution has the right to assume that the defendant is rational, then the answer is “no”. After all, assuming that the prosecution has shown the defendant to have the required criminal intention, what we need is a reason for thinking that in showing beyond a reasonable doubt that the defendant performed an act he believed to be a means to doing as he intended, the prosecution has thereby shown that the defendant was trying to commit the crime. If the prosecution can assume that the defendant is rational then the argument offered here gives us the reason we need. But does the prosecution have the right to assume that the defendant is rational? At least one reason why the prosecution has the right to assume this is that all that we can hope to use as evidence of the defendant’s mental state is the defendant’s behavior; but inferences from behavior to mental states are mediated by a presumption of rationality.³² Further, if the prosecution were required to establish the defendant’s rationality prior to drawing inferences about his mental states from his behavior, the prosecution would have to have some means of providing evidence of his mental states in order to show that those mental states comported with the defendant’s behavior in the way a rational agent’s mental states would. But this would require the prosecution to have some means of evidencing mental states distinct from the behavior of the defendant; an evident circle.

Premise 1 requires some clarification. To understand premise 1, it is important to distinguish between the following two claims:

- (a) $\forall m$ (m is a necessary means to C \rightarrow S tries to m).

³² Donald Davidson called this assumption the “principle of charity” or sometimes the “principle of rational accommodation”. See, for instance, Donald Davidson, ‘Radical Interpretation’ in *Inquiries into Truth and Interpretation*, Oxford: Clarendon Press, 2001, pp. 125–140; and ‘Three Varieties of Knowledge’ in *Subjective, Intersubjective, Objective*, Oxford: Clarendon Press, 2001, pp. 205–220.

(b) S tries to perform the following act: performing all the acts that are necessary means to C.

The consequent of premise 1 is (b), not (a). To see the difference between (a) and (b), consider an example³³: Say that S intends to rob a bank three months from now in New York. Right now S is trying to book a plane ticket to New York. S believes that booking the ticket is a necessary means to robbing the bank. S also believes that opening the bank's safe three months from now is a necessary means to robbing the bank, and this belief is true. And S is rational. It is manifest, however, that S is not trying *right now* to open the safe; after all, right now S is merely sitting at his computer browsing airline websites. And let's stipulate that, in fact, S will never try to open the safe: perhaps his plan will be foiled before he gets that far, or perhaps he'll give up his intention before the time to open the safe arrives. Thus, in this case (a) is false. Since the conditions specified in the antecedent of premise 1 are in place in this example, it follows that premise 1 would be false if (a) were its consequent. But notice that in the example (b) very well might be true: the fact that S is not right now trying to open the safe does not imply that he is not trying to perform all the acts necessary to robbing the bank. If someone were to ask S if he was trying to do everything necessary to rob the bank he might very well say that he was and the fact that he is not right now trying to open a safe would not invalidate his claim. Trying to do, at a particular time, everything that's necessary doesn't require trying to do, at that time, each of the individual acts that are necessary. The point here is really no more than the well-known point that co-referential action descriptions cannot be substituted for one another in sentences while preserving the truth values of those sentences; "is trying to —" is an opaque context. The sentence "S is trying to do all that is necessary to rob the bank" is true even though the sentence "S is trying to open the bank's safe" is false, and this is so despite the fact that

³³ Thanks to Michael Bratman for the example, which he offered, originally, as a counterexample to premise 1 when I had not yet formulated it in such a way as to avoid the ambiguity noted here.

among the acts *S* believes to be necessary to robbing the bank is the opening of the safe.

What this implies is that premise 1 is not a claim to the effect that a rational person who is trying to take one step towards doing what he intends is thereby committed to not changing his mind. Premise 1 would have that implication if (a) were its consequent. But, since (b) is its consequent, premise 1 has no implications whatsoever for what the rational agent will do after time t_1 , the time at which he is trying to undertake means to doing as he intends. Premise 1 can be thought of as a coherence constraint on the set of things which a rational agent is trying to do at a particular time. Just as it would be irrational for an agent to try to *A* while, at the same time, trying to refrain from *A*, it would be irrational for an agent to try to undertake a necessary means without also trying to undertake *all* the necessary means.

The success of the argument being offered turns on premises 1 and 2; if the argument is to succeed, both must be true, and both must be necessary truths. To see why premise 1 is true, consider, first, what the purpose of trying is. That is, consider what good is served by our being agents that have the capacity to try to do things. After all, it is a contingent fact that we do have that capacity. It's possible to engage in purposive behavior, end-directed behavior, without ever trying to do anything. Perhaps bees do this. Bees are not architects; they do not frame a conception of the hive before they build it and endeavor to bring it about that the world matches their conception of it. Yet they build hives. But human beings do this. We don't just plan; we also execute our plans. And the execution of our plans is itself an exercise of agency. Our plans don't simply translate themselves into conduct. Instead, we *try* to do as our plans direct. The point of trying is acting; we are trying creatures precisely because we are not the sort of creatures who, in many aspects of life anyway, will pursue ends except by actively doing so. Given the sorts of creatures we are, we need to have the capacity to try to do things if we are to reach our ends. What this suggests is that trying is an activity that is governed by certain norms of instrumental rationality. To be trying to do

something is to be subject to norms that, in general, need to be obeyed if one is to thereby succeed in acting. To take the obvious example already mentioned above: an agent who is trying to A is irrational if he is also trying to refrain from A-ing. The reason is because, generally, an agent who both tries to act and tries to refrain fails to act; since the point of trying to act is to act, an agent whose tryings are inconsistent in this way is irrational.

To reach premise 1, however, we need more than this. We need to show that the point of trying is generally undermined when an agent tries to take a means to an end without also trying to take all means to that end. What is the point of trying to take a means to an end? Part of the point is to, in fact, take that means. One of the goods that is served by trying to pull a trigger is that one, thereby, pulls the trigger. But the point of trying to do something under the description “a means to an end” is not just to do it; it is also to reach the end. Thus an agent who tries to take a means to an end—who endeavors, for instance, to pull a trigger not simply because he would be pulling a trigger, but also because pulling a trigger is one step in a course of conduct that he hopes will culminate in him killing another person—is under rational pressure to do all those things that he believes to be necessary to reach the end. But if he’s under rational pressure to do these things, he is also under rational pressure to try to do them. Premise 1 follows: a rational agent who performs one act as a means to his end tries to perform all the necessary means to his end. Since the argument for premise 1 just offered is purely conceptual, it follows that premise 1 is necessary.

To establish premise 2, we must first head off an objection that seems to show it to be false: Say that a person has false beliefs about what the necessary means are to reaching his end; he believes, for instance, that he can kill his boss by sticking pins in a voodoo doll. No amount of trying to do that is going to result in his killing his boss, even if he does have the ability and opportunity to kill him by, say, shooting him. Since what, precisely, a person who is trying to undertake all necessary means to an end directs himself towards doing depends on what he believes the

necessary means to his end to be—in our example, the man would direct himself towards sticking pins in a doll, not towards pulling a trigger—premise 2 is false of any agent with false beliefs about what means are necessary to doing as he intends. However, what this objection overlooks is that false beliefs can undermine one's abilities. False beliefs about what constitutes necessary means to an end can undermine one's ability to reach that end; hence, under the supposition that the agent has the ability to reach the end, we must imagine, also, that he does not have such radically false beliefs about what means are necessary as to prevent him from adopting the acts that are, in fact, necessary means.³⁴

The positive argument for premise 2 appeals to the concepts of ability and opportunity. It is obviously not always the case that performance of all the necessary means to an end is sufficient for reaching it: there might be some obstacle to reaching the end that the agent cannot hope to evade. However, in such a case the agent lacks either the ability or the opportunity to reach the end: if the obstacle derives from his lacking the appropriate properties, then he lacks the ability; if it derives from something in his circumstances, then he lacks the opportunity. However, if an agent has the ability and the opportunity to reach the end and yet fails to reach it must be because he didn't do one of the things that was necessary to reaching it; there must have been some necessary means that he failed to adopt. But if he tried to adopt that means, and yet failed, it must be because he lacked either the ability or the opportunity to adopt it. But if he lacked either of those, then he lacked the ability or opportunity to reach his end, contrary to hypothesis. Premise 2 follows, and, as in the case of premise 1, the

³⁴ Down this road lies an argument against criminalizing a certain class of radically impossible attempts. Someone who intends to kill someone while believing that voodoo will do the trick has beliefs that actually undermine his ability to reach the end by undermining his ability to perform acts that are actually likely to lead to completion of the crime; his beliefs commit him to using what resources he has on acts other than those that are actually necessary for doing as he intends. In such a case, possible worlds in which the agent has the ability to complete the crime are possible worlds in which he doesn't undertake the very acts that are to be construed as his substantial steps.

argument for it appeals only to the concepts invoked. What follows is that premise 2, also, is a necessary truth.

So, to summarize, acts taken as means provide evidence that the defendant, if he was rational, was trying to perform a means to the crime. The function of trying—namely, to lead to action—places a constraint on the set of things that a rational agent who is trying to perform a means to commit the crime is trying to do: he also tries to perform all the acts that are necessary means to committing it. But an agent who is trying to undertake that act—the act of performing every necessary means—and who has the ability and the opportunity to commit the crime, will commit it. It follows that substantial steps provide evidence for thinking that the agent would have committed the crime had he ability and opportunity and did not change his mind; they provide evidence for thinking (***) true. This is the mark of trying. Thus, steps taken as means to doing as specified by one's criminal intention provide evidence that the defendant was, in fact, trying to commit the crime. Since it is precisely that that we are criminalizing in criminalizing attempts, acts taken as means provide precisely the kind of evidence we need to justify conviction.

Recall that in the last section it was argued that while acts taken as means are construed as evidence of the culpability-constituting features of the completed crime, we lack an adequate justification of the requirement of substantial steps. The primary problem is that other things besides acts taken as means would serve as adequate evidence of those features and so it seemed that there was no acceptable justification for constraining the evidential route through which the defendant's guilt is to be established in the way that the requirement of substantial steps does. Does the proposed explanation for the evidential import of acts taken as means evade this concern? Yes. To see this, first imagine that some other fact F, distinct from the undertaking of a means M to the commission of the crime C, could provide evidence that the defendant was trying to M. F, then, would have to provide evidence that an instance of (***) is true; F would have to provide evidence, that is, that the agent would have performed M had he had, over the appropriate time interval, the ability and opportunity and not changed his mind. So, the question of what F could

be is just the question with which this section began, namely, what fact could serve as evidence that a person is trying to do something?³⁵ So the argument offered above for thinking that a person who is trying to undertake a means will succeed in reaching the end should he have ability and opportunity, cannot help someone who thinks that something other than acts taken as means could serve as evidence that a defendant is trying to commit the crime; such a person would still have to explain how there could be evidence of trying to act that didn't just amount to the taking of steps.

So, if something distinct from acts taken as means is to provide the same sort of evidence that substantial steps provide it must do so by providing evidence of something other than the trying to take means. And this other thing, whatever it is, must be assured with necessity to issue in action in an agent who has the ability and opportunity to commit the crime and does not change his mind over the relevant time interval. It's hard to imagine what this other thing could be if it is not another trying, a trying distinct from the trying to commit the crime. What else would intersect with ability and opportunity so as to necessitate action? If it is a trying distinct from the trying to commit the crime, then it would need to be linked to trying to commit the crime by virtue of the very concept of trying, as the trying to take means is so linked. But it seems unlikely that any other trying besides the trying to take means could be so linked. So, it seems that the crucial thing that we need to know when asking ourselves if an agent is trying to commit a crime is whether or not he was trying to take means to do so; and it seems that the only evidence we can hope to have that he was trying to do so is his actually doing so. Thus substantial steps don't merely provide the kind of evidence we need, it also seems very unlikely that anything else could provide the evidence we need.

³⁵ If F serves as evidence that the agent is trying to perform some means N to M, then it also serves as evidence that the agent is trying to perform C, since N is, necessarily, a means to C, given that it is a means to M and M is a means to C. But in that case, still, in meeting the need for evidence that the defendant is trying to C, the requirement of substantial steps is itself met.

It's important to explain how the solution proposed here differs from one of those rejected in section 1. Recall the proposal discussed in section 1 according to which the requirement of substantial steps is justified since trying is not merely a mental state but necessarily includes some kind of action in furtherance of the agent's aim. Thus, according to this view, since it is justified to criminalize attempts it is *ipso facto* justified to require that the guilty defendant has taken some means to doing as he intends. This approach was rejected on the grounds that the very problem that the criminalization of attempts was supposed to solve—it was supposed to provide for equal treatment of citizens who do not differ morally and whom there is no good reason to treat differently—was simply reintroduced. If trying necessarily involves action, and if what actions an agent performs frequently involves luck, then we should solve the problem of inequality not by criminalizing attempts, but by criminalizing something less than attempt that is nonetheless morally just as bad as the completed crime.

The view presented in this section does not require us to deny that trying necessarily includes action; in fact, it is probably true. But nor does the view presented here appeal to that fact in justifying the requirement of substantial steps. That is, according to the view presented here, the requirement of substantial steps is justified because the taking of means is probably the only good evidence we can have of trying to take means, and trying to take means is, given an allowable presumption of rationality, sufficient for trying to act criminally. Now it is possible that the taking of means *is* a component of trying to act criminally; perhaps there is no more to trying to act criminally than intending to and taking some means. But that fact plays no role in the justification of the requirement of substantial steps. Or, to put the crucial point in a different way, someone who appeals to the fact that trying includes action in order to justify the requirement of substantial steps cannot explain why the criminalization of attempts, given the denial of moral luck, doesn't result in morally equivalent people being treated differently by the state for no good reason; some people fail to act on their criminal intentions simply because of luck. But the view on offer can explain this. The difference

between those who fail to act on their criminal intentions and those who do act on them is that the latter, and not the former, provide us with the only kind of evidence that we can hope to have of their morally objectionable efforts. Since it is justified to treat defendants differently when we have different evidence of their guilt, the requirement of substantial steps is justified even if such defendants do not differ morally.

V. A SPECULATIVE SUGGESTION ABOUT INSUBSTANTIAL STEPS

In the first section of this paper I eluded to a famous problem faced by judges and legislators in their efforts to draw lines between those who attempt crimes, on the one hand, and, on the other, those who intend to act criminally, perform some acts that are necessary to doing as they intend, but who do less than is needed for attempt. The problem is ordinarily framed as one of deciding whether the defendant's actions were merely preparatory or, instead, constituted substantial steps towards performance of the crime. Was Duke's act of driving to the parking lot and flashing his lights a substantial step towards performing the sexual battery he intended, or was it less than that? This section offers a word about the implications for this issue of the evidentialist justification of the requirement of substantial steps defended here. Since there's a question as to whether the commonsense notion of "merely preparatory" acts maps precisely on to the category of "insubstantial steps"—acts taken as means to doing as one intends but which are insufficient for the act element of attempt—I will be using the term "insubstantial steps" and will be asking how the line is to be drawn between substantial and insubstantial steps.

The argument offered in Section 3 implies that *any act* undertaken as a means to an intended act constitutes an indispensable form of evidence to the effect that the agent is trying to perform the act he intends to perform. Acts expressive of criminal intention—such as verbal reports of one's intention—are not, ordinarily, undertaken as means to doing as one intends, and so such acts do not constitute the kind of evidence of attempt that acts undertaken as means provide; this is why they do not count

as substantial steps. However, acts expressive of intention are not, ordinarily, even insubstantial steps. Ordinarily, one does not need to express one's intent in order to do as one intends (exceptions are cases in which other agents need to know what one intends if they are to do their part in aiding one to accomplish one's intention). Thus, while we so far have an explanation for why acts expressive of intent do not amount to substantial steps, that same explanation does not extend immediately to explain why there are any insubstantial steps. Why isn't any act performed as a means to doing as one intends enough for the act element of attempt?

Perhaps the right response is that there are no insubstantial steps. In effect, to offer this response is to defend the "First Act" test under which a defendant has taken a substantial step just in case he has taken any step, has adopted any act as a means to doing as specified by his criminal intention. It is possible that this is correct. It is also possible that although the justification of the requirement of substantial steps does not rule out the approach implied by adopting the First Act test, there might be other independent reasons to prefer an approach that allows for some acts to be insubstantial steps. Perhaps, that is, there are reasons other than their evidential value for taking some acts taken as means to be insufficient for the act element of attempt. This also might be correct, and nothing in the argument offered so far suggests that it is not.

However, the argument of the last section of this paper does not force one to choose between these two options; there's a third. To see this, distinguish between the following two claims:

- (a) Performing an act that is a means to doing as one intends is evidence of trying to perform that means.
- (b) Performing an act under the description "a means to doing as I intend" is evidence of trying to perform that act.

Claim (a) is thrown into some doubt, although perhaps not invalidated, by cases in which agents perform acts that are means to doing as they intend, but who don't perform them for that reason, or under that description. Say that a man intends to commit murder using a screwdriver; he purchases a screwdriver.

His purchasing the screwdriver, let's stipulate, is necessary for him to do as he intends and is, in fact, a means to his doing so. But that doesn't mean that that's why he buys the screwdriver; he might buy it to work on his car. This doesn't show that (a) is false, for his buying the screwdriver might be *evidence* of his trying to take a means to killing someone with a screwdriver, even if it is defeasible evidence. Still, if it is shown that he buys the screwdriver *for the reason* that doing so is a means to killing someone with it—if it is shown that he performs the act under the description “a means to doing as I intend”—then that is far better evidence, perhaps even undefeasible evidence, that he is trying to undertake a means to doing as he intends. Here's the point: The argument of the last section of this paper showed that any act performed *as a means* to doing as one intends constitutes distinctive evidence that the defendant is trying to commit the crime. But it does not show that performing any act *that is a means* to doing as one intends constitutes such evidence; some acts that are means are better evidence that the agent is trying to take means than others, for there can be varying degrees of evidence that the agent is performing the act under the description “a means to doing as I intend”.

Perhaps, then, the substantial/insubstantial step distinction can be drawn by appeal to the degree to which the fact that the agent took an act that is a means to doing as he intends provides evidence that he was trying to take a means. In particular, perhaps the distinction should be drawn like so, where M is a means to doing as S intends:

M is a substantial step *if and only if* The fact that S performed M provides evidence beyond a reasonable doubt that S tried to take a means to C.

Under this suggested criterion, the fact that an act furthered more than one of the defendant's intentions—the fact that buying the screwdriver served both his end of fixing his car and his end of killing another person—could provide a reasonable doubt as to whether or not the defendant was trying to perform a means to doing as specified by his criminal intention. Such a doubt would then undermine the claim that the purchase of the screwdriver was a substantial step precisely because it would undermine the claim that the defendant was trying to take a

means to doing something criminal, and so would undermine the claim that he was trying to commit the crime in question. And, perhaps other things besides multiple reasons of this sort could also provide such reasonable doubt.

I am not entirely certain that this criterion for distinguishing substantial and insubstantial steps will survive scrutiny. The point for our purposes, however, is just this: since there is a difference between performing an act that is a means to doing as you intend, on the one hand, and performing an act *as* a means to doing as you intend, on the other, the argument of the last section allows room to distinguish among acts taken as means and so it neither commits one to the First Act test, nor to rejecting that test on grounds that do not appeal to the evidential value of acts that are means to doing as specified by the defendant's criminal intention. Some of those acts have greater evidential import than others for some evidence the fact that the defendant is trying to take means, and others do not.

VI. CONCLUSION

The elements of crimes are treated by the law as constitutive of those crimes; they are treated as necessary, and not just causally necessary, for performance of the crime. This is what makes an evidentialist justification of the requirement of substantial steps difficult, for acts undertaken as means are an element of the crime of attempt, and evidence is, except in rare cases, merely correlated with that which it is evidence of; evidence is not ordinarily necessary for the fact that it evidences. But the case of attempt, the case of trying, is not ordinary in this regard. Undertaking means to an intended act is evidence of trying to undertake such means, and a rational agent's trying to undertake means is sufficient for trying to perform the intended act. The requirement of substantial steps is justified precisely because of the evidence that substantial steps provide; in this case, the evidence is so closely related to the fact that it evidences, that it makes sense for the law to treat it as an element of the crime. And it is because this is so that the requirement of substantial steps can be justified even by those who deny that there is moral luck. Those, then, who deny moral luck should be happy with the standard doctrine of

criminal attempt; it enshrines in law deep features of our ordinary conception of trying and does so in a way which is consonant with the thought that there can be no moral difference where there is only a lucky difference.³⁶

School of Philosophy
University of Southern California,
Los Angeles, CA, 90089-0451,
USA
E-mail: yaffe@usc.edu

³⁶ Earlier drafts of this paper were presented at the Analytic Philosophy and Law conference at UCLA and to the Social Ethics and Normative Theory group at Stanford, both in April, 2006. I especially owe thanks to Larry Alexander, Michael Bratman, Claire Finkelstein, Stephen Finlay, Agnieszka Jaworska, Christopher Kutz, Barbara Herman, David Hills, Barbara Levenbook, Andrei Marmor, Connie Rosati, Seana Shiffrin and Ken Taylor.