

R. CLARKE

COMMANDING INTENTIONS AND PRIZE-WINNING DECISIONS

ABSTRACT. It is widely held that any justifying reason for making a decision must also be a justifying reason for doing what one thereby decides to do. Desires to win decision prizes, such as the one that figures in Kavka's toxin puzzle, might be thought to be exceptions to this principle, but the principle has been defended in the face of such examples. Similarly, it has been argued that a command to intend cannot give one a justifying reason to intend as commanded. Here it is argued that ordinary agents in ordinary cases can have justifying reasons for deciding that are not and will not be justifying reasons for doing what, in making those decisions, they come to intend to do. The paper concludes with some brief observations on the functions of decision-making.

When one makes a decision, one actively forms an intention to perform an action of a certain type. Some decisions are decisions to do something in particular at some nonimmediate future time; we may call these future-directed decisions. Others – present-directed decisions – are decisions to do something right away. In either case, the decision is itself an action, a basic intentional action of intention-formation.¹

One generally has reasons for the decisions that one makes as well as for intending as one comes to intend when one makes a decision. And one generally makes decisions and thus comes to intend for reasons.

Typically, reasons that one has for deciding to *A*, and reasons one has for intending to *A*, are reasons one has for *A*-ing. And the reasons for which one makes a decision are, likewise, typically reasons that one has to perform the action that one thereby comes to intend to perform. When I decide to take my car to the repair shop, the reasons that I have for so deciding, and the reasons for which I make this decision, are typically the reasons that I have to take my car to the shop.

Of course, one sometimes has a reason to make a decision on a certain matter at or by a certain time, where those reasons are not reasons favoring any specific decision on that matter over any alternative. I might find myself distracted by the question whether to attend a meeting this afternoon and want to settle the matter now so that I can get it out of my mind and focus on my work. This kind of reason for making a decision is not a reason to do what I decide to do. But it also need not be a reason favoring any specific decision on this matter. It is with reasons of this latter kind that I am henceforth concerned here.²

It is a widespread view that the reasons an agent has for deciding – and for intending – are, in every case, identical with the reasons the agent has for doing what, in making the decision in question and so coming to intend, she intends (or would intend) to do. Elizabeth Anscombe apparently endorses this view when she asserts that “grounds of intention are only reasons for acting” (1963, p. 90). Similarly, Donald Davidson maintains that

the reasons an agent has for intending to do something are basically of the same sort as the reasons an agent has for acting intentionally: they consist of both desires (and other pro-attitudes) and beliefs. If someone intends to polish his right shoe, it must be because there is some value he wants to promote by polishing his right shoe (perhaps he has already shined his left shoe and he wants the two to match), and he believes that by shining his right shoe he has a chance of promoting what he wants (1985, pp. 213–214).³

Berent Enç argues in his recent book on action that “the reasons an agent has for an action are the same as the reasons she has for willing, deciding, or intending to perform that action” (2003, p. 32).⁴ And Stewart Goetz defends the thesis that “an agent’s justification at any time t for then deciding to do action A (or deciding not to do A) later is identical with his justification at t for doing A (or not doing A) later” (1998, p. 206).⁵

Thomas Pink has rejected this view about the relation of reasons for deciding or intending and reasons for doing what one intends. He distinguishes between the psychological states

– he calls them reasons – that *motivate* one to decide or otherwise act and what *justifies* deciding or otherwise acting. Although one’s reasons for deciding (and thus coming to intend) are generally the very same psychological states that motivate one to perform the intended action, there are, Pink argues, cases where this is not so.⁶ And one’s justification for deciding (and thus coming to intend), he maintains, is never the same as one’s justification for doing what, in making the decision in question, one comes to intend to do (1991; 1993; 1996, ch. 8; 1998).

As Pink sees it, “agency is justified in terms of the likelihood of its furthering desirable ends” (1996, p. 137).⁷ What would justify some action *A* would be the likelihood that *A*-ing would further some desirable end. But since making a decision to *A* is itself performing an action, deciding to *A* is justified by the likelihood that *so deciding* will promote some desirable end (1993; 1996, p. 145).

Pink nevertheless argues that justifications for deciding are subject to the following restriction: “any end *E* that justifies deciding to do *A* must, supposing that decision is taken, also provide at least as much justification for doing *A*” (1996, p. 153; cf. Owens, 2000, pp. 81–82). Decision-making, he holds, serves to apply practical reason to subsequent action. “The whole point of taking decisions about which actions we shall perform... is to ensure that we end up performing the right actions” (Pink, 1996, p. 179).⁸ Decisions to act must then be motivation-perpetuating and rationality-preserving: rationally deciding to *A* must dispose one to *A* rationally thereafter (1996, p. 111–118). As he puts it, “the rationality of deciding to do *A* guarantees the rationality of doing *A* thereafter” (1996, p. 93), provided that there is no new relevant information.⁹ The rationality of deciding on a course of action could not provide this guarantee, he maintains, if the restriction on justifications for deciding did not hold (1996, p. 153).¹⁰

Pink regards an action as rational just in case it is “motivated by desires for ends which, through their desirability, also sufficiently justify the action” (1996, p. 140). Along these

lines, one might say that a justifying reason for performing a certain action is a desire for an end that is likely to be furthered by that action and which, through its desirability, provides some degree of justification for performing that action. Then, despite Pink's rejection of the widespread view about reasons for deciding, he apparently agrees with those who hold that view on the following restrictive thesis: any justifying reason for deciding must be, once the decision is made, a justifying reason for doing what one, in making that decision, comes to intend to do.¹¹ Indeed, since he maintains that all intentions are actively formed in decision-making, Pink seems committed, as are holders of the widespread view, to the following, more comprehensive, restrictive thesis:

(R) any justifying reason for deciding to *A*, and any justifying reason for coming to intend to *A*, must be, once the decision is made or the intention acquired, a justifying reason for *A*-ing.

Insofar as they recognize any practical reasons as justifying reasons, the several writers I have identified all appear to be committed to **R**. I argue here that this restrictive thesis is mistaken. An agent can have a justifying reason for making a certain decision, or for acquiring a certain intention, that is not and will not be a justifying reason to do any of what the agent thereby comes (or would come) to intend to do.

My argument against **R** takes desires (of a certain sort) to be justifying reasons. This is the view of justifying reasons that is suggested by Pink. There is, of course, a wide variety of competing views about what justifying (or normative) practical reasons are. I do not mean to commit myself here to any particular view on this matter. I have framed the discussion in terms of Pink's view for the following reasons: he has set out the most qualified version of the restrictive thesis against which I am arguing; he has offered a rather extensive discussion of the issues surrounding this thesis; and, consequently, I want to make it clear that my argument applies to the thesis understood in terms of his conception of justifying reasons.

But it should be evident, as I proceed, that the cogency of the argument does not depend on this particular view of justifying reasons. A key element of my argument is the presentation of certain examples (or situations, or cases); and these should convince proponents of any of several widely held views of justifying reasons that **R** is mistaken. The argument should be acceptable to those who take justifying reasons to be mental states of certain types, or rational mental states of those types, or the contents of such states, or states of affairs, or aspects of situations.¹²

I shall close the paper with a brief comparison of my case against **R** with those of some other writers who at least implicitly reject it, and with a few observations on the point of deciding.

1. PRIZES AND COMMANDS

Defenders of thesis **R** – at least those holding standard views of justifying reasons for actions that carry out prior intentions – must take a certain stand on Gregory Kavka's (1983) toxin puzzle. Imagine that an eccentric billionaire offers you a million dollars for intending by midnight tonight to drink a toxin tomorrow afternoon. The toxin will make you painfully ill for a day but will have no lasting harmful effects. If you have the indicated intention by midnight – the billionaire will be able to tell whether you do – the prize will be awarded tomorrow morning, and it will then be yours whether or not you subsequently drink the toxin. The puzzle concerns whether it would be rational to acquire the intention in question and, indeed, whether you would be able to do so.¹³

Given the alleged restriction on justifying reasons for intending, your desire to win the prize is a justifying reason for you to come to intend to drink the toxin only if (once you so intend) it is also a justifying reason for you to drink the toxin. But since, if you have the intention by midnight, the prize will already have been awarded before the time for drinking the toxin arrives, your desire to win the prize is not and will not be a justifying reason for you to drink the toxin.

Defenders of restriction **R** typically hold, then, that your desire to win the prize is not a justifying reason for you to intend – or to acquire the intention – to drink the toxin. And more broadly, they tend to hold that desires for prizes offered for intending – prizes that can be won regardless of whether one does what one intends – are never justifying reasons for intending. (See Pink, 1996, pp. 147–159; Owens, 2000, pp. 81–82; Enç, 2003, pp. 33–34).¹⁴

In defending the alleged restriction, Pink (1996, pp. 154–155) also rules out the justification-giving status of any command to decide or intend to perform a certain action. A command such as “Decide to raise your arm” or “Intend to raise your arm,” he maintains, makes no sense.¹⁵ It cannot provide you with a justification for making that decision or acquiring that intention. This is so, he holds, because a decision or intention command cannot, once the command is obeyed, justify performing the action that, in obeying the command, one comes to intend to perform.

Prizes for deciding or intending can justify so deciding or intending; desires to win such prizes can be justifying reasons for so deciding and intending. And a command to decide or intend can make sense; it can provide one with a justification for doing what one is commanded to do.

2. WINNING A DECISION PRIZE

Suppose that Beta offers you a prize of a million dollars for doing tomorrow at noon whichever of *A*-ing or refraining from *A*-ing that you decide today to do then. Apart from the prize offer, neither *A*-ing tomorrow nor refraining from *A*-ing tomorrow is either especially promising or especially onerous. (*A*-ing might be something like raising your arm.) Alpha now offers you a prize of the same amount for deciding today to *A* tomorrow at noon. If there are no other relevant considerations, you should be able to win both prizes.

Since you have ordinary desires concerning million-dollar prizes, Beta’s offer gives you a justifying reason to make some decision today with respect to *A*-ing tomorrow at noon

and then to carry out that decision. You can win Beta's prize by deciding today to *A* tomorrow at noon and then *A*-ing tomorrow at noon. Or you can win Beta's prize by deciding today to refrain from *A*-ing tomorrow at noon and then refraining from *A*-ing tomorrow at noon. Alpha's offer gives you a justifying reason to start on the first route to winning Beta's prize, for by taking that route you will win twice as much. With no other relevant considerations in the mix, it would be foolish – irrational – of you not to take advantage of the offers and win both prizes.

Alpha's offer gives you no justifying reason to *A* tomorrow at noon, since, if you decide today to *A* tomorrow at noon, Alpha's prize will be awarded whether or not you *A* tomorrow. But if you decide today to *A* tomorrow at noon, then you will have a strong justifying reason to *A* tomorrow – viz., your desire to win Beta's prize. Your desire to win Alpha's prize is a justifying reason for you to decide today to *A* tomorrow at noon, it is a justifying reason for making this decision that is not and will not be a justifying reason for *A*-ing tomorrow, and you can make the decision (in part) for this reason.

Suppose that a few days ago your friend Charles received a similar pair of offers from Alpha and Beta. Charles won Beta's prize by deciding to refrain from *A*-ing and then subsequently so refraining. Curious, you ask him why he did not decide to *A* and subsequently *A*. Did he not see gaining the additional million dollars as a desirable end? "Sure," he replies, but he did not see that the likelihood that deciding to *A* would promote that end justified so deciding. As he sees it, it was a rationally indifferent matter which route he took to winning the prize offered by Beta. He flipped a coin. Charles is a strange fellow – and a million dollars poorer than he ought to be.

Consider the objection that it is not Alpha's offer, but rather the combination of Alpha's and Beta's offers, that provides the justification for deciding to *A*. It is not the likelihood that deciding to *A* would win you a million dollars

that justifies your so deciding; it is the likelihood that deciding today to *A* and then *A*-ing tomorrow at noon would win you two million dollars that provides the justification. And the desirable end here will be achieved only if you do tomorrow what you decide today to do tomorrow. The only justifying reason you have for deciding today to *A* tomorrow, then, will be, once you so decide, a justifying reason to do what you have thereby come to intend to do.

Undoubtedly, the combination of offers provides a justification for your deciding to *A*. But it favors your so deciding only because Alpha's offer provides a justification for that decision. Beta's offer does not itself favor your deciding to *A*. The combination does. The combination would not favor your deciding to *A* if Alpha's offer did not provide a justification for your so deciding.

3. GETTING BACK TO WORK

Few of us are so lucky as to be presented with a pair of prize offers like those from Alpha and Beta. But it is not so extraordinary to be in a situation with an analogous structure of practical reasons.

Suppose that Ann is distracted by the question whether to attend a meeting this afternoon. She wants to settle the matter now so that she can get her mind off it and get some work done. She finds it unpleasant to attend meetings of the sort that is planned, but she is highly curious about who will say what at this one. As she sees the matter (and we may suppose she is right about this), she does not have better reason favoring either alternative. But she expects that if she decides now to go to the meeting, that decision will effectively put the matter out of her mind, allowing her to get something done in the time between now and the meeting, whereas if she decides not to attend the meeting, she might well be distracted throughout the day with thoughts about what she will miss at the meeting. Ann decides to attend the meeting. One justification that she has for so deciding is that the decision is more likely than the alternative to further the end of getting

some work done prior to the time of the meeting. But even when Ann has so decided, this end is certainly not something that will be furthered by doing what, in making this decision, she has come to intend to do – viz., attend the meeting. Since Ann desires this end, she has a justifying reason for deciding to attend the meeting that is not and will not be a justifying reason for attending the meeting.

Note that in making this decision, Ann does not come to intend to get some work done in the time prior to the meeting. She has had that intention all along; there has been no uncertainty in her mind about what to do between now and meeting time. Thus, the end of getting back to work, while something that will be furthered by Ann's decision, is not an end that will be furthered by doing anything that Ann, in making this decision, comes to intend to do.

In arguing against the view that justifications for deciding to *A* are identical with justifications for *A*-ing, Pink appeals to the role that future-directed decisions play in coordinating one's behavior over time. In most cases, the coordinating function of future-directed decisions justifies one in making some decision or other on a certain matter, without justifying any specific decision on that matter. But in some cases, the goal of coordinating one's behavior over time justifies a specific decision without, until that decision is made, providing any justification for the action decided upon. In a case of this sort, one will have, at some time *t*, a justification for making a certain decision that is not, at *t*, a justification for doing what one would thereby decide to do.

Pink imagines a stuntman, Dan, who must decide now whether to attempt a certain stunt in six months. Dan very much wants to avoid attempting the stunt without first having publicized it, and he wants as well to avoid not attempting the stunt having publicized it. At present, due to a recent mishap, Dan is risk-averse and so disinclined to attempt the stunt. But he realizes that as time passes, his risk aversion will wear off. There is a significant possibility that, even if he now decides not to attempt the stunt, when the time comes,

he will attempt it. In contrast, he is sure that if he now decides to attempt the stunt, he will do so. Wanting to avoid a mismatch between prior publicity and subsequent action, and realizing that his decision will determine whether he publicizes the stunt, Dan decides to make the attempt. The end of avoiding a mismatch is more likely to be furthered by – and so justifies – this decision. But until he makes the decision, that end is no more likely to be furthered by Dan's attempting the stunt than it is likely to be furthered by his not attempting the stunt (1991, pp. 351–353; 1993, pp. 330–331; 1996, pp. 234–236).¹⁶

My argument against restrictive thesis **R** appeals to another function of future-directed decisions: making such a decision generally allows one to settle some practical question and then turn one's attention elsewhere, devoting limited mental resources to other matters. In many cases, as I noted earlier, this function of decision-making provides one with a justification for making some decision or other on a certain matter without justifying any specific decision on that matter. But in some cases, the goal of settling some practical matter so that one can turn one's attention elsewhere justifies one's deciding that matter in a specific way, without ever justifying one's doing what one thereby decides to do. Ann's case is one of this sort.

4. OBEYING A DECISION COMMAND

A brain researcher seeking to observe the neural realization of your intention to raise your arm might prompt the intention simply by telling you to raise your arm. When you intentionally raise the arm, you will have intended to do so.

A command of this sort makes perfect sense. But, of course, it is not a command to decide or intend.

Suppose now that the researcher wants to see whether there is a difference in neural realization between present-directed and future-directed intentions. She tells you to decide now to raise your arm tomorrow at noon. As luck would have it, Beta had earlier offered you a million dollars for

doing tomorrow at noon whichever of raising your arm or refraining from raising your arm that you decide today to do then. You had your stint as research subject coming up, so you put off making the decision. Now it strikes you that you can both win the prize and give the researcher what she wants. You decide to raise your arm tomorrow.

If you have a general justification for complying with the researcher's commands, and if you rationally desire to do so, then it is not a rationally indifferent matter which route you take to winning Beta's prize. The rational thing for you to do is to win it by deciding now to raise your arm tomorrow at noon and then doing so. Your making that decision is rationally favored because only your so deciding will further the desired end of complying with the command. The likely furtherance of that end justifies your deciding to raise your arm but will not, even after your decision is made, justify your raising your arm tomorrow. Given your desire to cooperate with the researcher, you have a justifying reason to decide now to raise your arm tomorrow at noon, one that is not and will not be a justifying reason to raise your arm tomorrow.

5. SOME OTHER REJECTIONS OF **R**

I have argued that restriction **R** is mistaken. I am not the first to advance the view that one can have independent justifying reasons for intending and deciding – justifying reasons that are not and will not be justifying reasons for doing what one thereby intends (or would intend) to do. My case against **R** is, however, different in some important ways from others that have been advanced.

For one, in the examples presented here as undermining **R**, it is rational for the agent to do what she intends to do, and she has adequate justification for doing so, even though she has some further justification for intending that is not and will not be justification for doing what is intended. These features should make these cases less controversial – and should make them more effective in undermining **R** – than are some other cases that have been thought to count against such a thesis.

For example, Kavka (1978 [1987]) argues that, in certain narrowly characterized situations, an agent would be justified (by the end of deterrence) in forming the conditional intention to retaliate if she is herself subjected to a first-strike, even though she would not be justified (and she knows she would not be justified) in carrying out the intended retaliatory attack. (Implicitly rejecting **R**, he takes the likely furtherance of the end of deterrence to provide the agent with justification for forming the intention to retaliate if attacked, even though carrying out that intention will not further that end.)¹⁷ Kavka (1984) judges that it would be rational for the agent to form the intention but not rational for her to carry it out. David Lewis (1984) concurs with these assessments.

Defenders of **R** often respond with a denial that the rationality of forming an intention and the rationality of carrying it out can come apart in this way, or that “there could ever be sufficient justification for forming an intention to act which, thereafter, there is no justification for executing” (Pink, 1996, p. 177, note 3). No such denial is relevant to the cases I present here, since in these cases, the agents have ample justification for doing what they come to intend to do, and it is quite rational for them to do so.¹⁸

Second, it is at best doubtful that, in the kind of deterrence case presented by Kavka, a rational agent *can* form the intention in question – doubtful, at least, that a rational agent can do this without first changing or expecting to change in certain ways. (Of course, the same doubt arises with regard to the toxin case.) For the agent fully believes that she will have overwhelming reason not to carry out the intention, and (assuming that she expects to remain rational and to retain this assessment of her reasons) she will thus believe that she will not carry it out even if the circumstances for doing so (a first-strike) arise. Arguably, she cannot both intend to retaliate if attacked and believe that she will not do so. In contrast, rational agents in the cases I have presented should be able to form the intentions in question.

Alfred Mele (1995) rejects **R** in discussing a case, similar to one of mine, in which an agent is presented with two offers of prizes for intending from two eccentric billionaires. (The offers differ from those presented to the agent in my case.) Like my case, Mele's is of a sort that no one is ever likely really to be in. However, my case involving Ann's decision about whether to attend the afternoon meeting shows that there are quite ordinary cases with a similar structure of justifying reasons.¹⁹

6. THE FUNCTIONS OF DECISION-MAKING

Sensible decision commands, though possible, are not common. But prizes specifically for deciding – in a broad sense that includes a decision's likely furtherance of a desirable end that will not be furthered by the agent's doing what, in making that decision, she would come to intend to do – decision prizes of this sort may not be so rare. Such prizes can justify decisions. Ordinary agents in rather ordinary cases sometimes have justifications for deciding that are not and will not be justifications for doing any of what, in making the decisions in question, they come (or would come) to intend to do. Given desires for such prizes, these agents have justifying reasons for deciding that are not and will not be justifying reasons for doing what they thereby intend (or would intend) to do.

Does this conclusion imply that decision-making does not serve to apply practical reason to subsequent action? It implies, at least, that deciding can serve some purpose distinct from this one.

Deciding is, essentially, actively coming to be committed to or settled on performing a certain action. Being so committed is being disposed to do what one is committed to do. Typically, at least part of one's point in committing oneself to some specific course of action is ensuring that one will, in doing what one decides to do, do what one has best, or better, or good enough, reason to do. Deciding, then, typically applies practical reason to one's subsequent behavior. But making a decision sometimes has some other point as well.

An intention to perform a certain action at a later time might, as Kavka observes, “produce effects independent of its actually being carried out” (1984, p. 155). In some cases, the production of such effects may be at least part of the agent’s point in forming the intention. Ann’s case, in which deciding to attend the meeting has the effect of enabling her to get back to work, is one illustration of this phenomenon.

If by the function of deciding we mean the purpose for which the capacity to make decisions was favored by natural selection, then it seems likely that deciding has no single function. But Pink apparently takes the function of decision-making to concern what legitimate point an agent may have in making a decision, or what can justify deciding. To this question, too, we should expect a pluralist answer. Ends that can justify acting as decided play the lead role. But we do not have a one-man show.²⁰

NOTES

¹ On decisions as basic intentional actions, see, for example, Mele (1997, 2000) and Pink (1996). Not all intentions are actively formed in decision-making. One makes a decision only when there is some question in one’s mind about what to do. But there is no such question in some cases of intentional action. With routine or habitual actions – such as opening the door to one’s car – the intention that one executes in acting may be non-actively acquired, arising from one’s beliefs and desires without one’s having actively to form it. On this point, see Audi (1973, pp. 394–395), Bishop (1989, pp. 181–182), and Mele (1992b, p. 141).

² Bratman (1987, p. 103) and Pink (1991, p. 350) similarly distinguish between reasons for making some decision or other on a certain matter and reasons for deciding the matter a certain way.

³ The Anscombe and Davidson passages are cited by Pink (1991, pp. 344–345).

⁴ In arguing for this view, Enç (2003, pp. 30–37) aims to cast doubt on volitional theories of intentional action, views according to which (i) every intentional action either is or begins with a basic mental action such as a decision, a volition, or an act of will, and on which (ii) these basic mental actions cannot be analyzed in terms of event causation. While I argue here against Enç’s view about reasons for deciding and intending, my conclusions provide no support for either (i) or (ii).

⁵ Farrell (1989, p. 287) may also endorse this view when he asserts that “an ideally rational individual would commit herself to future action on the basis of her estimates of how reasonable the relevant actions will be at the time of action rather than on the basis of her estimates of the utility to her, at the time of forming that intention, of its being the case that she intends to perform those actions.”

Pink (1991, p. 344; 1996, pp. 141–142) takes Bratman to hold the widespread view as well. Bratman (1987, p. 103) does maintain that “what we deliberate *about*, when we deliberate about the future, is what to do then. This means that in such deliberation about the future the desire-belief reasons we are to consider are reasons for various ways we might *act* later.” However, in the discussion in which this claim appears, he seems to allow that an agent might have reasons for intending that would not be reasons for doing what is intended. The quoted passage simply asserts that no such reasons would be considered in deliberation about the future action. Bratman (1998, pp. 62–63) subsequently allows that reasons of this type can be considered in deliberating about future actions.

⁶ In a later section of this paper, I describe the main case employed by Pink in arguing for this claim.

⁷ “By furthering an end I mean something like raising the objective probability of the end’s attainment. This might be through causing the end’s attainment – but end furtherance need not involve causation. Doing *A* might further end *E* by helping *constitute* the attainment of *E*: attaining *E* might, for example, consist simply in *A*’s being done.... By ‘likely’ I mean what the evidence suggests, or what is epistemically as opposed to objectively probable” (Pink, 1996, p. 139).

⁸ Pink (1996, pp. 70–71) follows Bratman (1987) in drawing attention to the ways in which decision-making allows a persisting agent to coordinate her behavior over time and to the benefits of making decisions in advance, given our bounded rationality and intellectual resources. Pink sees these roles as subordinate to the main function of decision-making to ensure our rational unity through time, a function that he expresses with the statement quoted in the text here.

⁹ “If a decision to act is taken rationally, then, in the absence of new information which warrants revising the assumptions on which the decision was based, the action which executes the decision will be rational too” (1996, p. 93).

¹⁰ Goetz (1998, p. 205) also appeals to rationality-preservation in arguing for the stronger thesis that he holds, one that identifies justification for deciding to *A* with justification for *A*-ing.

¹¹ The restriction would be thought to hold *provided* that the desire in question – the justifying reason – survives (or would survive) the making of the decision. A similar proviso applies to the restriction **R** that follows in the text.

¹² My argument against **R** does assume that one possible view of justifying reasons is incorrect. Let us say that a “course of action” is a decision and a subsequent action (or series of actions) that carries out that decision. Some writers (e.g., Gauthier, 1994, 1998a, b) hold that the rationality of at least some decisions and the actions that carry them out is a matter not of their individual expected consequences but of the expected consequences of courses of action comprised of them.

Now, this is a view about the rationality of (at least some) decisions and decided-upon actions, and not explicitly about justifying reasons. But consider the following further thesis: something is a justifying reason for deciding to perform a certain action, or for performing a previously decided-upon action, only insofar as that something is a justifying reason for engaging in the course of action that includes the decision or subsequent action in question. Plainly, given such a view of justifying reasons, **R** comes out correct.

Gauthier apparently does not hold this further thesis (see, for example, his 1998a), nor do some other writers (e.g., Robins, 1997) who advance similar views of practical rationality. And none of the figures identified in the text here as committed to **R** endorses it.

¹³ The billionaire’s offer rules out your providing yourself with external incentives to drink the toxin, such as by hiring someone to break your legs if you fail to drink it. Gimmicks such as hiring a hypnotist to implant the intention or forgetting key facts about the situation are also disallowed.

¹⁴ Owens’s claim is stated in terms of justification: “a prize for taking the decision to consume a toxin will not justify that decision because it will not justify the action decided upon” (2000, pp. 80–81). But with justifying reasons characterized in the way suggested by Pink, it seems likely that Owens would accept the claim in the text regarding desires to win decision prizes.

¹⁵ O’Shaughnessy (1980, p. 300) similarly holds that there is no such command as “Decide to raise your arm.” He takes the alleged nonexistence of decision commands to show that forming an intention – deciding – is never an action. I take it for granted that deciding is a basic intentional action. On this point, I am in agreement with most defenders of the restrictive thesis.

¹⁶ Recall (note 7 above) that the relevant notion of likelihood is epistemic. Which action will avoid a mismatch depends on which decision Dan makes, and we are to imagine that prior to his making the decision, the available evidence does not settle the question of which decision he is more likely to make.

¹⁷ Similarly, as Kavka sees it, the agent in the toxin case has a reason to intend to drink the toxin but has no reason to drink it. (Gauthier [see note 12 above] rejects this last claim; so does McClennen (1990, pp. 227–231). For replies, see Bratman (1998) and Mele (1996)).

¹⁸ Another feature of Kavka's deterrence case is that the intention involved is conditional: it is an intention to retaliate if attacked. Moreover, it is what Kavka calls self-stultifying: if it fulfills the agent's purpose, "it ensures that the intended...act is not performed, by preventing the circumstances of performance from arising" (1978 [1987], p. 20). As my cases (and others) plainly show, reasons for rejecting **R** are not found only in cases involving intentions of these sorts.

¹⁹ Another case of Mele's (1992a) that, as he sees it, counts against **R** involves an agent who has a couple of highly unusual features: he is cursed to drink – either intentionally or unintentionally – any liquid toxin in his vicinity, and his intentions always persist unless he either actively drops them or carries them out. My cases, like others of Kavka, Lewis, and Mele, show that there may be independent justifying reasons for deciding or intending in cases with quite ordinary agents.

²⁰ I wish to thank Alfred Mele, Thomas Pink, Sarah Wright, and an anonymous referee for this journal for their generous comments on drafts of this paper. Versions of the paper were presented at the University of Georgia, a conference on Action and Agency at the University of Florida (February 2005), the 2005 Pacific Division Meeting of the American Philosophical Association, and the 2005 Annual Conference of the Australasian Association of Philosophy; I am grateful to audience members for their comments.

REFERENCES

- Anscombe, G.E.M. (1963): *Intention*, 2nd ed. Ithaca: Cornell University Press.
- Audi, R. (1973): 'Intending', *Journal of Philosophy* 70, 387–403.
- Bishop, J. (1989): *Natural Agency: An Essay on the Causal Theory of Action*, Cambridge: Cambridge University Press.
- Bratman, M. (1987): *Intention, Plans, and Practical Reason*, Cambridge, Mass: Harvard University Press.
- Bratman, M. (1998). 'Toxin, Temptation, and the Stability of Intention', in J.L. Coleman and C.W. Morris (ed.), *Rational Commitment and Social Justice: Essays for Gregory Kavka*, Cambridge: Cambridge University Press, pp. 59–83.
- Davidson, D. (1985). 'Replies to Essays I-IX', in B. Vermazen and M.B. Hintikka (eds.), *Essays on Davidson: Actions and Events*, Oxford: Clarendon Press, pp. 195–229.
- Enç, B. (2003): *How We Act: Causes, Reasons, and Intentions*, Oxford: Clarendon Press.

- Farrell, D.M. (1989): 'Intention, Reason, and Action', *American Philosophical Quarterly* 26, 283–295.
- Gauthier, D. (1994): 'Assure and Threaten', *Ethics* 104, 690–721.
- Gauthier, D. (1998a). 'Intention and Deliberation', in P.A. Danielson (ed.), *Modeling Rationality, Morality, and Evolution*, Oxford: Oxford University Press, pp. 41–54.
- Gauthier, D. (1998b). 'Rethinking the Toxin Puzzle', in J.L. Coleman and C.W. Morris (eds.), *Rational Commitment and Social Justice: Essays for Gregory Kavka*, Cambridge: Cambridge University Press, pp. 47–58.
- Goetz, S. (1998): 'Reasons for Forming an Intention: A Reply to Pink', *Mind* 107, 205–213.
- Kavka, G.S. (1978) [1987]. 'Some Paradoxes of Deterrence', *Journal of Philosophy* 75, 285–302. Reprinted in his *Moral Paradoxes of Nuclear Deterrence*. Cambridge: Cambridge University Press, pp. 15–32.
- Kavka, G.S. (1983): 'The Toxin Puzzle', *Analysis* 43, 33–36.
- Kavka, G.S. (1984). 'Deterrent Intentions and Retaliatory Actions', in D. MacLean (ed.), *The Security Gamble: Deterrence Dilemmas in the Nuclear Age*, Totowa: Rowman & Allanheld, pp. 155–159.
- Lewis, D. (1984). 'Devil's Bargains and the Real World', in D. MacLean (ed.), *The Security Gamble: Deterrence Dilemmas in the Nuclear Age*, Totowa: Rowman & Allanheld, pp. 141–154.
- McClellan, E.R. (1990): *Rationality and Dynamic Choice: Foundational Explorations*, Cambridge: Cambridge University Press.
- Mele, A.R. (1992a): 'Intentions, Reasons, and Beliefs: Morals of the Toxin Puzzle', *Philosophical Studies* 68, 171–194.
- Mele, A.R. (1992b): *Springs of Action: Understanding Intentional Behavior*, New York: Oxford University Press.
- Mele, A.R. (1995): 'Effective Deliberation About What to Intend: Or Striking It Rich in a Toxin-Free Environment', *Philosophical Studies* 79, 85–93.
- Mele, A.R. (1996): 'Rational Intentions and the Toxin Puzzle', *ProtoSociology* 8/9, 39–52.
- Mele, A.R. (1997): 'Agency and Mental Action', *Philosophical Perspectives* 11, 231–249.
- Mele, A.R. (2000): 'Deciding to Act', *Philosophical Studies* 100, 81–108.
- O'Shaughnessy, B. (1980): *The Will: A Dual Aspect Theory, Part 2*, Cambridge: Cambridge University Press.
- Owens, D. (2000): *Reason Without Freedom: The Problem of Epistemic Normativity*, London: Routledge.
- Pink, T. (1991): 'Purposeful Intending', *Mind* 100, 343–359.
- Pink, T. (1993): 'Justification and the Will', *Mind* 102, 329–334.
- Pink, T. (1996): *The Psychology of Freedom*, Cambridge: Cambridge University Press.
- Pink, T. (1998): 'Reply to Goetz', *Mind* 107, 215–218.

Robins, M.H. (1997): 'Is it Rational to Carry Out Strategic Intentions?',
Philosophia 25, 191–221.

Department of Philosophy
Florida State University
151 Dodd Hall
Tallahassee, FL 32306-1500
USA
E-mail: rkclarke@fsu.edu