THREE WAYS OF SPILLING INK TOMORROW

LUCA FERRERO

University of Wisconsin at Milwaukee / Stanford University

ABSTRACT

Many traditional accounts of the nature and role of future-directed decisions fail to do justice to them by conflating them with either merely causal or manipulative forms of distal self-control. More promising views are in terms of either non-manipulative decision-based reasons or non-manipulative psychological propensities of non-reconsideration. The problem with the latter views is that they are not sufficiently comprehensive, since they focus exclusively on one among the many mechanisms of non-manipulative diachronic agency. I argue that a principled account of the role of genuine future-directed decisions must acknowledge that decisions need not necessarily be effective in influencing future conduct. The truly distinctive feature of the diachronic agency associated with genuine decisions is the reliance on the temporal stability of the reasons in support of the decided upon action. This stability underlies our entitlement to the memory of prior deliberations, which is often the primary source of the non-manipulative distal effectiveness of future-directed decisions.

Key words: diachronic agency, future-directed decisions, intentions, memory.

1. Introduction

1.1. When an agent makes a decision to do φ at a future time *f*, it is expected that, everything else being equal, she will do φ at *f* and do so, at least in part, because she has earlier decided to φ at *f*. What is the exact nature of the difference that the agent's future-directed decision appears to make in her future conduct? The answer would be straightforward if decisions were simply the triggering events of mere causal mechanisms that reliably force the agent to φ at *f*. A decision, however, does not appear to work like the igniting of a fuse that will later cause the explosion of a bomb.[1] It rather seems that an agent can make herself do φ at a later time—say, spilling ink tomorrow—

---

[1] The contrast between future-directed decisions and lit fuses is due to Velleman (1997), 47. Reprinted in Velleman (2000).

just by making a future-directed decision to do so. As J.H. Sobel writes, "*By dint of sheer will*, a person may be able to 'tie her hands' and make necessary and inevitable some action… In order to be sure that she does not answer the door, a person could, instead of handcuffing herself to the stove, do that sort of thing 'all in her head'".[2] The ability to tie one's own hands 'mentally' merely by an act of will appears to be a wizardly power, yet we use it quite frequently and with remarkable success. As Jerry Fodor once remarked, no one would ever get his feet wet if we were as good at predicting the weather as we are at predicting future behavior simply on the basis of the declaration of intentions, even if "the etiology of the weather must surely be child's play compared with the causes of behavior".[3]

1.2. Despite its ubiquity and reliability, distal self-control via future-directed intentions is a puzzling phenomenon. To begin with, the mere passage of time appears to void the force of the agent's decision, even when he is not weak-willed. Sartre made this point vivid in the famous case of the gambler. Sartre's gambler, who in the past has resolved not to gamble anymore, is once again presented with the opportunity to gamble. When this happens he realizes that his past resolution, just because it is a *past* one, is no longer exerting force on his present action. In the gambler's own words:

> It seemed to me that I had established a *real barrier* between gambling and myself, and now I suddenly perceive that my former understanding of the situation is no more than a memory of an idea, a memory of a feeling... After having patiently built up barriers and walls, after enclosing myself in the magic circle of a resolution, I perceive with anguish that *nothing* prevents me from gambling.[4]

The gambler is not weak willed. The problem that he is facing affects all of his past resolutions about present action. A past resolution creates a 'magic circle' that is supposed to keep him from acting against his original decision. But the force of magic is no force at all. A mere memory of a past decision cannot bind one's present conduct. As the gambler's predicament shows, the temporal transmission of the force of a past

---

[2]   Sobel (1994), 247 (my emphasis): "A person can have a capacity to set her mind adamantly, and to make decisions for the future that are not only firm but irrevocable. *By dint of sheer will*, a person may be able to "tie her hands" and make necessary and inevitable some action, while making impossible every action that would otherwise have been a possible alternative to it… In order to be sure that she does not answer the door, a person could, instead of handcuffing herself to the stove, do that sort of thing "all in her head".

[3]   Fodor (1987), 3-4: "If someone intends to be at your airport on the 3 p.m. flight, then, ceteris paribus, he will produce behavior of a sort which will eventuate in his arriving at that place at that time, barring mechanical failures and acts of God. And this works not just with people whose psychology you know intimately: your closes friends, say, or the spouse of your bosom. It works with *absolute strangers*; people you wouldn't know if you bumped into the. And it works not just in laboratory conditions—where we can control the interacting variables—but also, indeed preeminently, in filed conditions where all you know about the sources of variance is what commonsense psychology tells you about them. Remarkable. If we could do that well with predicting the weather, no one would ever get his feet wet: and yet the etiology of the weather must surely be child's play compared with the causes of behavior".

[4]   Sartre (1956), 70.

resolution is not a matter of course. One cannot simply claim that a past resolution creates by itself a real barrier that makes the agent act at the later time in conformity with the resolution.

Even if the force of a resolution could be transmitted over time, there is a further problem: A decision appears to be always revocable by mere dint of sheer will.[5] What it takes to untie one's hands 'mentally' is no different from what it originally took to tie them. This raises the worry that a decision might have no binding force at all. The authority of a self-imposed constraint might vanish as soon as the agent realizes that she is both the source and the target of this authority.[6] We face what Richard Moran calls the 'quandary of binding and unbinding oneself': "Insofar as I recognize myself as someone who can bind himself, I must recognize that I can also always unbind myself."[7]

2. The Three Ways of Diachronic Agency

2.1. Many attempts have been made to come to terms with the peculiar and puzzling features of future-directed decisions and intentions so as to explain their alleged binding force on future conduct. Unfortunately, none of the many theories currently available provides a satisfactory account. All of these theories fail to do justice to the distinctive character of the diachronic agency produced by our reliance on future-directed decisions and intentions. The standard views mistakenly equate this form of diachronic agency to the one produced by forms of distal self-control *other* than genuine future-directed decisions. Or so I will argue in the first part of this paper. An alternative account is thus called for. I will try to sketch one in the second part of this paper.

2.2. Let's consider *A*, a philosopher who is supposed to lecture tomorrow on intentional action. Today, *A* realizes that intentionally spilling some ink during the lecture tomorrow would be a very effective way, indeed the best possible one, to illustrate the central claim of her lecture. She decides, therefore, that she will spill ink at the lecture tomorrow.

---

[5]   See Bratman (1999), 2: "Future-directed intentions and plans are, after all, revocable: They do not control one's future conduct by way of some mysterious action at a distance: and many times, in the face of new and relevant information, we recognize that it would be folly to stick rigidly with our prior intention." See also McClennen (1997), 232: "Being resolute involves merely 'mentally' tying one's own hands by an 'act of will' and it seems to leave open the possibility that one can, in the future, perform another act of will that undoes the former resolve."

[6]   See Kant (1797), 214: "If the I *that imposes obligation* is taken in the same sense as the I *that is put under obligation*, a duty to oneself is a self-contradictory concept. ... One can also bring this contradiction to light by pointing out that the one imposing obligation (*actor obligationis*) could always release the one put under obligation (*subjectum obligationis*) from the obligation (*terminus obligationis*), so that (if both are one and the same subject) he would not be bound at all to a duty he lays upon himself. This involves a contradiction." See also Cohen (1996).

[7]   Moran (2001).

There are several ways in which *A*'s present decision can influence her future conduct. According to the first one, *A*'s decision sets in motion a *mere* causal mechanism (either internal or external to her body) that is going to cause the spilling of the ink tomorrow. This future ink-spilling is not going to be the direct product of *A*'s future exercise of her agential governance, since tomorrow *A* will be under a physical constraint to spill the ink. She will not be acting of her own contemporaneous accord at that time. Once the causal mechanism is set in motion, at most she might have only an indirect control over it. Were she to judge, some time between now and tomorrow, that she has no longer reason to spill the ink, this judgment would not by itself stop the mechanism. She could stop it only by *tampering* with it.

2.3. The second way for *A*'s decision to make her spill ink tomorrow is by way of a so-called *precommitment*.[8] A precommitment is some causal device, psychological mechanism, institutional or social practice that induces the agent at the time of action *f* to choose to do as she had originally decided. There are two ways of accomplishing this. First, the precommitment might change the payoffs of the feasible courses of action at *f*, so that the agent will have either an extra positive incentive to φ or an extra disincentive to refrain from φ-ing. For instance, an agent might precommit herself to stop drinking by taking a drug like disulfiram, which makes her especially sensitive and unpleasantly so to even a small amount of alcohol. Another example is an agent who precommits to keeping some money invested for future retirement by depositing it in a retirement account that imposes high penalties on early withdrawals. Second, precommitments might affect the feasibility of courses of action alternative to φ-ing. For instance, when one attempts to resist the temptation to drink by locking the liquor cabinet and throwing away the key, or puts money into a retirement account that does not allow for early withdrawals.

Although sometimes the word 'precommitment' is used to cover cases of mere physical constraints, like Ulysses's being tied to the mast of his ship, I think it is better to reserve the term for those strategies that, even if they involve the work of a mere causal mechanism, still let the agent exercise *rational governance* at the time of action; so that she does choose to φ at *f*, but she does so over a set of alternatives that have been modified by the precommitment, which has either restricted the number of alternatives open to the agent or changed their pay-offs. In my sense of the term, reliance on precommitments always respects the agent's *diachronic* rational governance, since the agent who does φ as a result of a precommitment exerts rational governance both at the time of decision *and* at the later time of action. Imagine somewhat fancifully that *A*—who cares most about her life—hires someone to kill her in case she fails to spill the ink at the lecture tomorrow. In this situation, the self-generated threat of death will work as a precommitment and make her spill the ink tomorrow. This spilling is something that

---

8   The term 'precommitment' was first introduced in Elster (1979).

A does at *f* of her own contemporaneous accord, unlike the case in which, say, she takes a pill that reliably induces a delayed onset of convulsive movements at *f* that *A* expects to make her hit the inkpot and spill the ink.

A's present decision to rely on the effects of a precommiting device exerts only an *indirect* influence on *A*'s future action. *A* must take some steps at the present time to bring about those conditions that will induce her to choose to spill ink tomorrow. Secondly, and most importantly, the influence of a precommitment is of a *manipulative* sort. The considerations that will induce *A* to choose to spill ink are at least in part extraneous to the original merits of the case, i.e., to the reasons that supported her original decision to spill ink. Novel incentives have been introduced or alternative courses of action have been made unfeasible. Hence, when *A* chooses because of her precommitment to spill ink tomorrow, she might end up spilling the ink *reluctantly*. This is because she would not spill ink now if she were to choose now what to do independently of the effects of her prior precommitment. In this sense, she can see herself as being coerced into spilling the ink, even if the coercion is one that ultimately originates in a decision of her past self (that is, in the past decision to adopt the precommiting device).

The manipulative effects of precommitments are not optional, even if in some cases they might not be perceived as such (since they simply overdetermine the agent's choice or the agent is willing to let herself be determined by them). The point of the precommitment is to be able to induce the agent to φ at *f regardless* of whether the agent at that time would still choose to φ in the absence of the precommitment. This is the reason why precommitments are especially suited for distal self-control when the agent at the time of the original decision expects to resist following through on her decision in the future.

2.4. The third and final way for a decision to make a difference to future action is the case in which *A* spills the ink tomorrow just *out of* her past decision to do so. This is the case in which I am interested here: The case of genuine future-directed decisions. Let's first characterize the way in which they operate by contrast with the ways of distal self-control presented above. First, in acting out of a past decision, one is not under a physical constraint; one is not—to use Velleman's expression—under the 'remote volitional control' of the past decision.[9] Second, genuine future-directed decisions, unlike precommitments, do not manipulate the agent. In acting out of a decision to φ, the agent does φ because at *f* she finds that φ-ing is choiceworthy on its own merits at that very time. Future-directed decisions are not tricks for self-directed manipulation.

A positive characterization of the third way of spilling ink tomorrow, the way of genuine future-directed decisions and intentions, can be offered in terms of the agent's *diachronic autonomy*. In acting out of a decision, the agent not only exerts her rational

---

9   The expression 'remote volitional control' is from Velleman (2000).

governance both at the time of decision and at the time of action (something already guaranteed by precommitments), but she exerts it in response to considerations that have not been manipulatively altered. Acting out of a future-directed decision is thus the paradigmatic instance of a particular kind of temporally extended intentional agency, the agency that respects the agent's diachronic autonomy. Let's call it *autonomous intentional diachronic agency*, or *AIDA* for short.

This is something that we must keep in mind in our assessment of the standard views of future-directed intentions: No account can be considered an adequate theory of genuine future-directed decisions (rather than of some alternative form of distal self-control) if it fails to acknowledge and make room for the distinctive contribution of decisions to *AIDA*. On the basis of the considerations that I have presented in the last few sections, let's now turn out attention to the standard theories of future-directed decisions (unless otherwise noted, from now on I will use 'decision' to stand for 'future-directed decision').

3. Decision-Based Reasons

3.1. For reasons that should by now be clear, the most inadequate attempts at accounting for future-directed decisions would equate decisions to devices of either mere physical constraint or precommitment. In either case, there would be the violation of the basic requirement of diachronic autonomy.[10] Theories of this sort would be patently inadequate in their failure to recognize the distinctive features of genuine future-directed decisions.

It takes a lot more work, however, to assess the adequacy of those theories that account for decisions in terms of so-called 'decision-based reasons'.[11] The basic idea is that, by deciding to φ at *f,* the agent gives her future self at *f* a (usually) conclusive reason to φ at that time. The attractive feature of these views is that they respect the agent's exercise of rational governance at the time of action. As shown above, however, this is not sufficient for an adequate theory of decisions. After all, precommitments are also sources of reasons that support φ-ing at *f* in the respect of rational governance. What must be shown in addition is that a decision-based reason is not manipulative, i.e.,

---

[10]   Versions of the view of future-directed decisions as precommitments can be found in Schelling (1980); Ainslie (1992); Ainslie (2001); Holton (1999); and Rachlin (2000).

[11]   See Anderson (1996), 542: "We can frame coherent intentions only under a conception of ourselves as autonomous agents. We act autonomously only if we act on reasons that we regard as warranting our actions. But agency is exercised over time. Autonomy thus requires that decisions we make now be able to control our future actions. But we cannot act autonomously at that future time unless the sort of control our prior selves exercised is compatible with our acting for good reasons at that time. This is possible only if we regard the commitments we made then as giving us reasons to fulfill them now." See also Velleman (1997), 46. Reprinted in Velleman (2000).

that it respects the stronger condition of diachronic autonomy. This can be established only by looking at the details of how these various theories explain the source of the decision-based reasons. One cannot simply rest content with the idea that a decision generates a reason literally out of nothing else but itself. To claim that 'I have a reason to do such-and-such now simply because I have earlier decided to do such-and-such' is just to restate the problem of decisions, not to solve it—unless, that is, one is convinced that decisions are really works of magic.

3.2. Several proposals have been advanced to explain how a decision-based reason can be generated. First, there are the views that take decisions to be the *intra*personal version of some familiar and fairly well understood *inter*personal phenomenon. This is the case of the view of decisions as self-directed commands or injunctions.[12] Then, there are the views of decisions as self-directed pieces of advice (i.e., as instances of sincere and cooperative Gricean communication between a present self and a future one).[13] According to these theories, the earlier self gives to the later self a usually conclusive reason to φ, in analogy with the reason to φ that one would give to another person by either commanding or advising her to φ.

According to another family of accounts, the agent has a decision-based reason to φ because acting on an earlier decision satisfies some standing preference of the agent. According to one such view, the preference is for one's own steadfastness; we care about our reputation as agents who keep their word and thus we have a reason to act as we said, or suggested, we would.[14] According to another proposal, following through on past decisions satisfies our standing desire for self-knowledge. According to such view, a future-directed decision is first of all a prediction about our future. Hence, by acting in conformity with it, we secure its truth and thereby satisfy our preference for

---

[12] See Kenny (1963), 216-220; Castañeda, (1975), 42-43, 155; see also Velleman, (1989), 99: "My theory of intention can be read as a variation on a familiar theory that identifies intentions with commands issued by the agent to himself... My theory of intentions is quite similar. For although I do not identify intentions with commands, I identify them with expectations whose utterance would convey commands." For a criticism of the self-directed command view, see Rundle (1997), 189. Notice that according to the view presently under consideration, the authority of the self-command depends on the generation of a command-based *reason*. There could be other versions of the self-command theory that do not qualify as versions of the decision-based reason view. This is true if one thinks that commands work by the internalization of a psychological constraint, as a sort of brainwashing.

[13] See Rovane (1998), 149: "As someone engaged in long-term planning, I form intentions about what actions I will take in the future. So by forming these intentions, I intend to elicit a response in my future self… Insofar as I form long-term intentions bearing this in mind, I intend that my later recognition of my present intentions will be my reason to do then what I now intend." See also Hinchman (2003), 1.

[14] See Sobel (1994), 249: "That [an agent] had committed herself to doing something, or resolved to do it, would be for her a new reason for doing it, much in the way in which promises to others provide most people with new reasons for doing what they have promised to do. This is at least part of the way in which advance decisions work for most people. Most people value being steadfast, and perhaps even more would so if they though about it and appreciated its place in the economy of an integrated psyche."

self-knowledge. For a view of this sort, a decision turns out to be a sort of self-fulfilling prediction.[15]

A further suggestion on the source of decision-based reasons is the idea that a presently available record of a past decision gives to the agent an *evidential* reason to φ without rehearsing the full deliberation at *f*. This is because the record of her past decision is evidence that, if she were to deliberate again, she would reach the same conclusion she reached when she originally decided.[16] A different proposal is that decisions work by a 'snowball effect'.[17] As a result of having made a decision to φ at a future time, the agent makes investments and coordinates with her future φ-ing. When the time of action comes, the costs of undoing the coordination and the investments give the agent novel and usually decisive reasons to stick with her past decision.[18,19]

3.3. It is undeniable that each of these theories describes some actual and effective form of distal self-control. They are, however, problematic as accounts of genuine decisions and the role of these decisions in *AIDA*. Here I am not going to raise criticisms directed

---

[15]   See Velleman (1989), 98: "Intentions to act... are the expectations of acting that issues from reflective theoretical reasoning. These are self-fulfilling expectations of acting that are adopted by the agent from among potentially self-fulfilling alternatives because he prefers that they be fulfilled, and they represent themselves as such... These expectations... are, and represent themselves as being, self-fulfilling expectations of acting that the agent adopts out of a desire for their fulfillment... Intentions differ from ordinary predictions in their psychological relation to the evidence, but they are future-tense beliefs that rest on evidence, and so they will deserve to be called predictions."

[16]   See Joyce (1999), fn 16: "A person who plans to do A may thereby give her future self as a *reason* to do A; that is, in the future she might take the fact that A accords with her past plan as a consideration in favor of doing it and this makes it more likely that one does it. One of the easiest, and I suspect most common, ways for this to happen has to do with the fact that adopting a plan to do A can give one's future self *evidence* for thinking that one would decide to do A were one to think about the matter in detail. This, of course, gives one's future self a reason *not* to think about the matter in detail."

[17]   The term 'snowball effect' is introduced by Bratman (1987), 82, who credits it to John Etchemendy.

[18]   See Pink, (1996), 130, see also 75: "By deciding on my actions in advance, I effectively increase the likely stake on a bet that I shall perform the act decided on. … Decisions to do A later... give rise to persisting intentions. And so they make it more likely that before doing A, I shall have performed other actions on the assumption that I may or shall do A. Such actions tend to increase the benefits I derive from doing A. I decide to go on holiday well in advance of going. My holiday is thereby better prepared. So the holiday provides me with more enjoyment. On the other hand, taking a decision to do A later, because of this coordinatory influence, also increase the likelihood of my performing actions I shall regret performing unless, in the end, I do A."

[19]   In this paper, I do not discuss the decision-based reasons that are allegedly generated by a future-directed *selection* among two or more equally choiceworthy options. I discuss them more extensively in my "Decisions, Diachronic-Autonomy, and the Division of Deliberative Labor" (manuscript). Briefly put, the problem with this solution is twofold. First, it is not clear how a future-directed selection is supposed to exert its influence over time, which is the specific problem with which I am concerned here. Even if it is true that the original decision-as-selection introduces a novel and conclusive reason in support of the selected action, the question still remains of how such reason can exert its influence over time. What makes the tie-breaking decision effective over the temporal distance cannot just be the fact that the agent picked one particular option over another in the past. To this extent, the theory must be supplemented with some of the other views discussed in the paper. Second, there is a further consideration against the idea of a selection-based reason as the basis for a general theory of decisions: Not all cases of future-directed decisions involve some selection or tie-breaking over equally choiceworthy actions.

at the specific suggestions about the origin of the decision-based reason. I rather want to discuss a general problem with the very idea of decision-based reason: Such reason appears to be one reason too many. On the one hand, if the case for φ-ing stays the same over time and the agent is able to appreciate this case directly at the time of action, then there is no need for the agent to rely on the decision-based reason. On the other hand, if at the time of action the agent would no longer find φ-ing choiceworthy on its own merits, the decision-based reason appears to have what Bratman calls a 'bootstrapping effect'.[20] This effect is troublesome. First, the decision appears to be imposed on the agent at *f*. Second, this imposition is particularly problematic if the agent's change of mind about the choiceworthiness of φ-ing is indeed supported by good reasons (such as the detection of an error in the past deliberation or the acquisition of novel relevant information).

3.4. Notice that the bootstrapping by itself is not necessarily undesirable. One might want to be bootstrapped into φ-ing whenever one is expected at *f* to resist following through on one's earlier decision. This is, however, exactly how precommitments are supposed to work. It is not the way decisions work, though. Therefore, although bootstrapping is not necessarily a problem for the agent, it is certainly a problem for a theory of decisions. Another way to show that decisions are not supposed to be bootstrapping is to consider that decisions (as opposed to promises or commitments that intentionally produce assurance in other subjects) are very easy to revoke.[21] Therefore, at the time of action, it takes nothing for the agent to repudiate her earlier decision, and a decision-based reason, as opposed to a precommitment, should not provide any resistance to such repudiation. But if the decision-based reason can be so easily set aside, one wonders how it can really exert any normative force on the agent at the time of action. On the other hand, if the decision-based reason is going to have some binding force on the agent, it seems that it can do so only manipulatively, that is, by introducing a consideration that was not part of the original case for φ and that cannot be removed simply at the agent's whim. This would make the reason bootstrapping in Bratman's sense, since the decision-based reason to φ could continue to hold even if the original case for φ-ing no longer stands. This, however, violates the requirement of diachronic autonomy. Notice that such violation is present even in cases in which the original case for φ-ing at *f* does not change over time. The problem is that the agent could be said to be manipulated into φ-ing even if the original case for φ-ing is still valid. At issue is not *what* the agent ends up doing, but *how* she is led to the doing of it. If the decision-based reason determines the agent to φ but is extraneous to the original merits of the case, then such reason is manipulative.

---

[20]   See Bratman (1987).
[21]   See Davidson (1980), 90: "If an agent does not do what he intended to do, he does not normally owe himself an explanation or an apology".

3.5. Theories of decisions in terms of decision-based reasons appear, therefore, to face a dilemma between two unappealing alternatives. On the one hand, they might bite the bullet and embrace a truly revisionist account according to which decisions do not really respect diachronic autonomy. On the other hand, they might salvage diachronic autonomy by making decision-based reasons to be at best redundant and, in any event, so easily revocable to put into doubt that they really have any binding power on the agent. This, however, goes against the intuitive idea that motivated the investigation into the nature of decisions, namely, that future-directed decisions are somehow effective at influencing future conduct.

This dilemma does not yet show that any theory in terms of decision-based reasons is doomed. It poses however a very stark challenge. The theory must show that there is a way to generate decision-based reasons that are effective at controlling future conduct, but can be immediately revoked or suspended at the agent's own discretion, and are neither bootstrapping nor manipulative.

Not all of the proposals discussed above seem equipped to live up to this challenge. This is especially so for those accounts formulated in terms of the satisfaction of a standing desire either for steadfastness or for self-knowledge via the fulfillment of self-directed predictions. This is because such standing desire is always in place and can only be fulfilled by acting as originally decided, regardless of whether the decision is still correct. As far as steadfastness and self-knowledge are concerned, once a decision is made, any deviation from the intended course of action is necessarily a failure of either steadfastness or the self-knowledge promised by one's prior predictions. Problems arise also for the snowball-effect theory. The snowball effect is a clear case of bootstrapping. Because of this effect, the agent at f is stuck with her past decision. It is either too late or too expensive for her to backtrack, even if it is true that she would no longer decide to φ if she were to choose what to do without the threat of being crushed by the snowball generated by her past decision.[22]

The theory of decisions as self-directed commands suffers from the bootstrapping problem. Let's see how. The theory explains the force of a decision on the model of interpersonal commands. In a case of interpersonal commands, the agent is to φ simply because she has been commanded to do so. To this extent, a standard command is always bootstrapping. This is, after all, the main point of the command: To give the agent a reason to do as commanded for that reason alone. As such, bootstrapping

---

[22]   I am not suggesting that decisions either cannot or do not generate decision-based reasons. My point is only that these reasons, even when present, are not what—under normal conditions—explains the agent's autonomous diachronic agency. Nonetheless, whenever there is the possibility that such decision-based reasons be generated, the agent should be aware of their potential bootstrapping effects. The agent, in other words, must pay attention to the *indirect* consequences of her decisions, such as the snowball effect described above (which the agent would normally want to avoid, but in some cases might encourage as a sort of insurance policy against a future irrational reversal of preferences—in which case, however, the agent would no longer be acting out of her decision, but rather on the snow-ball effect of her earlier decision).

is not a problem for interpersonal command. But it becomes troublesome when we take decisions to work as the intrapersonal analog of command, since decisions are not supposed to be bootstrapping. A possible solution is to claim that self-directed commands can always be revoked and as such they do not really bootstrap the agent. This solution, however, points to another serious problem with the idea of a decision as a self-directed command. If the agent really has the power to revoke a self-directed command, it is no longer clear in which sense she is ever bound by the command that she herself has issued. No such problem arises for interpersonal commands, since the sovereign and the subject are separate persons, and only the sovereign has the prerogative to revoke the command. But in the intrapersonal situation the source and the target of the authority are the same and this appears to deprive the self-directed command of any real authority.[23] A decision as a self-directed command could therefore avoid the bootstrapping effect on future conduct, but only at the price of having no effect on such conduct.

3.6. The theories of decisions as generating self-directed advice or evidential reasons appear to fare better. The agent need not get stuck with doing φ just because of a piece of advice or an evidential reason in support of φ-ing. Either kind of consideration can be defeated when there is a suspicion that it favors a course of action not conclusively supported by independent considerations. Moreover, it is plausible to claim that these decision-based reasons are not necessarily redundant. For they might help the agent with determining what she is to do at *f* when she is either unable or unwilling to embark in a full deliberation at that time. Therefore, among the theories considered thus far, those formulated in terms of either self-directed advice or evidential reasons appear the most promising.

I must sound a note of caution. If we observe the standard instances of *AIDA*, it is rather unusual to see an agent who φ's upon consulting her past decision as a piece of evidence about what she would choose to do now if she were to engage in full deliberation. Likewise, it is rare to observe an agent who uses the record of a past decision as a piece of trustworthy advice or an instance of Gricean communication coming from her past self. Of course, theories in terms of evidential reasons, self-directed advice, or self-directed transtemporal communication need not be taken as claiming that we *explicitly* and *reflectively* engage in self-advising or the consultation of evidential reasons. Hence they should not be faulted by the mere observation that we do not seem openly to engage in any of these practices. Nonetheless, it is still implausible to claim that, even if only implicitly and non-reflectively, when we act out of our past decisions, we are always either consulting evidential reasons or heeding self-directed advice offered to us via transtemporal Gricean communication.

---

[23]   See footnote 6 above.

I am not denying that there might be cases that can be properly described in the terms of these theories. These cases, however, do not appear to be the norm, and our present concern is with finding a *general* account of decisions. This criticism would be unjustified if the theories were meant to enter a weaker claim, namely, to suggest that decisions work somehow in *analogy* with interpersonal phenomena like command, advice, or communication. As such, the theories would not be meant to describe the actual mechanisms responsible for generating and grounding decision-based reasons. The analogy would be meant only to illustrate the structure of decision-based reasons; to show that decision-based reasons are (a) in place only once we make a decision, (b) usually conclusive, yet defeasible, (c) not redundant, yet not manipulative. In other words, the theory would simply claim that decision-based reasons are very much like advice-based or Gricean-communication-based reasons, except that they need not be the product of actual episodes of intrapersonal advising or communication. The latter disclaimer, however, deprives such proposals of real explanatory power. What we are left with is only the indication of how decision-based reasons are supposed to look like if they are to play an acceptable role in the account of future-directed decisions and *AIDA*. This would amount to a restatement of the challenge, not to a solution of it. The theory would say nothing about how these reasons are supposed to arise, how often they do so, and in virtue of what; which are the questions that we are trying to answer here. We must thus conclude that even the more promising among the decision-based theories turn out to be disappointing. In principle, this still leaves open the possibility of formulating a theory in terms of decision-based reasons, but none of the standard views has proved satisfactory.

4. The Psychological-Propensity View

4.1. A radical reaction to the worries raised by the idea of decision-based reasons is to do away with the idea altogether. This reaction is most famously associated with the work of Michael Bratman. His positive suggestion is that the distal influence of decisions is due to features of our psychological make-up: To a set of underlying habits, skills, dispositions, and propensities that, under normal conditions, make us not reconsider our past decisions and follow through them without further ado when the time of action is ripe.[24] The agent neither takes herself to have a decision-based reason to φ, nor does she engage in full deliberation at *f* about what she is to do then. The internalized propensities explain the seeming 'inertia' of decisions; the fact that *by default* the agent is going to follow through her past decisions.

4.2. It is undeniable that our psychological make-up plays an important role in the explanation of the influence of decisions on later conduct. But this platitude does not

---

[24] See Bratman (1987), 16-17, 52, 60, 65-66.

yet amount to a theory of decisions. For that we need a more specific characterization of the alleged contribution of these propensities. And we must make sure that the more specific account does not run into the troubles encountered by the views considered thus far. In particular, we must make sure that it does not conflate genuine decisions with either physical constraints or precommitments.

4.3. Here is one suggestion on how to develop the psychological-propensity view (or propensity view, for short) of decisions. Some have suggested that decisions work by instilling mechanisms of psychological repression and inhibition.[25] The suggestion is that at f the agent does φ simply as a result of a psychological inability to entertain the possibility of acting otherwise. This inability was originally instilled by the agent's past decision via some psychological mechanism of repression or inhibition that prevents her at f from entertaining considerations that might lead her to choose not to φ.

A repression mechanism of this kind might be actually operating in certain situations (and we might even have reasons to secure its existence), but it cannot be part of an acceptable account of genuine decisions and *AIDA*. The repression operates in violation of diachronic autonomy. It has a manipulative effect, although different from the one of bootstrapping. The repression alters the situation of choice not by adding extraneous considerations, but by affecting the agent's ability at f to appreciate considerations that have the potential to induce her not to φ. The suggestion is that a decision works not by giving the agent a leg up or an arm twist (as a decision-based reason theory might have it), but by brainwashing her. Like any effective manipulative devices, the repression mechanism is not immediately responsive to the agent's judgments about the choiceworthiness of action. The inhibition can thus be overcome (if at all) only by tampering with the underlying propensity.

4.4. A better version of the propensity would have to avoid the claim that a decision makes the agent *unable* to reconsider whether to φ. It should rather maintain that the agent has a propensity not to reconsider, but that the decision is always open to reconsideration, if only the agent sets herself to do so. In this sense, the propensity must be judgment-sensitive, so as to be different from either mere physically constraining devices or the psychological effects of the likes of inhibition, repression, brainwashing, or hypnosis.

---

[25]   See Shapiro (2001): "Rules normally prevent agents from "willing" disobedience to their demands, not by physically disabling the will's expression [i.e. by setting constraints on action]. How might these constraints on the will arise? First, the rule might prevent the agent from considering the reasons for disobeying the rule. The reasons for breaking the rule would, in other words, be 'repressed' by the rule. This repression blocks the possibility of intentional action contrary to the rule—the agent, being unaware of reasons for breaking the rule, would be unable to break the rule for a reason. Second, the rule might disable the person's normal psychological inhibitions. An agent in the grip of a rule might no longer be able to withstand certain emotional pressures, such as guilt and shame, even though the agent would be aware of the important reasons for breaking the rule."

This improved version might still run into problems. The issue concerns how the propensity for nonreconsideration of a past decision to φ at *f* relates to the agent's φ-ing at *f*. One tempting suggestion is that the propensities for nonreconsideration are also responsible for leading the agent to act as originally decided. According to this proposal, because of our psychological make-up, once we have decided to φ at *f*, the decision typically produces a bias in the unfolding of our future agency: unless we *actively* contrast the propensity by reconsidering our past decision, we are naturally led to φ at *f*.

4.5. This solution, however, is problematic. It seems to suggest that following through a past decision is a sort of automatism, a sort of *ballistic* phenomenon. I say a 'sort of' ballistic phenomenon since the phenomenon does not have exactly the same features as the unfolding of an inertial physical process, such as a tennis swing. This is apparent by comparing what happens when one tries to stop the two kinds of processes. When the tennis swing is under way, there is a point past which the agent can no longer directly stop or deviate the swing, if not by some external interference or tampering (for instance, by interposing a physical barrier that prevents the completion of the swing). A propensity to notreconsider and to follow through a past decision, instead, can always be *directly* stopped—up to the last moment before the inception of the action. The agent can do so simply by actively reconsidering the past decision. This means that there is still some role for the agent's rational governance in the unfolding of the propensities that are responsible for the influence of prior decisions on later conduct. This rational governance, however, is exerted only when the agent *stops* the propensity, not in its normal unfolding. According to this proposal, there is no active sustaining of one's prior decision. The process is, so to say, a semi-passive one: It is passive by default, but it leaves room for immediate active interruption. In other words, it has an agential profile like that of yawning. Typically, a yawning unfolds passively as a result of some underlying physiological process, but the agent is usually able to take control over the yawning in that she can actively stop it if she wants to.

As Bratman himself has correctly remarked, the agent "normally retains rational control over what she does when the time [of action] comes. Following through with one's plan is not, after all, like following through with one's tennis swing."[26] It does not seem that the semi-passivity illustrated above, despite its difference from that of the tennis swing, is sufficient to capture the extent of the agent's rational control over her φ-ing at *f*. The semi-passivity of the yawning-version of the propensity view violates the requirement of diachronic autonomy. In acting out of a past decision, the agent is normally in direct active control of both the initiation and the unfolding of the action. Whatever bias a prior decision exerts on future conduct, it is not by way of causing a default *absence* of agential governance. This is reason enough, in my view, to reject

---

[26]   Bratman (1999), 60.

the improved version of the propensity view. (Furthermore, even if it were true that in some instances of *AIDA* we follow through on a prior decision passively, these cases cannot be the paradigm around which to formulate a general theory of decisions.)

4.6. Notice that in the previous discussion I have not attributed the yawning-view to Bratman. This is so even if (a) Bratman appeals to psychological propensities in order to explain the nonreconsideration of reasons at the time of action, and (b) he is explicit about the semi-passive status of nonreconsideration, as can be seen from the following quotation.

> Nonreflective reconsideration of a prior intention is typically not itself an intentional action of refraining from reconsidering. Rather, nonreflective nonreconsideration of a prior intention typically consists merely in the *absence* of reconsideration of that intention, an absence explained by appeal to relevant general habits and dispositions.[27]

Despite these explicit statements on the semi-passivity of nonreconsideration, it is not clear whether he takes the propensities of nonreconsideration to be also propensities to follow through on a past decision in analogy with yawning. The latter position is not forced on him simply by his account of nonreconsideration. It is possible, at least in principle, to keep separate the issue of the nonreconsideration of a past decision from the issue of the retention of a decision over time (which is the primary concern of the present investigation). First, a decision can be retained even as it is undergoing reconsideration. Second, a decision can be retained over time, even it is undergoes a repeated reconsideration. If the reasons supporting the decision are stable over time, the retention of the decision is not affected by the reconsideration. Moreover, it seems that the basic form of retention of a decision has psychological priority over the nonreconsideration. In a sense, a past decision must be retained up to f included in order for it to be a possible object of nonreconsideration at f. Given the potential separation between retention and reconsideration, one might grant the semi-passivity of the latter, while leaving open the possibility that acting on a retained decision be an entirely active process in the respect of diachronic autonomy.

This conclusion leaves open the possibility of a version of the propensity view that appeals to two different sets of propensities, one underlying nonreconsideration, the other underlying retention. This version would improve over the analogy with yawning. But at this stage of the analysis, it could hardly offer the final word on the nature of decisions and *AIDA*. This is because the theory would be centered the an uncontested, but uninteresting claim that there are some yet unspecified propensities underlying the retention of decisions and whose operations are compatible with diachronic autonomy. No one is likely to contest this, but absent a more detailed specification of the nature of

---

[27]   Bratman (1987), 61, his emphasis.

these propensities, the best that can be said about the propensity view is that it is a mere placeholder for detailed explanations still to come.

5. An Inclusive Account of Future-Directed Decisions

5.1. In the previous sections, I have shown the limitations of the standard views of decisions. The most common problem is that many views fail to recognize the distinctive role of decisions in their contribution to diachronic agency. These views account for decisions in ways that turn decisions into instances of some different strategy for distal self-control, such as manipulative precommitments. I think that this conflation is usually unintended, since these theories are not meant to deny the distinctive nature of genuine future-directed decisions. They simply fail to appreciate that they do not really account for it.

This is not true of all theories. Some of the views discussed above (such as the evidential reason view, the Gricean-communication, and certain versions of the propensity view) account for decisions in ways that make decisions truly different from other forms of distal self-control. These views, however, suffer from another problem. Each of these accounts, individually taken, does not have sufficient scope to offer a general theory of decisions. If one were to attempt to generalize any of these accounts to cover for all cases of future-directed decisions, one would impose an unrealistically univocal explanation that only fits some instances of decisions. As indicated above, there are cases in which one can say that one consults a memory of a decision as a source of evidential reasons about what she would conclude to do if she were to engage in full deliberation now; but this is not always so. Likewise, there are cases in which one might truly engage in Gricean-communication with one's future self; but this is not always so. Finally, there are cases in which psychological propensities explain the inertia of decisions; but once again not all decisions appear to work this way.

One might attempt to generalize the views by taking them to offer not a literal account of the actual mechanisms responsible for *AIDA*, but a more abstract characterization of the distinctive features of decisions by analogy with phenomena such as Gricean communication or the inertia of psychological propensities. This is a troublesome move, however, since it deprives the views of their alleged explanatory power. These views would simply restate the problem, by telling us the ways in which decisions are different from other forms of distal self-control, rather than telling us *how* they can so operate in diachronic agency.

5.2. A different suggestion, which would avoid the risk of explanatory emptiness, is that of combining the detailed accounts of the more promising views into a more comprehensive theory. This is possible, since there is no incompatibility between the mechanisms of evidential reasons, Gricean-communication, and psychological

propensity. Different mechanisms might be at work in different contexts. Each mechanisms would be a token realization of *AIDA*, the type of diachronic agency distinctive of genuine future-directed decisions. This combined view would gain in scope. However, like all disjunctive theories obtained by the mere combination of satisfactory but local accounts, the combined view is explanatorily weak. We lack a principled way of determining which mechanism should count as a realization of *AIDA* and why.

In the remainder of this paper, I will suggest a way to put some organization into the heterogeneous mechanisms that appear to underlie *AIDA*, so as to avoid a mere mixed-bag-of-tricks account of decisions. I will proceed in a somewhat unorthodox way. All the standard accounts of decisions share the assumption that future-directed decisions have to be *effective* in influencing the agent's future conduct. Contrary to this assumption, I will suggest that the primary focus of the investigation should be on the fundamental structure of *AIDA*, leaving open whether, and if so to which extent, future-directed decisions contribute to *AIDA* as effective tools for the distal control of conduct. According to my strategy, the first question to ask is whether we can engage in the kind of diachronic agency that we normally associate with future-directed decisions even if we lack future-directed decisions with effective powers to influence future conduct. I will argue that AIDA need not depend on effective decisions. This claim will help us get a better understanding of the true nature of the contribution of future-directed decisions to diachronic agency, which in turn will help us find a principled way to make sense of the disparity of mechanisms that have been suggested by the standard views of decisions.

5.3. Let's return to the case of spilling ink discussed at the outset. Let's imagine that I am the one who is supposed to deliver the lecture tomorrow and that the following conditions hold (I will refer to this scenario as *S*). Imagine that all the relevant considerations concerning the advisability of my future ink-spilling are easily accessible to me not only now, but also at any other time from now until tomorrow. This means that whenever I direct my attention to the question of what I should do tomorrow at the lecture, it is immediately clear to me—on the basis of the directly available considerations—that I ought to spill the ink.

In a scenario like this, it does not really matter whether my original decision to φ tomorrow makes a difference to my future conduct. Given that the relevant considerations are always immediately and easily accessible to me, whenever I need to determine what I am to do at *f,* I simply consult these considerations and immediately realize that I am to φ. This realization is not influenced by any previous assessment about the choiceworthiness of φ-ing at *f*, nor is it meant to exert any influence on either future assessments or future conduct. The only case in which the assessment makes a difference is at *f*, since at that point the assessment concerns what I am to do *now*. In this case, the determination of the choiceworthiness of φ-ing is nothing other than the

present-directed and immediately executive decision to φ now. The original decision, instead, is not executive. It is merely an assessment that I ought to φ at *f*. In this respect, the original decision is no different from any of the other assessments about φ-ing at *f* that I might make between the time of the original decision and the time of action. Given that, as Hume correctly remarked, "the will has an influence only on present actions" (Treatise, III.ii.5) none of these assessments has a direct influence over the future action. Each assessment can have direct executive power only in determining what I am to do *at the time of that assessment* (for instance, do whatever is immediately necessary in order to coordinate with my future ink-spilling).

5.4. Notice that in this scenario *S*, diachronic autonomy is respected. First, I exert rational governance about my assessments and my conduct both at the time of the original decision and at the later time of action. Second, rational governance is exerted at both times in response to the original, non-manipulated merits of the case. This means that diachronic autonomy, which lies at the heart of *AIDA*, does not require that a future-directed decision exert any kind of executive power at the future time of action, not even of a mediated and indirect sort. It should not come as a surprise that diachronic autonomy is secured in *S*, since the various assessments are made in isolation from each other and are not supposed to adulterate the original merits of the case.

5.5. In the following sections I want to show how *S* offers an interesting model for studying the nature of genuine future-directed decisions under the less than ideal conditions of the real world. Before doing so, however, I must consider a serious challenge to this strategy. It could be argued that *S* cannot be used to model the working of real-world intentions because it is missing an indispensable element of the idea of acting out of future-directed decisions; namely, the capacity of future-directed decisions to *settle* what the agent is going to do in the future.

5.6. The intuition is that deciding to φ at *f* 'closes the question' of what to do at *f*. Central to the idea of this closing is that in the future the agent is not going to raise (either absolutely or by default) the issue of what to do at *f*. It is undeniable that this is missing in *S*. In *S*, the question of what to do at *f* is always raised in full at the time of action as if for the first time. In this sense, the original decisions-as-assessment does not settle what the agent is going to do; it is not at the origin of a mechanism that makes future reconsideration either unlikely or impossible.

5.7. There is, however, another and ultimately more fundamental sense in which the question whether the agent is going to φ at *f* is said to be 'closed'. The question is closed when one can rest reasonably assured that the agent is going to φ at *f*, and thus coordinate with the expected future φ-ing. Normally, we think of this assurance as the product of a tendency of nonreconsideration, which in turn is produced by an effective decision. But in *S* we can get the assurance without either of the two other elements. In *S*, the

future-directed decision does not settle what the agent will do at *f*. It is nonetheless true that the *agent is settled* on her doing φ at *f*, which is ultimately what really matters.

In *S*, there is a guarantee that the assessment in support of φ is renewed every time that the question of what to do at *f* arises. Therefore, the agent can be said to be settled on φ-ing at *f* even if her original future-directed assessment does not influence, even indirectly, her φ-ing at *f*. The agent can reliably anticipate that she will φ, even if there is no truly *effective* future-directed decision. This means that, at least under the special conditions of scenarios like *S*, there is no need that the φ-ing at *f* be the product of some process of temporal *transmission* or *preservation* of the initial assessment. No such transmission or preservation is necessary in order to guarantee the agreement between the assessments made at the different times. The assessments can be mutually isolated and yet the agent can be settled on φ-ing. The agreement between the assessments is not the outcome of a mutual influence or determination, but of their individual tapping on a shared source, viz., the reasons that constitute the complete and conclusive case for φ-ing at *f*. The assessment agree with each other because they originate in the repeated appreciation of the same set of reasons.

5.8. The conclusion of the previous discussion is that an agent can take herself to be settled on φ-ing at *f* as long as she is entitled to believe in the temporal *stability* of both (a) the complete case in support of φ-ing at *f*, and (b) her capacity to appreciate this case. The agent's sense of this stability in grounded in the combination of two sorts of evidence. On the one hand, there is the theoretical and inductive evidence that warrants the agent's belief in the continued existence of the preconditions for the pursuit and success of φ (say, in the example of spilling ink, that inkpots do not pop in and out of existence unpredictably, that the agent is not going to die before the lecture, etc.). On the other hand, there is what might be called 'deliberative evidence'. The agent has what George Wilson calls "a continued sense of the intrinsic force and authority of the reasons for action which have been decisive for him in his practical assessment".[28] This evidence is of a deliberative kind because it depends on the agent's taking a deliberative, rather than a theoretical/contemplative stance, toward the justification for spilling ink tomorrow.

Thus, as long as the agent's assessments are accompanied by her sense of the stability both of the reasons for the future action and of her appreciation of these reasons, she can take her present assessment to warrant a prediction about her future conduct, namely, that she will do as she has originally decided to do—even if this decision has non executive power. This is how, in scenarios like *S*, a future-directed decision (together with the future-directed intention thereby acquired) can be taken as a source of reliable predictions about the agent's future conduct (thereby making it possible for the agent to coordinate in advance with the 'fixed point' of her future φ-ing).

---

[28]   Wilson (2000), 18.

5.9. One possible problem with the present account is that the reliance on the repeated renewal of assessment exposes the agent to the danger of fickleness. This is because, with each renewal of the assessment about what to do at *f*, the agent has the opportunity to reconsider and possibly change her mind about whether to φ. Nonetheless, as long as both the case for φ and the agent's ability to appreciate this case are stable, there is no reason to worry that the agent will revise her assessment and come to do something other than φ. Nevertheless, diachronic autonomy requires that the agent be given the opportunity for a revision, were he to wish to make one, thereby exposing the agent—at least in principle—to the danger of an arbitrary change of mind. This danger, however, is the necessary price we have to pay in order to enjoy diachronic autonomy and engage in *AIDA*. (This does not yet tell us whether this is a price worth paying, which is an important question, but not one that can be addressed in this paper.)[29] In any event, it should be noted that the danger of fickleness is not specific to my account of future-directed decisions. It is present for any account that relies on diachronic rational governance, even if exerted under manipulated conditions. The success of a precommitment to φ, as much as that of genuine future-directed decision to φ, is always hostage to the ultimate acceptance of the agent at the time of action. This agent is in principle always free to refuse to φ, even if this were irrational of her. The only way to avoid this dependency on the future acceptance and its attendant risk of arbitrary rejection is to exert control via mere physical constraints, thereby restricting the risk only to the possible unreliable functioning of the causal mechanism.

5.10. There is another important concern about the solution that I am exploring. It is not hard to see that, in the special scenario *S*, the distinctive features of *AIDA* are secured even if future-directed decisions are only isolated, non-executive assessments made under the assumption of stability. But what does this tell us about future-directed decisions and *AIDA* in general? What I have discussed thus far is *S*, a rather special and rare scenario. One might be concerned that this discussion has no import on the understanding of normal cases of *AIDA*. This concern is misplaced. The picture of the nature of future-directed decisions and *AIDA* derived from *S* offers the basis for a theory of decisions that can be applied to all situations. This is because the distinctive feature of S—viz., the repetition of mutually isolated, non-executive assessments against the background of the stable sense of the reasons for action—offers a *regulative ideal* of the sort of diachronic agency that we take to be distinctive of genuine future-directed decisions. This ideal offers the organizing principle for the general theory of decisions. First, it allows us to draw a distinction between the basic forms of diachronic agency. The exclusive association between *AIDA* and genuine decisions is due to the fact that *AIDA* is the only kind of diachronic agency that is informed by the ideal illustrated by *S*. No such ideal informs, instead, the distal self-control of physically constraining mechanisms or precommitments. Second, the ideal offers the principle around which

---

[29] For a more extended discussion of this issue, see my "Personal Identity and Diachronic Agency" (ms).

to organize the otherwise disparate collection of mechanisms that forms what I have earlier called the 'combined view', i.e., the mere disjunction of the explanations offered by the most promising of the standard theories. It is now possible to see the various mechanisms as different but possibly equally successful strategies that limited beings like us can adopt in order to instantiate the ideal of *AIDA*.

5.11. Before articulating my proposal, I have to consider a radical objection to one of its central ideas, the reliance on the stability of reasons and rationality. The objection is that usually we cannot assume this stability. Given that human beings are often capricious and inconstant, the work of decisions appears to be that of counteracting or compensating for our changeability, so as to secure consistency in our conduct over time. Therefore, the objection continues, it is erroneous to think that decisions and intentions are at work exclusively, or at least primarily, under conditions of stability of reasons and rationality.

I contend that this objection is based on a serious misunderstanding of the fundamental role of genuine future-directed decisions and *AIDA*. It is a mistake to think that the work of decisions is either to create or surrogate the stability. Rather, the stability is presupposed in acting out of future-directed decisions. Hence, engagement in *AIDA* is a *manifestation* of the underlying stability of reasons and rationality.

5.12. Velleman has raised the worry that, by relying on stable rationality, one might trivialize the contribution of decisions and intentions to diachronic agency.[30] This worry arises, however, only for someone who embraces two erroneous ideas: (a) that future-directed decisions *necessarily* contribute to diachronic agency as effective tools for the control of future conduct, and (b) that such control must be 'remedial' of expected future irrationality. If one thinks of decisions as satisfying these two conditions, one is mistakenly identifying them with those tools, such as precommitments, that are specifically designed to address an expected instability by compensating for or providing insurance against it. Genuine future-directed decisions and intentions are not to be confused with these 'remedial tools', including certain forms of resolutions, which use bootstrapping or self-directed manipulation to address an expected future irrationality.

My insistence on the difference between decisions and 'remedial tools' should not be interpreted as a denial of any effective contribution of decisions. My point is only that this contribution is neither necessary (at least under the ideal conditions of *S*) nor

---

[30] Consider Velleman (1997), 46 where he suggests that it is a mistake to study commitments by adopting "an axiology that guarantees compatibility between the interests of our present and future selves". This is because by adopting this axiology "commitments undertaken rationally would always be seconded by the rational deliberations of later selves, whose interests would necessarily harmonize with those which made the commitments rational to undertaken. But this axiology wouldn't help to explain our actual capacity for future-directed commitments; it would merely wish away the real conflicts of interest in which the true nature of that capacity is revealed".

remedial. And even when decisions make an effective contribution, the contribution is offered only against the background of stability, since the decision is only meant to *facilitate* the agent's continuous appreciation of the stable reasons for action.

In saying that decisions rest on stability, I am not suggesting that stability is a given, something to be taken always for granted. My present point is only that this stability is presupposed in particular instances of *AIDA* and thus a necessary assumption in the account of particular decisions and intentions. A discussion of what conditions must be in place in order to assume such stability cannot be exhausted in a few lines and must be left for another occasion.[31] Finally, here I make no claim about how frequently we can rely on the stability of reasons and rationality. For my present purposes, I just need to draw the distinction between decisions (which rest on expected stability) and the manipulative remedial tools of distal self-control (which rest on expected instability). If it turns out that human nature is indeed for the most part unstable, instances of genuine decisions and *AIDA* are going to be quite rare, possibly even more than we initially thought. But this would not be a problem for my account, which has nothing to say about how often we can in fact rest on the stability.[32]

6. A Psychologically Plausible Account

6.1. Let's consider now how we can make the special scenario *S* relevant for an understanding of future-directed decisions and *AIDA* in non-idealized circumstances. The difference between *S* and the normal situations does not lie in the reliance on stability, which—as I claimed in the previous section—is an essential and distinctive feature of *AIDA*. What makes *S* unusual is rather the assumption that at any one time the agent can effortlessly and immediately access the complete case in support of her φ-ing at *f*. This is what allows the agent in *S* to be able to repeat her deliberation from scratch whenever the question of what to do at *f* arises. Of course, this is not something to be expected in most actual scenarios. The access to the complete case for φ-ing usually

---

[31] See my "Personal Identity and Diachronic Agency" (ms).

[32] The standard case in which *AIDA* finds no application is that of a reversal of preferences at *f* that is not justifiable in light of the opinions, preferences, and values of the agent at the time of the original decision. This is the typical case of instability. A different sort of case is that in which the reversal in preferences is due to a peculiar temporal structure of the situation of choice, so that the agent can find the reversal acceptable even at the time of the original decision. These are the cases (like the toxin puzzle) in which the agent is going to obtain *all* of the beneficial effects of her decision to φ at *f before* she actually has the opportunity to φ, so that rationality appears to dictate that, once she has obtained these benefits, she should no longer φ at *f*. According to my view, it is an open question whether one can engage in true AIDA in situations of this sort. If indeed it is the case that the reasons are temporally unstable as standard theories of rationality suggest, then it follows from my theory that in these cases the agent cannot have a genuine future-directed decision to φ at f. However, if the standard theory of rationality is shown to be incorrect and there is a set of stable reasons to φ, then I am committed to claiming that one can indeed have a genuine future-directed decision to φ.

requires the expenditure of scarce resources, including time. The constant repetition of full deliberation would put an unbearable strain on the agent's scarce deliberative resources. A scenario like *S* suggests that engagement in *AIDA* is the product of a sort of Cartesian 'continuous creation'. If this were true of all instances of *AIDA*, however, we would have to be agents with Herculean cognitive and deliberative powers, which of course we are not.

6.2. In the following sections, I will show how agents with limited deliberative resources and rationality like us can engage in genuine *AIDA*. We can do so, I will argue, to the extent that we rely on psychological mechanisms, techniques, and devices that allow us to overcome these limitations in a non-manipulative fashion, that is, in keeping with the spirit of the simplest and purest form of *AIDA* exemplified in *S*: In the spirit of the constant renewal of immediate responsiveness to the full set of stable reasons for action.

6.3. A first step toward a more inclusive and psychologically plausible account of decisions is to set aside a worry about those *assessments* that I claim should be renewed over time. These assessments that the agent ought to φ at *f* are not to be narrowly interpreted as explicitly entertained judgments that φ ought to be done at *f*. They are rather to be understood broadly. They are the manifestations both in thought and action of the acknowledgment of the conclusive case for φ-ing. Usually, this acknowledgment is not exhibited in the awareness of a set of articulate and explicit considerations in support of φ. The standard and basic form of acknowledgment of reasons consists in the direct and immediate shaping of one's thinking and acting in conformity with one's normative appraisals. The responsiveness to reasons is primarily manifested *in* one's own conduct and attitudes, not necessarily in either conscious or explicitly formulated judgments. When I speak of the agent assessing that she ought to φ at *f*, I only mean that *at the time of the assessment* she shapes her acts and attitudes as required by the fact that she is to φ at *f* (say, she might do something now to coordinate with her future φ-ing, take the presently available necessary means, make preparatory arrangements, avoid doing thinks that would make her future φ-ing too costly, or just let the natural course of events go on undisturbed since it is not interfering with φ-ing at *f*, etc.). My view does not require the renewal of assessment to be an explicitly entertained and consciously attended judgment that one ought to φ at *f*. Rather, it requires the renewal of a non-manipulated and direct acknowledgment of the reasons for φ-ing *in* the agent's shaping of her mental and bodily conduct.

Notice that the *immediacy* of this shaping is the mark of responsiveness to reasons in the mode of autonomy. This immediacy is to be contrasted with the indirectness of the shaping of one's attitudes and acts as a result of manipulation. Autonomous responsiveness consists in taking *direct responsibility* for one's acts and attitudes as a result of the immediate judgment-sensitivity of these acts and attitudes, to be contrasted with *taking of steps to the effect that* one's acts and attitudes will be forced to conform to one's normative appraisals.

7. Memory and Decisions

7.1. At the core of the idealization of the scenario *S* is that the renewal of a complete assessment comes at no cost to the agent. Unfortunately, this is not generally true of limited agents like us. Unlike the Herculean deliberators, we cannot shoulder the costs of a constant *rediscovery* of the complete case for φ. In the idealizing scenario, the repeated acknowledgment of reasons is the product of a renewal of both the *discovery of* and the *assent to* the reasons for φ. The element of acknowledgment that is taxing for limited agents is the discovery of reasons. If the product of discovery can be preserved over time, the burden of renewal will be limited to *assenting*, which is obviously much less costly. This distribution of the burdens is possible if the agent has a faculty of memory, i.e., a faculty for the preservation of her earlier discoveries of what the reasons for action are and what they require her to do.[33]

7.2. The intuitive way of picturing the role of memory in *AIDA* is to see it as making available to the agent at *f* the outcome of her past deliberation, when she directly acknowledged the complete case for φ. After the original decision, whenever the question of what to do at *f* arises, the agent can simply consult the record of her past conclusion and act accordingly. It is tempting to think that this memory is *all* that there is to future-directed decisions. We deliberate, reach a conclusion (the decision), store it, remember it at the time of action, and then act on it. Things might not be that simple, however. To begin with, consider again Sartre's worry as quoted in the opening of this paper. What is it about a record of a past decision that can move us to act as indicated by it? The memory of a past psychological *event* (that my past deliberation terminated in a certain assessment) does not by itself have any immediate bearing on what I am to do now. At most, I can take that event to be of indirect significance, to include it in the evidence that I take into account in my present deliberation of what I am to do now.

As long as we picture the memory as being about past psychological goings-on, the worry is correct. But what the memory primarily contributes is not a record of our past mental history. What memory retains first and foremost is not the fact that in the past a judgment with a given content was made, but the content of that past judgment. It retains the content *as* a content that was *judged* to be correct, not the *judging* of that content.

7.3. It is important to stress that the memory preserves content as judged-to-be--correct. What this means in the case of a practical assessment that one ought to φ at *f* is that the content comes to the agent at *f* with a default (although defeasible) *claim* that the agent ought to φ straightaway. To this extent, Sartre's worry is misplaced, since it overlooks that the claim does not come from the mere psychological event of judging.

---

[33] For the idea of memory as a faculty of content preservation, which has inspired my present treatment, see Burge (1993) and Owens (2000).

Nonetheless, there is a different aspect of Sartre's worry that is justified. Even if the judged content is transmitted with a default claim to be immediately accepted, the agent still needs to accept it. The force of the retained claim is not a force that makes the agent at *f* a mere tool for the execution of the past judgment. Memory is not the faculty by which the past judgment extends its executive control over time by bypassing the future exercise of rational governance. The exercise of rational governance at *f* takes the form of the agent's *assent* at *f* to the retained content. The agent at *f* acknowledges the content, once again, by immediately shaping her conduct at that time in accord with the content, i.e., by her φ-ing at *f*.

7.4. The faculty of memory does not spare us the need to renew the assessment, even if at the time of action it requires no costly discovery, but only an effortless assent. The memory helps with the renewal of assessments demanded by *AIDA* by offering to the agent at *f* a content that is, so to say, readymade for default approval. Memory spares us the need to rehearse at *f* the complete deliberation, but not the need to take the very last step of the deliberation, i.e., the *acceptance* of the deliberation's conclusion. The faculty of memory, as a faculty of content preservation rather than as a mere causal recording mechanism, offers to the agent at *f* a content with a claim of immediate significance for the immediate shaping of her contemporaneous conduct in the mode of autonomy. Still, in order for the claim's provisional justificatory force to determine the agent's conduct, she must *assent to* the claim at the present time.[34]

7.5. The present assent to the claims presented by memory seems to be the standard form in which limited beings renew their assessments over time in accordance with the demands of *AIDA*. The assent to the deliverances of memory requires no Herculean deliberative powers. It is a sort of 'degree zero of deliberation'. It is the minimal and effortless form of acknowledging reasons, since it is usually an assent by omission. The agent simply refrains from either rejecting or questioning the claim entered by memory and thereby lets this claim guide her present conduct. Notice that omitting to reject a claim is an exercise of agency. In this sense, the agent actively sustains her decision over time, even if effortlessly so.

7.6. The content offered by memory is offered as having-been-judged-correct, hence readymade not just for assent, but for immediate assent *by default*. If it were otherwise, there would be no point in relying on memory. For there would be no reduction of the burdens of the acknowledgment of reasons. One should beware, however, of the degeneration of default assent into an automatic and uncritical acquiescence. This degeneration could ultimately lead to a loss of diachronic autonomy and thus to

---

[34] See Buss (1999), 405: "In order for the resolution to constitute the gambler's present intention, he must determine (once again) that he has good *reason* not to gamble. True, his earlier resolution reflects this very evaluative judgment. Yet as one of the facts he now considers in trying to decide what to do…, its effects on his future behavior can be attributed to *him* only if this effect is mediated by his present judgment regarding its relevance to what he does next."

the complete alienation of present control to one's past self. This outcome would be unfortunate. First, there would no longer be a guarantee that one's present point of view is taken into account in the determination of what one is to do presently. Second, a systematic and complete enslavement to one's former self would ultimately be a denial of the continuous exercise of one's free agency, since one would turn merely into a conduit for the temporal transmission of the effects of one's past will.

7.7. The default omissive acceptance of the preserved content explains what Bratman calls the 'inertia' of decisions, the fact that once a decision to φ at *f* has been made, the agent at *f* is expected to φ *by default*. This does not mean, however, that one follows through a past decision in a non-agential manner, that is, as a result of the passive unfolding of some underlying psychological propensities (even if, following through might depend on the operations, sometime inertial ones, of subpersonal psychological mechanisms). In the paradigmatic case, the inertia is due to the default *active* refraining from reconsideration, rather than the absence of reconsideration. The agential character of this refraining does not mean that it is either reflective or effortful; it is simply an omission. What is important is that this omission counts as an instance of *contemporaneous* responsiveness to rational considerations, even if it takes the minimal form of the acceptance of readymade preserved content and, as such, it offers no articulation of the reasons supporting the claim to which one is presently assenting.[35]

7.8. The role of memory explains the sense in which a future-directed decision makes a difference in future conduct, even if it exerts no distal causal restraint, volitional control, or executive authority over time. The decision influences by contributing the content that is to be retained by memory. The original deliberation operates the first discovery of the reasons to φ, a discovery that, because of our limited time and resources, we are unable to repeat from scratch at *f*.

This influence is compatible with diachronic autonomy. It is not manipulative. Memory does not introduce any novel and extraneous consideration in the situation of choice at *f*. It simply offers the preserved content as the digested summary of the original merits of the case. The memory of the original judgment is a *stand-in* for the case for φ as this case was articulated at the time of the original deliberation. Hence, the

---

[35]   Diachronic autonomy requires more than the agent's answerability to reasons over time. It calls rather for the renewal of what I called the 'responsiveness' to reasons. The latter concerns the agent's shaping of her acts and attitudes in direct response to her normative appraisals (even if only implicit and non-reflective ones). It is the immediate manifestation of the agent's rational agency. Answerability to reasons, instead, can be a more indirect affair. The agent's acts and attitudes can be answerable to reasons to which she is not presently directly responding (for instance, the inferential consequences of her present assessments). Likewise an agent who is presently unable directly to respond to normative appraisals (say because she put herself under a physically constraining device) might still be answerable to reasons, in that she is expected to be able to justify her choice to suspend the exercise of her rational governance. In any event, answerability to reasons requires that the agent, if challenged to live up to its demands, be directly and correctly responsive to those reasons.

memory allows the agent at *f* to acknowledge, although indirectly, a better articulated case for φ. This acknowledgement, even if indirect, is an instance of immediate rational responsiveness to reasons. It takes the form of the default acceptance of the preserved content as manifested in the agent's φ-ing at *f* without further ado.

8. The True Memory of The Will

8.1. In the previous sections, I argued that the combination of memory as content preservation with the default renewal of assent can successfully account for *AIDA*. This is not to say that memory is essential to *AIDA*. Memory is often an indispensable tool for limited beings like us. But it only facilitates our exercise of *AIDA*, which in principle can do without memory. At the core of *AIDA* is not content preservation, but the active renewal of the acknowledgment of the original, unadulterated merits for φ. This renewal is manifested in how the agent shapes her conduct over time in her immediate response to a continuous acknowledgment of the stable case for φ. The acknowledgment is often done with the help of retained content, but this is not essential to *AIDA*. The acknowledgement might just as well be constantly renewed *ex nihilo*.

To use Nietzsche's famous expression, the 'true memory of the will'[36] that is required by *AIDA* is not the ability to remember past assessments. It is rather the ability *consistently* to take the actions and attitudes that, at each moment, are required to sustain one's φ-ing at *f*; and to do so out of a *constantly* renewed autonomous acknowledgment of temporally stable, non-manipulated reasons for φ-ing. This is exactly what happens in the scenario *S* illustrated before, which is the reason why I have used this scenario to illustrate the structure of *AIDA* in its purest form.

8.2. The reliance on memory is thus dispensable in principle, but occasionally in practice as well. Future-directed decisions are not necessarily efficacious. They do not have to make a difference in order for the agent to be engaged in *AIDA*. At the same time, when they do make a difference, this need not be via the sort of content preservation illustrated above. What matters for *AIDA* is that diachronic autonomy be respected in a non accidental way. The facilitation of the renewal of assessment can take a variety of forms, as long as they are not manipulative. The default renewal of the acceptance of preserved content is an important way, possibly the most common one, in which limited agents like us engage in *AIDA*. But it is not the only way.[37] Genuine instances of *AIDA* might be produced by some of the mechanisms championed by the traditional accounts.

---

[36]   Nietzsche (1887).

[37]   I illustrate some of these ways in "Decisions, Diachronic-Autonomy, and the Division of Deliberative Labor" (ms).

9. A Reassessment of the Traditional Views

9.1. The traditional views of future-directed decisions describe actual cases of distal self-control. Some of these cases respect diachronic autonomy and thus count as genuine instances of *AIDA*. But is it a mistake to elevate any of these accounts to the status of a general and comprehensive theory of decisions. When they are generalized, the traditional accounts run into two problems. First, they are phenomenologically implausible, since only a portion of the actual instances of *AIDA* conforms to their explanation. Second, the traditional accounts fail to understand the ultimate structure of *AIDA*. As a consequence, they do not to realize (a) that the mechanisms they champion are only *some* among the many different instruments for the facilitation of *AIDA*; and (b) that, under favorable conditions, all these instruments are dispensable.

9.2. There are also difficulties specific to the various views. Let me illustrate some of them here. Consider first the accounts based on the idea of decision-based reasons. These accounts correctly stress the active dimension of the renewal of the acknowledgment of reasons. However, they do so by offering an account centered around a rather special scenario, that of a reflective agent who is considering whether she is justified in relying on the mechanisms that transmit the conclusions of her prior deliberations. If this is how one pictures the predicament of the agent at the time of action, the idea of a decision-based reason is indeed very attractive. This is because the reflective agent, once presented with a record of her past decision, will raise the question: What should I do now with this record and why? In response to this, she might think of the record as equivalent to a self-directed command or advice, or a source of evidential reasons about what she would to do now, where she to engage in full deliberation. By drawing an analogy with what happens in the interpersonal counterparts of these scenarios, the agent takes herself to have a conclusive reason to act as originally decided. In the predicament of the reflective agent, the acknowledgment of the decision-based reason is a crucial explanatory factor of her following through on her past decision. This does not mean, however, that this explanation covers all cases of *AIDA*. To think otherwise is to offer an overintellectualized picture of the nature of future-directed decisions; a picture that is both phenomenologically implausible and explanatorily unconvincing.

The theories in terms of decision-based reasons are not to be confused with the theory formulated in terms of the *entitlement* to the content preserved by memory. In the latter case, the acceptance does not depend on the agent's contemplation of a decision-based reason. The agent might appeal to such a reason only in case she were challenged to justify her reliance on the record of her prior judgment. In the absence of this challenge, however, the agent acts on this record without taking into account any decision-based reason. As a stand-in for her prior conclusion, the record offered by memory allows the agent at *f* to be *directly* responsive, in her φ-ing, to the considerations that supported her prior conclusion. There is no need for the agent to appeal to any

intermediary decision-based reason unless she wants to articulate the grounds for her entitlement to rely on the deliverances of either memory or some other device that records and transmits her past decisions.[38]

9.3. The previous conclusion appears to support the propensity view. According to it, under normal conditions, the agent φ's at *f* without consulting any decision-based reason. She simply let herself be guided by some underlying propensity to grant that the records of her past decisions are correct and to be acted upon. Hence, the agent does φ at *f* without further ado; she φ's without reconsidering her past decisions. The default acceptance of past records and, more generally, of the working of the mechanisms of content-preservation are indeed the outcome of underlying psychological propensities.

This is not a vindication of the propensity view of the traditional sort. The standard interpretation of the propensity view is not formulated in terms of the specific mechanisms that underlie the work of the faculty of memory. Rather, the standard view makes a more generic appeal to the existence of a set of psychological dispositions, habits, and propensities that make the agent at *f* act in accordance with her prior decision. The problem with this view is that the generic appeal to the existence of underlying psychological propensities is not sufficiently discriminating between the various forms of temporally extended agency. The agent's psychological habits, disposition, and propensities play a role in the agent's diachronic agency in any of its forms. The goal of a theory of decisions, however, is to offer an account of a specific form of diachronic agency. Not all the ways in which our psychological propensities lead us to act in accordance with our past decisions can be counted as genuine instances of *AIDA*. This is the point of the criticism I raised above when I contrasted the inertial passivity of the yawning picture of decisions with the agential activity of the renewal-of-acceptance picture. According to both models, the agent's φ-ing at *f* is largely determined by the unfolding of her psychological propensities. But it is only in the renewal-of-acceptance scenario that the operation of the propensities amounts to an active, personal-level, non-inertial involvement of the agent in the initiation of the action at *f*. This sort of involvement is necessary if the reliance on the propensities is to count as a genuine instance of *AIDA*. This is not to deny that there can be inertia in the working of our subagential psychology. There is bound to be, if only because of the ultimate physical nature of the psychological processes. The sort of inertia that is troubling for diachronic autonomy is inertia at the agential level. This is because diachronic autonomy demands that the agent actively exercise both her rational responsiveness and governance in

---

[38] A further problem with some of the decision-based reason views, on which I am not going to dwell in this work, is that they characterize decisions as the intrapersonal analog of interpersonal phenomena such as issuing commands or Gricean communication. The risk is that, if this suggestion is literally interpreted, it fails to recognize the self-directedness of decisions as a truly distinctive property of them.

following through on her past decisions. This is thus the sense in which Sartre is right in claiming that there is no "inertia in consciousness".[39]

10. Conclusion

10.1. In this paper, I have argued that we can understand the role of future-directed decisions in shaping future conduct only by investigating their contribution to a distinctive form of diachronic agency, what I called 'autonomous intentional diachronic agency' or *AIDA*. This form of diachronic agency is different from manipulative distal self-control (produced by either physical coercion or precommitment). At the core of *AIDA* lies the respect of the agent's diachronic autonomy. Diachronic autonomy is secured in a paradigmatic form by the agent's renewal over time of the direct and full acknowledgment of stable reasons for action. In principle, *AIDA* does not require the contribution of effective future-directed decisions. Nevertheless, because of the limitation in our deliberative resources, in most cases we use a variety of techniques to distribute over time the costs and burdens of the repeated acknowledgment of reasons. Such techniques count as implementations of *AIDA* only to the extent that they do not jeopardize the continuity of our responsiveness to the stable reasons for action, a continuity that is demanded by diachronic autonomy.

There is a variety of ways in which *AIDA* can be implemented (the most common being the retention of past assessments via memory). The basic fault of the standard theories is that they ignore this variety by focusing exclusively on one of the possible implementations of *AIDA*. In actuality, there is really no privileged or unique mechanism underlying the contribution of genuine future-directed decisions. What matters is only that the various mechanisms help limited beings like us with securing the respect of diachronic autonomy, under the guidance of the regulative ideal of the purest form of *AIDA*. Each of these mechanisms is thus only instrumental to securing the essential feature of *AIDA*, viz., the non-manipulated renewal of the acknowledgment of stable practical reasons. Each of the mechanisms is in principle dispensable, not only because alternative mechanisms might be available, but also because the agent might be under favorable conditions that allow her to renew the full acknowledgment of reasons at no cost.

A comprehensive theory of future-directed decisions must make room for the variety of ways in which decisions contribute to *AIDA*. This is the suggestion that I have outlined in this paper. I have not attempted, however, a systematic survey of all the ways in which *AIDA* can in fact be implemented.[40] My main concern has rather

---

[39] Sartre (1956), 104.

[40] There are circumstances in which a temporary suspension of diachronic autonomy causes no serious loss and we might find easier to let some inertial or manipulative processes take care of the progress of our

been twofold. First, to outline what lies at the core of *AIDA*—i.e., the renewal of the acknowledgement of stable reasons— and, second, to show that there is no need for either a unique or privileged way of securing the essential structure of *AIDA*. Not to mention that, in the purest instances of *AIDA*, the agent is expected to do without effective future-directed decisions altogether.

10.2. By bringing to light the basic structure of autonomous diachronic agency, it becomes possible to explain the source of the puzzling features of decisions mentioned at the outset. Once we appreciate that autonomous intentional diachronic agency requires the *renewal* of our acknowledgement of reasons, we realize that diachronic agency in the mode of autonomy requires that our projects be *actively sustained* over time. If we thought that taking a future-directed decision would relieve us of this burden, we would be surprised by Sartre's claim about the effect of the passage of time on the force of our resolutions. But by now we should be able to appreciate in which sense a decision is not truly binding on our conduct independently of our renewal of the acknowledgement of the reasons that support it.

Sartre is also right in claiming that there is something *ex nihilo* in the renewal of our past decisions. To see this, consider once again the two components of the acknowledgment of reasons, what I have called the 'discovery' and the 'acceptance' of reasons. The discovery is often onerous. Hence, limited beings like us, who lack Herculean deliberative powers and resources, must take advantage of the various techniques for the preservation and transmission of their prior discoveries. The combined operation of these techniques accounts for the effectiveness of future-directed decisions. We do not get, however, a comparable help with the burden of renewing the acceptance of reasons. To be released from that burden would amount to relinquishing the exercise of our distinctive form of temporal agency, viz., temporally extended intentional agency in the mode of diachronic autonomy. In this sense, the renewal of acknowledgement is *ex nihilo*. We continue to be under the bond of a past decision only as long as we continue to renew our acceptance of the stable practical reasons that supported that decision.[41] Hence, as temporally-extended

---

extended projects. An example would be that in which the future φ-ing is a relatively simple action and there is nothing that the agent needs to do between now and the future time *f* but to wait for the time of action to be ripe (she needs not prepare for φ-ing, remove obstacles, or fend off interferences). In situations of this sort we might revert to manipulative techniques of distal self-control (like an internal delaying mechanism, a lit-fuse or the likes), without any dramatic loss to the ultimate success of the project and the preservation of our overall diachronic autonomy. Although, strictly speaking, these cases are not instances of *AIDA*, they can be quite good and useful *surrogates* of it. When speaking in the 'loose and popular' way, we might be excused for taking them as instances of genuine future-directed decisions. This is not troubling as long as we understand that they are just degenerate forms of *AIDA*, which have been adopted only in order to promote the regulative ideal of diachronic autonomy. What we should resist, instead, is the use of these special and simplified scenarios as providing the model cases for the investigation of intentional diachronic agency, something that unfortunately occurs too often in the contemporary discussion.

41   See Moran (2001), 82: "The resolution is mine to keep or to break at any time. For me the fact of my

autonomous agents we can do without the strengths of Hercules, but we always need some of the active perseverance of Sisyphus.[42]

REFERENCES

Ainslie, G., (1992), *Picoeconomics: The Strategic Interaction of Successive Motivational States Within the Person,* Cambridge: Cambridge University Press.

Ainslie, G., (2001), *The Breakdown of Will,* Cambridge: Cambridge University Press.

Anderson, E., (1996), "Reasons, Attitudes, and Values: Replies to Sturgeon and Piper", *Ethics,* 106, 538-544.

Bratman, M., (1987), *Intention, Plans and Practical Reason,* Cambridge: Harvard University Press.

Bratman, M., (1999), *Faces of Intentions,* Cambridge: Cambridge University Press.

Burge, T., (1993), "Content Preservation", *The Philosophical Review*, 102, 457-488.

Buss, S., (1999), "What Practical Reasoning Must Be If We Act For Our Reasons", *Australasian Journal of Philosophy*, 77, 399-421.

Castañeda, H-N., **(1975),** *Thinking and Doing: the Philosophical Foundations of Institutions*, Dordrecht: Reidel.

Cohen, G.A., (1996), "Reason, Humanity and the Moral Law", in Korsgaard, C., (ed.), Th*e Sources of Normativity*, Cambridge: Cambridge University Press, 1996.

Davidson, D., (1980), *Essays on Actions and Events,* Oxford: Oxford University Press.

Elster, J., (1979), *Ulysses and the Sirens,* Cambridge: Cambridge University Press.

Ferrero, L., "Personal Identity and Diachronic Agency" (ms).

Ferrero, L., "Decisions, Diachronic-Autonomy, and the Division of Deliberative Labor" (ms).

Fodor, J., (1987), *Psychosemantics*, Cambridge: MIT Press.

Hinchman, E., (2003), "Trust and Diachronic Agency", *Noûs*, 37, 25-51.

Holton, R., (1999), "Intention and Weakness of Will", *Journal of Philosophy*, 96, 241-62.

Joyce, J., (1999), *The Foundations of Causal Decision Theory*, Cambridge and New York: Cambridge University Press.

Kant, I., (1797), *Metaphysical Principles of Virtue,* Cambridge University Press, 1996.

resolution cannot be something on which I may base a confident prediction, because I must recognize that the resolution only exists as a fact for anyone to base a prediction on insofar as I continue to endorse it. And my endorsement is in principle answerable to how I relate to the reasons in favor of some course of action. They are what brought me to the decision in question in the first place and the decision is only decision is only as strong as my hold on those practical reasons".

Kenny, A., (1963), *Action, Emotion and Will,* London: Routledge.

McClennen, E., (1990), *Rationality and Dynamic Choice,* Cambridge: Cambridge University Press.

McClennen, E., (1997), "Pragmatic Rationality and Rules", *Philosophy and Public Affairs*, 26, 210-58.

Moran, R., (2001), *Authority and Estrangement, An Essay on Self-Knowledge,* Princeton: Princeton University Press.

Nietzsche, F., (1887), *On the Genealogy of Morality,* Indianapolis: Hackett, 1998.

Owens, D., (2000), *Reasons without Freedom,* London: Routledge.

Pink, T., (1996), *The Psychology of Freedom*, Cambridge: Cambridge University Press.

Rachlin, H., (2000), *The Science of Self-Control,* Cambridge: Harvard University Press.

Rovane, C., (1998), *The Bounds of Agency,* Princeton: Princeton University Press.

Rundle, B., (1997), *Mind in Action,* New York: Oxford University Press.

Sartre, J.P., (1956), *Being and Nothingness*, New York: Pocket Books.

Schelling, T., (1980), *The Strategy of Conflict*, Cambridge: Harvard University Press.

Shapiro, S., (2001), "Judicial Can't", *Philosophical Issues*, 11, 530-57.

Sobel, J.H., (1994), "Useful Intentions" in *Taking Chances,* Cambridge University Press.

Velleman, D., (1989), *Practical Reflection,* Princeton: Princeton University Press.

Velleman, D., (1997), "Deciding How to Decide", in Cullity, G. and Gaut, B., (eds.), *Ethics and Practical Reason,* Oxford University Press, (1997). Reprinted in Velleman, D., (2000), *The Possibility of Practical Reason*, Oxford: Oxford University Press.

Wilson, G., (2000), "Proximal Practical Foresight", *Philosophical Studies*, 99, 3-19.