A Hybrid View of Commitment[1]

Facundo M. Alonso
Miami University, Ohio

(Penultimate draft. Final version forthcoming in *Oxford Studies in Agency and Responsibility*, David Shoemaker (ed.). Vol. 9. Oxford: Oxford University Press.)

Abstract

We often appeal to the notion of an agent's commitment to action to characterize, for example, an agent's faithfulness to a promise she has given to another, her robust disposition to pursue a goal she values or cares about, and her determination to stick to that goal. In the philosophy of action, that notion is often associated with the idea of an agent's intention to act. In ethics, it is associated primarily with the idea of an agent's commitment to, or endorsement of, a certain norm. This chapter argues that both these ideas, of intention and of normative endorsement, are central to the –or, at least, an—ordinary notion of an agent's (psychological) commitment to action. In so doing, the chapter also sheds light on important features of intention.

Keywords: commitment – intention – normative endorsement – stability across contexts – coordination

1. Introduction

The notion of an agent's commitment to action is central to much philosophy of action and ethics, and yet it appears at times mysterious and hard to pin down (see, Arruda 2022 for a recent overview). We often appeal to this idea to explain, or characterize, among other things, an agent's faithfulness to a promise she has given to another, her robust disposition to pursue a goal she values or cares about, and her determination to stick to that goal. In the philosophy of action, the notion of an agent's commitment to action is often associated with that of her intention to act (Harman 1976, Bratman 1987, Velleman 2007; see, Davidson 1978 for dissent). In ethics, talk of an agent's "commitment" to action is associated first and foremost with her commitment to, or "endorsement" or "acceptance" of, a certain norm (Korsgaard 2008; Cf. Gibbard 1990, Railton 2006). I think that both these ideas, of intention and of normative endorsement, are, when properly fleshed out and interconnected, central to the –or, at least, an—ordinary notion of an agent's

[1] A version of this chapter was presented at the Seventh New Orleans Workshop on Agency and Responsibility (NOWAR). I am grateful to that audience for valuable discussion. I am indebted to Santiago Amaya, Gunnar Björnsson, Karin Boxer, Eric Brown, Luca Ferrero, Sebastián Figueroa Rubio, Erin Kelly, Taylor Madigan, Ben Mitchell-Yellin, Doug Portmore, Peter Railton, Jules Salomone-Sehr, David Shoemaker, Kenneth Silver, Michael Smith, Manuel Vargas, and Eric Wiland for insightful comments and suggestions. Special thanks to Sarah Paul for helpful comments on an earlier draft.

(psychological) commitment to action.[2] In this chapter, I articulate and defend a view of commitment along such lines. I call this the "hybrid view" of commitment.

Commitment is a more robust phenomenon than is mere intention. When one intends to act in a certain way, one is willing to persist in this attitude until carried out, so long as one does not acquire relevant new reasons to do otherwise. But when one is committed to so act, one is willing to retain the commitment even in the face of some of those reasons. It is a central feature of commitment that it tends to persist even when circumstances change in relevant ways. Because of this, commitment is a more effective tool for coordination. We live in a changing world, a world that constantly produces new incentives for action. Therefore, in order to secure coordination in such a world, we need an attitude that is able to withstand such changes. Such an attitude, I submit, is commitment. The hybrid view, I argue, explains these features of commitment and in doing so provides us with insights not only into the nature of this attitude but also into the related attitude of intention.

2. Is commitment to action a mere intention to so act?

Recent discussions in the philosophy of action have called attention to the existence of a tight connection between intention and commitment (Bratman 1987; Cf. Velleman 2007). Michael Bratman, in particular, has elaborated on this connection with remarkable insight (1987, 1999, 2007, 2018). Bratman maintains that "intention … involves a characteristic kind of commitment [to action]" and that this commitment has two aspects (1987: 15). First, intention involves a commitment to engage in certain forms of practical reasoning that are conducive to its successful execution and to forgo engaging in others that are detrimental to it (16-17). In forming an intention to do something, one thereby *settles* the question of whether to do that thing in one's mind and is as a result resolved or determined to do it. Moreover, in intending to do that thing, one is disposed to *see* the question of whether to do it *as settled*.[3] This helps to explain why in intending to do something, one is disposed to refrain from reopening deliberation as to whether to do that thing, absent new relevant reasons to do otherwise. Provided that circumstances do not change in relevant ways, one will normally avoid reconsidering one's intention and retain it until it is carried out. Intention thus exhibits a sort of "inertia" (16). But not only does intention dispose one to avoid engaging in certain forms of reasoning, such as arbitrarily reopening deliberation about a question one has already settled. It also disposes one to engage in others. For example, in having an intention for a certain end, one is disposed to form intentions about means to that end. One is disposed to do this partly because, in intending to attain

[2] Here I put aside conceptions of commitment as a normative phenomenon or relation, such as those advanced by, respectively, Margaret Gilbert (2014) and Sam Shpall (2014).

[3] The connection between intention and settling is originally highlighted by Harman (1976: 438ff). I discuss aspects of it in Alonso (2017).

a certain end, one is settled on the question of whether to pursue that end and also tends to see that question as settled in one's mind.[4]

Things are different with desire. It is part of the normal functioning of this attitude that one may desire to do something and continue to reconsider the question of what to do even when circumstances do *not* change in relevant ways. This is because, unlike intention, desire settles nothing. Furthermore, since it is not part of the functional profile of desire to settle the matter of what one will do, it is not surprising that in having a desire for a certain end, one is not necessarily disposed to pursue the means to that end. For if one is not settled on the question of whether to pursue an end, why would one bother to settle on the means to it?

According to Bratman, the commitment characteristic of intention involves not only dispositions toward reasoning but also a disposition toward action (1987: 15-16). It is clear that both intention and desire involve (what we may call) a basic disposition to bring about action. Accompanied with relevant beliefs, each of those attitudes can move one to act. However, intention involves a stronger disposition to action than does desire. You may now desire to watch the last Scorsese movie tonight, believe that by going to the local theater you can accomplish this, continue to have that desire until the time of action, and fail to be moved to go then. (After all, you may have settled on an alternative course of action, say, reading a novel tonight.) All the while, your desire to watch the movie may be functioning properly. In contrast, were you now to intend –rather than merely desire—to watch that movie tonight and to have the cited belief, were your intention to survive until the time of action, and were nothing to interfere, you would normally at least try to go to the local theater then. For you would now be settled on watching that movie tonight, which would then normally lead to your settling that issue in the world.[5]

Undoubtedly, there is more to be said about Bratman's conception of the relation between intention and commitment. But what I want to highlight for present purposes is a central feature of this conception. As we have just seen, Bratman characterizes intention as involving a sort of commitment and understands the latter, roughly, in terms of a relevant cluster of dispositions toward reasoning and action. Therefore, talk of a "commitment" is, on this conception, talk of a cluster of dispositions that partly constitutes intention.

The phenomenon of commitment I am inquiring about in this chapter differs from what Bratman understands by "commitment" in his discussion of the nature of intention. What I have in mind is a

[4] Bratman thinks that the fact that intention involves such reasoning dispositions is not just a brute fact about this phenomenon, that is, it is not just that we are simply wired that way. It is also partly a result of the fact, he maintains, that, as intending agents, we endorse certain norms of rational coherence as governing this attitude (2018). These norms enjoin the agent to, among other things, persist in her intention unless she acquires strong reasons for revocability and/or reconsideration (diachronic stability) and intend the believed necessary means to her intended end (means-end coherence).

[5] On the idea that intention involves a stronger commitment to bringing about action than does ordinary desire, see also, for example, Velleman (1989: 281-84; 2000: 202-4). However, Velleman thinks, unlike Bratman, that this commitment involves belief in success.

phenomenon captured, I think, by our ordinary talk of an agent's "commitment to action" and, as I see it, this is, in a sense to be specified later, a more robust phenomenon than is mere intention. I argue that a commitment to action involves an intention to so act, but the latter is not sufficient for the former. Commitment to action is a complex psychological attitude that involves intention *and more*. So, while Bratman proposes that we understand intention in terms of a certain idea of commitment (understood as a cluster of dispositions), my proposal in this chapter is that we understand ordinary commitment partly in terms of intention. It seems clear to me that an agent's *ordinary* commitment to action –I drop the qualification in what follows—involves the cluster of dispositions commonly attributed to intention, and hence it is natural to think of the former phenomenon as being partly constituted by the latter. However, as suggested above, it also seems clear that commitment involves more than intention. A main reason for believing this, I contend, is that the former phenomenon exhibits greater stability than the latter does. In addition to possessing the "inertia" characteristic of intention, a commitment to action possesses a sort of counterfactual robustness. Not only does it tend to persist in the absence of relevant reasons to do otherwise, but it also tends to persist in the *presence* of some of those reasons. And this supports the conjecture that intention, although necessary for commitment, is not sufficient for it.

3. Two sorts of stability of intention

In order to substantiate these points, we need first to have a deeper look at the stability of intention. For reasons that will become apparent later, it is important that we begin by distinguishing between two cases of intention. Normally, when one forms an intention to act in a certain way, one does so on the basis of reasons. One deliberates about what to do, assesses the relative weight or normative force of (what one regards as) various competing reasons for action, endorses some of such reasons as supporting a certain option, and, on that basis, forms the corresponding intention. It is quite plausible that when you formed the intention to go to the movie theater, you did so by way of a process like this. You reflected on the matter of what to do this evening, assessed the considerations for and against the relevant options, and formed the relevant intention partly on the basis of the judgment that watching the last Scorsese movie would be entertaining –and also, perhaps, the best option. However, not all cases of intention formation are a result of a conscious process of deliberation about what to do, in which one assesses the relative weights of various competing reasons for action and endorses some of them as normative for one's intention. In some cases, one arrives at intention through a nondeliberative process: one comes to form an intention in ways that are unconscious and insensitive to reasons. For the sake of convenience, I call the former, cases of *reasons-based intention*, and the latter, cases of *mere intention*.

Let us take a closer look at the latter cases. To get a better sense of what such cases look like, we can find inspiration in Elizabeth Anscombe's discussion of the relation between acting intentionally and

acting for a reason. Anscombe proposed a nuanced conception of this relation. She claimed that while it is true that whenever one acts for a reason one acts intentionally, it is not true that whenever one acts intentionally one acts for a reason.[6] Anscombe maintained that intentional actions are partly distinguished from nonintentional ones by their being subject to a certain question: "they are the actions to which a certain sense of the question 'Why?' is given application; the sense is of course that in which the answer, if positive, gives a reason for acting" (1963: 9). Suppose that I ask you, "Why did you go into the local theater?" and you respond, "Because I wanted to watch the last Scorsese movie" or "Because I thought that watching that movie would be fun." In this case, the question "Why?" has application to your action of going into the movie theater, and your positive answer to it gives the reason for which you acted –that is, you wanted to watch the movie, or thought that doing so would be fun. Nevertheless, Anscombe also maintained that there are cases in which the question "Why?" has application and yet the answer to it does not identify "a reason for acting." These are cases in which the agent answers that question by saying things like, "'It was an impulse' or 'For no particular reason' or 'It was an idle action—I was just doodling'" (1963: 25). Each of these answers indicates that the agent had no reason for acting and yet the question "Why?" can have application, Anscombe explained, just as the question how much money I have in my pocket has application when I respond "None" (25).

Like Anscombe, I believe that one can do something intentionally "for no particular reason." More importantly for our purposes, I believe that one can also *intend* to do something "for no particular reason."[7] I should clarify how I understand this expression. My understanding differs from an influential interpretation offered by Donald Davidson.[8] In his seminal article "Action, Reasons, and Causes," Davidson suggested that what we mean to say when we speak of an intention to do something "for no reason" is that in the case in question "there is no *further* reason … no reason, in other words, besides [intending] to do it" (1963: 6. The emphasis is his.). In effect, according to Davidson, to say that an agent intends to do something "for no reason" is in fact to say that the agent has a reason for doing this and that this reason is provided, not by some independent fact, but by the very fact that she intends to do it. Clearly, Davidson's suggestion goes hand in hand with his well-known thesis that to intend to do something is partly to judge that one has a reason for doing it. Unlike Davidson, I reject the just mentioned suggestion and related thesis.[9] When I speak of an intention to do something "for no reason" what I mean to say, instead, is quite literally that "there is no … reason," that is, that the intention is the product, not of the assessment and endorsement of reasons, but of an unconscious or nondeliberative process. Granted, it is plausible that cases of intention

[6] This contrasts with a familiar conception, offered by Davidson (1963) as well as by many others, according to which to act intentionally *is* to act for a reason.
[7] Anscombe appears to agree on this (1963: 90). For discussion, see Wiseman (2016).
[8] Here I put aside the controversial question of how Anscombe herself understands this expression.
[9] For critical discussion of that thesis, see, among others, Velleman (1992).

formed for no reason are not as pervasive as cases of intention formed for reasons.[10] But they are nonetheless genuine cases of intention. According to functionalist views of the type invoked above, intention is partly distinguished by a relevant cluster of dispositions toward reasoning and action. And it seems to me that intention can exhibit these basic forward-looking (or downstream) features independently of the backward-looking (or upstream) issue of how it was formed.[11]

There are two main reasons why the contrast between cases of mere intention and cases of reasons-based intention matters to the question of the stability of intention. The first is that this contrast allows us to identify the stability of intention *as such*. Reasons-based intention involves the endorsement of reasons and it is intuitively obvious that such normative endorsement can contribute –in a way that needs to be specified—to the stability of this attitude. Therefore, if we want to accurately represent the stability of intention as such, it seems that we should focus on cases of mere intending. Were we to focus instead on cases of reasons-based intention, we would risk attributing to intention features that belong in fact to the endorsement of reasons. To my knowledge, discussions of the stability of intention rarely include, if ever, consideration of cases of mere intending. If what I said above is correct, this constitutes an important deficit in the literature. In what follows, I aim to remedy this deficit. The second reason the aforementioned contrast matters is that it allows us to differentiate, regarding the stability of reasons-based intention, between the contribution made to it by the intention itself and the contribution made to it by the endorsement of reasons. And this helps us to explain, in turn, why reasons-based intention is more stable than is mere intention. (I will say more about this below.)

Consider a version of Anscombe's case of doodling (1963: 25). Suppose that you are sitting at a coffee shop, having your morning coffee, and that at some point you start doodling. You do so for no reason. You are in control of your action in the sense that you are tracking its success. For example, were someone to distract you momentarily –say, by asking you for the time—you would be disposed to return to your activity and to continue with it for some time. Moreover, you are tracking the success of your action of doodling partly because you *intend* to doodle. So, during that time, you intend to doodle, and are doodling intentionally, for no reason. Now consider a variation on this case. Suppose that after a few minutes of being engaged in this activity, one of your neighbors –call him, Nice Guy—shows up at the coffee shop. You take this to give you a sufficient reason to stop doodling and to strike up a conversation with him, and you accordingly respond to it by replacing one intention with another. Furthermore, although in this scenario you abandoned your intention to doodle for a particular reason, it is likely that you would have

[10] Cf. Grice: "we do not, normally at least, find ourselves with intentions; we form intentions [for reasons]" (1971: 267).

[11] I suggest in sections 6 and 7, however, that when an intention is formed on the basis of reasons, this reshapes the attitude's basic dispositions toward reasoning and action. For talk of "forward-" and "backward-looking" features of intention, see Bratman (1987, Ch. 10).

abandoned this intention for many other reasons –say, if you saw a friend walking on the other side of the street or suddenly remembered that you had to run an errand. This leads to a more general observation. In cases of merely intending to φ –that is, of intending to φ for no reason—almost any (sufficient) reason one may newly acquire to pursue an alternative course of action, ψ-ing, may appropriately lead one to abandon one's intention to φ and form the intention to ψ. And this indicates that intention is *as such* less stable than what it is usually taken to be.

Consider next another variation on the original case. Suppose that, in a third scenario, a different neighbor –call him, Obnoxious Guy—enters the coffee shop. You find Obnoxious Guy annoying and think that by continuing to doodle you will be able to avoid him. You take this to give you a reason for doodling, which you did not antecedently have, and respond to it accordingly. This involves basically two things in this scenario. First, you incorporate this new reason in your practical reasoning. In particular, you take this reason as grounds for retaining your present intention to doodle. Second, you act accordingly, that is, you continue doodling. Your endorsement of and responsiveness to this new reason for action have an important effect on your intention in this third scenario, namely, they enhance the stability of this attitude. Your intention now exhibits a sort of stability that goes beyond the stability of intention as such. It is not just that you continue to be disposed to persist in your intention to doodle other things being equal, that is, to persist in it only insofar as you do not acquire relevant reasons to do otherwise. You are now also disposed to persist in your intention even if you acquire some of those reasons. Put simply, in having endorsed and responded to the aforementioned reason to doodle, you are now disposed to persist in your intention in circumstances in which you previously would not have.

To see this, imagine that in this third scenario, your spouse, who is sharing a table with you, proposes to leave the coffee shop –say, in order to run an errand with them—right at the time in which Obnoxious Guy walks in. It seems that in those circumstances you will likely ask your spouse to stay at the coffee shop for a little longer and continue to doodle, in order to remain unnoticed. But had your spouse made that request in the original scenario –in which you were idly doodling and not trying to avoid an unpleasant encounter—you would have likely acquiesced to it and abandoned your intention to doodle right away. In other words, in the third scenario, you are less willing to abandon your intention in response to your spouse's request than you would have been in the original scenario. The explanation is simple. Since in the original scenario your intention is not based on any reasons for action, you are disposed to easily change your mind. But, in the third scenario, you have created for yourself a barrier to doing this; you have endorsed a reason and based your intention on it. Indeed, in this latter scenario not only do you intend to doodle for a reason, but you also take this reason to outweigh some reasons to do otherwise. You deem it more important to avoid unpleasant company than to immediately acquiesce to your spouse's request. And, since you are responsive to reasons, you finally stick to your original plan.

The preceding remarks support a distinction between two sorts of stability a particular intention may exhibit. First, there is a sort of stability that is essential to this attitude. Intention is, as such, *stable within a context* –the context in which it is formed. This is basically what Bratman and others identify as the (rational) stability characteristic of intention, as mentioned in section 2. Intention tends to persist until the time of action without reconsideration, provided that circumstances do not change in relevant ways between the time in which it is formed and the time in which it is carried out. In other words, one will normally persist in one's intention only insofar as one does not acquire stronger or weightier reasons to do otherwise. We may call this the "ceteris paribus inertia" of intention, since it is the stability displayed by this attitude when other things are equal. This helps to explain why intention is as such more robust than is desire, for arguably this latter attitude need not exhibit a tendency to persist even when other things are equal. Having said that, I will argue below that although the fact that intention possesses ceteris paribus inertia portrays it as a robust attitude relative to desire, it does not portray intention under such positive light when compared against the related phenomenon of commitment.

We gain a more comprehensive view of the ceteris paribus stability of mere intention if we look at it from the perspective of an observation made in section 1. As noted, we live in a changing world and therefore it is unsurprising that the circumstances that follow the formation of an intention may differ quite dramatically from the circumstances in which it was formed. New circumstances may create reasons to pursue an alternative course of action, and sometimes an appropriate response to such a change will be to abandon one's prior intention before the time of action has arrived. Cheshire Calhoun makes a similar point in a related discussion (2009: 625). She notes that whether an intention persists until the time of action in the particular case is partly dependent on factors that are external to the intention itself; it partly depends, she says, on the relevant changes that may or may not occur in the circumstances as well as on one's awareness of and responsiveness to them. Now, reflection on cases of mere intention allows us to take the aforementioned observation one step further. It is surely important to acknowledge that the circumstances that follow the formation of an intention to do something may change in ways that give one a reason to do otherwise and that this may appropriately lead to the abandonment of that intention. But it is also important to recognize that this reason need not be particularly strong or weighty in order to justify abandoning that attitude. This is one of the lessons we learned from the doodling example above. In having a mere intention to φ –that is, in intending to φ for no reason—one is rationally disposed to abandon one's intention in the face of pretty much *any* sufficient reason to not φ. Therefore, the fact that one's intention exhibits ceteris paribus inertia does not make it a remarkably robust attitude, since it does not prevent one from rationally abandoning this attitude in response to the acquisition of a fairly weak reason to do otherwise.

Second, there is a sort of stability that is possessed to a higher degree by some particular intentions than by some others. Specifically, in addition to being stable within a context, an intention can be *stable*

across contexts. It can persist until the time of action not only in a scenario in which one's reasons for action remain the same but also in several counterfactual scenarios in which one acquires relevant reasons to do otherwise. When that happens, we say that the intention exhibits a certain degree of *counterfactual robustness*. The wider the range of counterfactual scenarios in which a particular intention tends to persist, the more counterfactually robust it is. We saw above that intention does not possess *as such* a high degree of counterfactual robustness. Provided one is rational, one will tend to abandon a mere intention if one simply acquires a sufficient reason to do otherwise. That is why when you intend to doodle for no reason, you are disposed to abandon your intention virtually at the drop of a hat. On the other hand, we also saw that there are other cases –namely, cases of reasons-based intention—in which the intention exhibits greater counterfactual robustness. In the third scenario above, you come to base your intention to doodle on a reason for action –namely, the reason of avoiding an unpleasant encounter. But in basing your intention on this reason, you acquire a disposition to retain this attitude in the face of certain reasons to do otherwise. So, when, a moment later, your spouse requests that you both leave the coffee shop in order to run an errand, you do not respond to their request –and, more specifically, to the reason it gives you to act accordingly—by immediately dropping your intention to doodle. Instead, you stick to it, at least for a while. The point generalizes. When one endorses a reason for action as normative for one's intention and is appropriately responsive to it, one's intention acquires as a result a higher degree of counterfactual robustness, that is, a tendency to persist not only in the scenario in which it was initially formed but also across a range of counterfactual scenarios. On this view, therefore, it is the endorsement of and responsiveness to reasons, or its lack thereof, that determines the stability of a particular intention across contexts and hence that explains why reasons-based intention is considerably more robust (or stable) than is mere intention.

4. A hybrid view of commitment

The preceding distinction between mere intention and reasons-based intention as well as the related distinction between two sorts of stability of intention speak in favor of the following conjecture about the nature of commitment. Namely, an agent's commitment to act in a certain way is partly constituted both by her intention to so act and by her endorsement of certain reasons for so acting as being normative for her intention. The idea, roughly, is that a commitment to act is nothing but a reasons-based intention to so act.

This conjecture might initially be met with skepticism. Return to the third scenario above, in which you intend to doodle for a reason, namely, the reason of avoiding an unpleasant encounter. It may seem strained to say that you are committed to doodling in that scenario. For, intuitively, it seems that an agent's commitment to action involves valuing or caring about the object of her commitment in a special way, and it is unclear whether such forms of valuing or caring are present in this scenario.

I am inclined to agree that in paradigmatic cases of commitment the reasons that support the agent's intention have a source in what she deeply values or cares about. They are regarded by the agent as *strong* reasons for action. But I also think, on the other hand, that there are obvious problems with specifying what counts, from the agent's perspective, as strong reasons. In particular, it is unclear that we can specify beforehand how strong the agent must conceive of a certain reason for action in order for her endorsement of it to count as partly constitutive of her commitment to action. A better strategy, I think, is to accept that commitment to action comes in degrees and that degree of commitment is partly a function of the agent's conception of the strength (or weight or status) of such reasons. So, I propose that we amend the previous conjecture thus. We may say that one's commitment to φ is partly constituted both by one's intention to φ and by one's endorsement of certain reasons for φ-ing as normative for one's intention to φ, where such reasons are seen by oneself as strong enough to outweigh or defeat various reasons for not φ-ing. According to this proposal, then, the stronger one's reasons for φ-ing are from one's perspective, the stronger and more counterfactually robust is one's commitment to φ. Since, on this view, a commitment to action is a hybrid of two elements, an element of intention and an element of normative endorsement, I call it the "hybrid view" of commitment.

The following examples can help to bring the idea of degree of commitment home. Consider the case of two college students, Claire and Harold, each of whom intends to attend law school after graduation. Claire and Harold have different reasons for intending to do this. Claire places a great value in protecting the rights of people, especially vulnerable people. Her dream is to become a public defender. Harold, on the other hand, thinks that it would be nifty to follow a family tradition. Both his parents are prestigious attorneys and so are his grandfathers and their fathers before them. Imagine, next, that Claire and Harold have taken relevant steps toward attending law school. They have each declared the philosophy major, taken classes pertaining to the pre-law track, and attended events organized by the Pre-Law Office. Imagine, also, that they are at the end of their junior year and believe that the forthcoming summer break is their last chance to prepare for the LSAT. Finally, imagine that during this time, they each receive an invitation to spend their summer break at their favorite holiday destination. Claire prioritizes her LSAT preparation and declines the invitation. Harold accepts it, as he realizes that his priority now lies in having fun. Harold also realizes that his new priorities are incompatible with his original plan of attending law school and accordingly decides to abandon this plan. I think it is plausible to suggest that all along Claire was more committed to attending law school than was Harold. This is because Claire regarded her intention to be based on more solid grounds. Unlike Harold, the value Claire placed on attending law school was sufficiently high to withstand the eventual costs, however significant, associated with her commitment to that end.

Now contrast the cases of Claire and Harold with the case of Tom. Like them, Tom is also a pre-law student. He just completed two pre-law courses in his first semester in college and unfortunately did not do as well as expected. This puzzles him and leads him to question whether attending law school is what he really values or cares about. He now realizes that (sadly) he had never given this question any serious thought and that his plan to attend law school was a product of the influence of an unconscious desire to please his parents. Now that he comes to reflect on it, Tom realizes that his true interests and prospects of success support instead a career in sports journalism and consequently abandons his original plan. I think it is appropriate to say that although in this scenario Tom initially intended to attend law school, he was not committed to doing so.[12] Since his desire to please his parents was unconscious, he did not treat it as having any normative force. And since his intention was not based on the endorsement of reasons, it did not amount to a commitment to attending law school. Tom's attitude was a mere intention and thus displayed low stability across contexts, so much so that when the context slightly changed, he decided to give it up.

Of course, we commit not only to long-term projects or life-plans. We also commit to ordinary, day-to-day, actions. Suppose that this morning you decide to go for a 5-mile run this evening. Compare between three different scenarios. In the first scenario, you arrive at such a decision on the basis of health considerations. You judge that running this evening will contribute to your health and general fitness. In a second, you decide to do this for financial reasons. You were offered a significant monetary reward to complete the run. Finally, in a third scenario, you make that decision for moral reasons. You promised a friend to join them for a 5-mile run at the park this evening. Suppose, next, that it has been a stressful day at work and that you now begin to value the option of stopping by the local bar this evening instead. The point I want to stress here is that your response to this new reason for action is likely to differ between such scenarios. In the first, you come to judge that, in such a context, considerations of entertainment outweigh considerations of health and consequently replace your prior intention to go for a run with a new intention to stop at the bar. In the second and third scenarios, you judge, in contrast, that entertainment considerations are outweighed by, respectively, financial considerations and moral considerations, and accordingly decide to stick to your original plan. In the three scenarios, the fact that you endorsed some considerations as normative for your intention endowed this attitude with a certain degree of counterfactual robustness. In all three cases, your intention exhibits more than ceteris paribus inertia. Yet, the counterfactual robustness of your intention is greater in the second and third scenarios than it is in the first. This is because in the latter two scenarios, but not in the first, you regard your reasons for running 5 miles as being sufficiently strong to outweigh or defeat several reasons to do otherwise –where this includes your reasons for stopping at the

[12] Calhoun considers a similar example and arrives at a similar conclusion (2009, 618).

bar.[13] And this supports our intuition that in the last two scenarios you have a stronger commitment to action than you do in the first.

5. In defense of the hybrid view

Typically, the question of an agent's (psychological) commitment to action is associated, quite naturally, with the issue of the rational motivation of action. There are at least two more specific questions we may have in mind by speaking of "rational motivation" in this context. On the one hand, we may be thinking of how, or the circumstances under which, an agent comes to rationally make, sustain, and abandon a commitment to act in a certain way. This is a generalization of the question of the rational stability of commitment discussed above. On the other, we may be thinking of how, once present, an agent's commitment can lead her all the way down to action, that is, we may be thinking of the functions it plays in seeking to bring the action about.

The hybrid view, I submit, makes progress toward answering both those questions. Regarding the first, it says that the fact that commitment is partly constituted by both intention and normative endorsement helps to explain why the agent will rationally sustain her commitment across a set of counterfactual scenarios. Regarding the second, it says that the fact that commitment is thus constituted helps to account for the roles it plays in leading the agent all the way down to action. In section 4, we saw how the hybrid view articulates, on the grounds of an important distinction drawn in section 3, a basic answer to the question of rational stability. In what follows, I elaborate further on this answer. I sketch the answer the hybrid view provides to the question of how commitment tracks action in sections 6 and 7.

The hybrid view states that it is one's responsiveness to the reasons one endorses as normative for one's commitment that makes this latter phenomenon counterfactually robust. Importantly, this statement builds on the idea that the endorsement of reasons is essential to commitment. This is an idea that is accepted not only by the hybrid view but also by other views of commitment. Indeed, in developing her influential theory of the will, Christine Korsgaard offers an argument that lends support to this idea.[14] Her argument concerns the question of how one might justify to oneself the pursuit of a certain end. This is how Korsgaard frames it: "If I am to will an end, to be and to remain committed to it even in the face of desires that would distract and weaknesses that would dissuade me, it looks as if I must have something to *say to myself* about why I am doing that—something better, moreover, than the fact that this is what I wanted yesterday" (2008: 63. The emphasis is hers).

[13] This is not to deny that obligation-based reasons typically enhance the stability of the intention of a rational agent in ways in which ordinary prudential reasons –where this includes financial reasons—do not. See, for example, Wallace (2019). That said, what I am highlighting in the text is a feature that is common between those two scenarios

[14] Korsgaard's argument is in fact offered in support of a stronger thesis, namely, the thesis that commitment is to be identified with a form of normative endorsement. I think we should reject this stronger thesis. Below I explain why.

For the moment let us bracket the issue of what Korsgaard means by "willing an end" in this context and concentrate instead on the question she calls attention to in that passage, namely, the question, which may arise in the context of one's pursuing, or intending to pursue, a certain end, of "why [one is] doing that." This is a question about the (subjective) justification of intention and action.[15] It is the question of how one can reasonably convince oneself to retain an intention to act in a certain way when confronted with, for example, the influence of opposing desires.[16] Korsgaard maintains that what one must say to oneself in response to that question is (at the very least) that one has a reason for so acting (2008: 63-64). When one commits to an end, Korsgaard says, one tends to "remain committed to [that end] in the face of desires that would distract and weaknesses that would dissuade [oneself]" (2008: 63). But one will "remain committed" only if one can justify to oneself why one is committed to so act, and one will be able to do this only if one takes oneself to have a reason for so acting. Therefore, commitment provides an answer to the substantive question of justification partly because it involves normative endorsement.

Return now to the question of what Korsgaard understands by "willing an end." If by this expression Korsgaard means intention, as some critics have suggested, then my agreement with Korsgaard's points above is only partial. For, as I have said above in relation to Davidson's view, I reject the thesis that intention is or involves a judgment about the good or reasons. I take cases of intending for no reason as serious counterexamples to it. Accordingly, I also reject the related idea –perhaps, Korsgaard's own idea— that seeing oneself as being able to give a satisfactory answer to the aforementioned question of justification is a condition of possibility for intention. In this respect, I agree with R. Jay Wallace's acute observation, in the context of his critical discussion of Korsgaard's theory of the will, that the identification of intention with normative endorsement "foreclos[es] genuine possibilities in the theory of action," such as the possibility of akratic action (2001: 5, 4-10).

In that discussion, Wallace makes another observation about Korsgaard's theory, which is relevant for present purposes. He claims that an appropriate response to Korsgaard's question, "Why am I doing that?", is to point out that one has formed (and currently possesses) a mere intention to do it. To support this claim, Wallace imagines the case of someone who, "in a spiteful and nasty mood," forms the intention, or "resolves," to burn their roommate's books, "without really supposing that [this] is good or justified in any way at all" (2001: 7). Wallace points out that if, in the midst of their action, this person comes to critically reflect on what they are doing, they will "have something better to say to [them]self than that burning the books is what [they] wanted yesterday, namely that it's what [they] have resolved to do" (2001: 8).

[15] For present purposes I will concentrate more specifically on the question "Why am I intending to do that?" as one may ask this question also in circumstances in which one has not begun to act.

[16] As I see it, Korsgaard's question differs from Anscombe's question "Why?" in that the latter is a question about the explanation of action, which need not also be a question about its justification.

Despite agreeing with Wallace on the need to distinguish between intention and normative endorsement, I am not persuaded by the idea that one can satisfactorily answer Korsgaard's question in the way he specifies. For, as suggested above, the question Korsgaard calls our attention to is a question about the justification of intention and the mere fact that one has formed this attitude is compatible with the possibility of this attitude's being, from one's own deliberative perspective, arbitrary or unjustified. In a sense, saying that one intends, or has resolved, to do something provides a justification no stronger than saying that one desires to do it. For just as much as one may desire to do something for no reason, one may intend to do something for no reason. Mere intention cannot provide a satisfactory answer to the question of justification.

The preceding remarks provide support for the idea that normative endorsement is necessary for commitment. This idea is central to Korsgaard's view of commitment. But Korsgaard accepts not only the idea that normative endorsement is necessary for commitment but also the idea that it is sufficient for it. She avers, for instance, that "willing an end just is committing yourself to realizing the end," where this "*is* an essentially first-personal and normative act," namely, the act of "giv[ing] [your]self a law" (2008: 57. The emphasis is mine). As mentioned above, I think, like Korsgaard, that normative endorsement is necessary for commitment. But I also think, unlike Korsgaard, that normative endorsement is insufficient for it. For in my view, commitment is also partly constituted by intention and normative endorsement need not secure this. One may endorse a consideration as a decisive or sufficient reason for action in a certain context and fail to be responsive to it in that context; one may fail to form the corresponding intention. Here I agree with a point stressed in work by Allan Gibbard (1990) and Peter Railton (2006), among others. Accepting a norm or endorsing a reason involves a disposition to conformity. But this is only a "mitigated" or defeasible disposition, and thus it can fail to manifest in the particular case.[17]

6. On the functional roles of commitment

The explanation of how commitment leads to action owes, according to the hybrid view, to the fact that it is partly constituted by intention and by normative endorsement. Each of those elements plays a relevant motivational role in leading the agent all the way down to action.

The functions that one's commitment to an end plays in one's practical reasoning and action in pursuit of that end are, in part, the functions of its partly constitutive intention. I say "in part" because there is also an important difference. We noted above that, unlike mere intention, commitment involves two elements, intention and normative endorsement, and that it is the influence of the latter element on the former that endows commitment with counterfactual robustness. This points to a difference in functional

[17] Talk of "mitigated disposition" in this context comes from Gibbard (1990).

roles between those attitudes. Since, unlike intention, commitment tends to be stable not only within a context but also across several contexts, it also has a tendency to play the aforementioned functional roles across a wider range of circumstances or counterfactual scenarios. Rational commitment plays those functions both in circumstances in which rational intention plays them and in circumstances in which the latter may not; and this is simply because, if rational, one may retain a commitment in circumstances in which one may not retain a mere intention. To put it differently, the fact that commitment is partly constituted by intention and normative endorsement produces a sort of motivational synergy that is absent in the case of mere intention: the stronger one deems one's intention for an end to be supported by reasons, the wider the set of circumstances in which one is willing to pursue that end.

To better appreciate the difference in functional roles between commitment and intention, let us delve deeper into one of the functions of this latter attitude, namely, the role it plays in tracking action. Intention plays this action-tracking role in virtue of involving some relevant dispositions. First, as mentioned in section 2, intention involves a fairly strong disposition to move one towards action. Bratman offers a precise characterization of this disposition: normally, he says, if one intends to act in the future, persists in one's intention until the time of action, and nothing interferes, one will have so acted –or at least one will have tried to so act (Bratman 1987; cf. Grice 1971). What is important for present purposes is to call attention to a key feature of the relation between intention and action, which is correctly captured in Bratman's characterization of that disposition, namely, that the success of intention in tracking action is partly dependent on the *persistence* of this attitude throughout that process. Once we acknowledge this, we see why commitment involves an even stronger disposition to action. Since the intention partly constitutive of commitment tends to persist across contexts, commitment tends to move one toward action across a wider range of circumstances than does mere intention.

Second, the disposition to action characteristic of intention involves a disposition to make several attempts, if necessary, in order to secure success –where this includes, in turn, a set of dispositions to, among other things, search for and intend to pursue alternative (sufficient) means to the intended action. This disposition of intention is not a disposition to try just once, regardless of success, or to try halfheartedly; it is instead a disposition to "seriously try" (Cf. Hampshire 1959: 112-13), where this includes, I take it, a disposition to make several attempts if necessary. Now, again, in the case of a mere intention to act, this disposition will typically manifest when other things are equal, that is, only when one registers no relevant changes in the circumstances. For such are the circumstances under which one will rationally retain such an intention. Things are different with commitment. If one has a commitment to act, one will thereby possess an intention to so act that tends to persist not only within the context in which it was formed but also across contexts, and thus one will be rationally disposed to make several attempts not

only in contexts in which one does not acquire a new reason to do otherwise but also in contexts in which one does acquire some of those reasons.

7. Intention, commitment, and coordination

As human beings, we have a need for coordination. We need to coordinate our activities within ourselves through time and also coordinate our activities with those of others. Several philosophers –and, again, most notably, Bratman—have argued that intention is particularly well suited to facilitate such forms of coordination, given the various roles this attitude plays in one's practical reasoning and action.[18] For example, since it involves a tendency to bring about action, intention typically grounds expectations of future action, and such expectations are key to familiar forms of coordination. It is partly because I intend to visit NYC this weekend that I can reasonably expect that I will be there then. And, it is partly because of this expectation, that I can now plan to attend the Picasso exhibition at MoMA on Saturday afternoon.[19] Similarly, it is partly because you know that I intend this and that this grounds a corresponding expectation that you can now plan to meet me at MoMA on Saturday afternoon.

I think it is undeniable that intention plays an *important* role in facilitating coordination, especially when compared with ordinary desire. For, normally, a mere desire on my part to visit NYC this weekend supports neither the assumption that I will be there nor making plans on the basis of that assumption. Furthermore, intention seems *necessary* to secure such forms of coordination. For it seems that in order to now ground the assumption that I will visit NYC this weekend, I must now have settled in my mind the question of whether to visit NYC then.

What I find unpersuasive, however, is the argument that goes from the claim that we have a need for coordination to the conclusion that intention is (conatively) sufficient to ensure it. Intention can facilitate coordination only if it tends to bring about action. But intention exhibits this tendency, we have said, only insofar as it exhibits a tendency to persist until the time of action. We know that one's intention to act exhibits, as such, a tendency to persist so long as one registers no relevant changes in the circumstances. That is, it exhibits what we called "ceteris paribus inertia." Therefore, we know that so long as one registers none of such changes, one's intention will tend to be sufficiently stable to facilitate coordination. The problem, however, is that it is quite common for those changes to occur and for one to register (at least some of) them.

As mentioned earlier, we live in a changing world, a world in which we ordinarily discover new opportunities for action. So, when we think about the challenge of coordination, we should remember that

[18] The main reference here is Bratman (1987). See also, for example, Grice (1971), Harman (1976), and Velleman (2007).

[19] Bratman (1987: 18). Cf. Grice (1971: 270) and Harman (1976: 446-48). For further discussion of forms of planning on a prior intention, see Alonso (2020).

we are thinking about the challenge of coordinating our activities in such a world, a world that is likely to bring relevant changes. My sense is that coordination in such a world requires an attitude *more robust* than mere intention. It requires an attitude that is sufficiently stable across contexts, rather than simply within a context, and therefore an attitude that can facilitate coordination across various contexts, rather than only within a context. It requires, in other words, an attitude that tends to persist even in the face of some new reasons to do otherwise. This attitude, I submit, is commitment. In the changing world we live in, it is commitment to action, rather than a mere intention to so act, that *normally* provides a sufficiently robust fixed point for coordination.

This point is strengthened when we connect it to our previous example. Consider how my intention to attend the Picasso exhibition can facilitate interpersonal coordination. Contrary to what the standard story of the coordinating role of intention indicates, I suggest that the mere fact that I intend to attend the exhibition on Saturday may provide little support for the belief that I will perform that action then and thus little support for coordination. To see why, consider the following variation on this case. Imagine that I formed the intention to attend the exhibition for no reason –perhaps I acquired it through a nondeliberative process. Imagine, also, that you know this. This supports neither assuming that I will in fact attend the exhibition nor planning to meet me at the museum on the basis of that assumption. For you know that I only possess a *mere* intention to attend the exhibition and might drop it as soon as I come to reflect on the matter. What is more, you know that it is likely that I would come to reflect on the matter in such a context, as that is a context in which I am prone to encounter new opportunities for action and acquire new reasons to do otherwise. In the paradigmatic or typical case, my intention to attend the Picasso exhibit will be based on strong reasons –or so it seems plausible to say. It will be an intention I am disposed to retain in the face of some new opportunities for action. It will be a commitment to attend the exhibit. And this commitment will typically be robust enough to facilitate coordination.

Of course, the limitations that mere intention displays in facilitating interpersonal coordination are more salient in high stakes cases in which the agents have no initial reason to cooperate with the other. So, consider a version of Hume's example of the two farmers (1979). Imagine that I need your cooperation in order to harvest my crops in a timely manner and that you need my cooperation in order to harvest yours. Imagine, also, that each of us is moved solely by considerations of self-interest. In such a context, none of us seems to have a reason to cooperate with the other. I know that if I help you to collect your crops today, you will have no reason to help me collect my crops tomorrow. For, by then, you would have already reaped the benefits of cooperation, and helping me in return would only bring about a cost to you. But knowing this, I have no reason to help you today. For helping you today would only bring about some costs and no benefit to me in the future. Clearly, cooperation is difficult to come by in such a context. All this is familiar. However, the point to be made here is that this coordination problem would not likely be resolved were

each of us to unconsciously acquire an intention to help the other and were this transparent between us. For each of us would know that if the other only possessed a *mere* intention to help, she would drop it as soon as she comes to reflect on the matter. Typically, in order to be willing to cooperate with you in such cases, I will need to know more than the fact that you have a mere intention to do your part; I will also need to know that you take yourself to have strong reasons for doing so. I will need to know that you are *committed* to doing your part, and vice versa.[20]

The previous observations should not lead us to conclude, however, that mere intention is never sufficiently robust to facilitate coordination. As suggested above, one's mere intention to act can facilitate coordination when one registers no relevant changes in the circumstances. If one does not come to see oneself as having stronger reason to do otherwise, one will be disposed to retain one's intention and also be disposed to engage in further practical reasoning and action conducive to such coordination.

This prompts a final observation. It seems clear to me that when we (as theoreticians) describe the coordinating roles of intention, we are rarely describing the coordinating roles of intention as such. For the cases we generally appeal to in order to illustrate the coordinating role of intention are not, in fact, cases of mere intention. Think, again, of the coordination afforded by forms of prior-success planning, as illustrated in the first example above. This involves forming an intention to φ (here: attend the MoMA exhibition) in the further future on the assumption that one's prior intention to ψ (here: visit NYC) in the nearer future will be successfully executed. When we characterize this form of planning, we rarely consider cases in which the prior intention is based on no reason. This raises a question as to whether the standard story of the coordinating role of intention does not often mistakenly attribute to this attitude, and solely to it, a role in facilitating coordination that is in fact played not only by this attitude but also by the endorsement of reasons. My impression is that it is usually the more complex attitude of commitment, rather than mere intention, that is at work in ordinary cases of coordination.

8. Concluding remarks

In this chapter I have proposed a hybrid view according to which an agent's commitment to action is partly constituted by a structure of attitudes that includes her intention to act and her endorsement of certain reasons for action as being normative for her intention. I have argued that the fact that commitment is so constituted helps to account for important features of this attitude and also for some differences with intention. It helps to explain why, unlike intention, commitment tends to persist across a wider range of circumstances and, consequently, why it also tends to track action across such circumstances. This in turn

[20] Cf. Kolodny and Wallace (2003).

helps to explain, I have suggested, why commitment is a more robust fixed point for facilitating coordination than is mere intention.

References

Alonso, Facundo. (2017). "Intending, Settling, and Relying," in *Oxford Studies in Agency and Responsibility*, David Shoemaker (ed.), Volume 4, pp. 50-74. (Oxford: Oxford University Press).

———. (2020). "Planning on a Prior Intention." *Journal of Ethics and Social Philosophy* 18: 229-265.

Anscombe, Elizabeth. (1963). *Intention*. (Ithaca: Cornell University Press).

Arruda, Caroline. (2022). "Sticking to It and Settling: Commitments, Normativity, and the Future," in *Time in Action*, Carla Bagnoli (ed.), pp. 149-72 (London: Routledge).

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

———. (1999). *Faces of Intention*. (Cambridge, MA: Cambridge University Press, 1999).

———. (2007). *Structures of Agency*. (Oxford: Oxford University Press).

———. (2018). *Planning, Time, and Self-Governance: Essays in Practical Rationality*. (Oxford: Oxford University Press).

Calhoun, Cheshire. (2009). What Good is Commitment? *Ethics* 119: 613-41.

Davidson, Donald. (1963). "Actions, Reasons and Causes," reprinted in his *Essays on Actions and Events*, pp. 3-19. (Oxford: Oxford University Press, 1980).

———. (1978). "Intending," reprinted in his *Essays on Actions and Events*, pp. 83-102.

Gibbard, Allan. 1990. *Wise Choices, Apt Feelings*. (Cambridge, MA: Harvard University Press).

Gilbert, Margaret. (2014). *Joint Commitment: How We Make the Social World*. (Oxford: Oxford University Press).

Grice, Paul. (1971). "Intention and Uncertainty". *Proceedings of the British Academy* 57: 263–79.

Harman, Gilbert. (1976). "Practical Reasoning". *Review of Metaphysics* 29: 431–63.

Hume, David. (1979). *A Treatise of Human Nature,* L. A. Selby-Bigge and P. Nidditch (ed.) 2nd edition. (Oxford: Oxford University Press).

Kolodny, Niko & Wallace, R. Jay (2003). Promises and Practices Revisited. *Philosophy and Public Affairs* 31 (2):119-154.

Korsgaard, Christine. (2008). "The Normativity of Instrumental Reason," reprinted in *The Constitution of Agency*, pp. 17-68 (Oxford: Oxford University Press).

Railton, Peter. 2006. "Normative Guidance," in *Oxford Studies in Metaethics*, Russ Shafer-Landau (ed.), Volume 1, pp. 3-33. (Oxford: Oxford University Press).

Shpall, Sam. (2014). “Moral and Rational Commitment,” *Philosophy and Phenomenological Research* 8: 146-172.

Velleman, J. David. (1992). “The Guise of the Good,” *Noûs* 26: 3-26.

———. (2007). “What Good is a Will?” in *Action in Context*, pp. 193–215. A. Leist. (ed.). (Berlin: Walter de Gruyter).

Wallace, R. Jay (2019). *The Moral Nexus*. (Princeton, NJ.: Princeton University Press).

Wiseman, Rachael. (2016). *Routledge Philosophy Guidebook to Anscombe’s Intention*. (New York: Routledge).