

# Intention as Belief\*

John Schwenkler  
Florida State University

(I)

What's the relationship between:

- (i) *Intending* to do something,
- (ii) *Believing* that you are going to do this, and
- (iii) Its being the case *that you are going* to do the thing in question?

A natural answer to give is that while these three categories often coincide, the reality of each doesn't depend on either of the others. A person might, for example, intend to stop by the bookstore on the way home from work but not believe that she's going to do it (see Bratman 1999, p. 37): this would be an example of (i) without (ii); and if we imagine instead that the person *believes she is going* to skip the bookstore, despite not intending to do so, then we'd have a case of (ii) without (i) instead. A person can intend to do, or believe that she's going to do, something that's actually quite impossible for her, like square a circle or travel faster than the speed of light: this would give us (i) or (ii) without (iii). And, finally, things might be such that a person *is going* to do a certain thing—step on a snake, say, or have another cigarette—but without any idea of this, or without any resolve that it happen: the former case gives us (iii) without (ii), and the latter gives us (iii) without (i). The prospects for aligning our categories do not appear very promising.

Despite these appearances, several philosophers have recently developed accounts of future-directed intention according to which some of these categories coincide after all. For example, in a recent paper Berislav Marušić and I made a case for *strong cognitivism* about intention, according to which 'To intend to do something is neither more nor less than to believe, on the basis of one's practical reasoning, that one will do it' (Marušić and Schwenkler 2018, p. 309). And Philip Clark (2020) has recently proposed a view on which intention is identified with the 'conative state' that he calls 'being going to do' the thing that one intends (Clark 2020, p. 316). Clark's position identifies (i) above

---

\* I presented a version of this paper at Baylor University and learned a lot from members of the audience there, especially Charity Anderson, Yoav Isaac, Anne Jeffrey, Alex Pruss, and Christopher Tomaszewski. Thanks also to Phil Clark, Eric Marcus, Beri Marušić, and Marshall Bierson for discussion of these matters, and to a referee for this journal, who I later learned was Phil Clark, for an exceptionally helpful report that led to significant improvements.

with (iii) and denies that they are identical with (ii), while Marušić and I identified (i) with (ii) without taking any direct stance on the relationship of these to (iii). In this paper, I propose a position on which all three categories, correctly understood, amount in the fundamental case to the very same thing.

Here is how I will proceed. Section II summarizes the case for strong cognitivism and presents Clark's main objections to it. Sections III and IV evaluate Clark's case for noninferential weak cognitivism, first offering a friendly amendment to it and then raising an important objection. Section V then defends the cognitivist identification of intention with belief while also arguing that, in the conceptually central case, the belief that you are going to do something, when grounded in your practical reasoning, makes the case that you are going to do the thing in question.

## (II)

In our paper, Marušić and I offered three main arguments in defense of the identification of intention with belief.<sup>1</sup>

First, we argued that this position best explains why beliefs about one's future actions should be *transparent* to practical reasoning. If asked whether I'll speak at a conference next February, I will address this question through reasoning that considers the pros and cons of doing so, rather than reasoning that considers evidence for and against the judgment that I'll do this. And if I reason, 'The conference will be in a nice location and on an interesting topic; my schedule is free and the trip won't inconvenience my family; so I'll go', then my belief *that I'll go* seems to be arrived at through the same reasoning by which I arrive at the intention to do this.<sup>2</sup> This phenomenon is easy to account for on a view according to which belief and intention are one and the same.

Second, Marušić and I observed that the usual way to *express* one's future-directed intentions is through a statement that describes *what one will do*, or *what one is going to do*. For example, if I tell you that I'm going to be there for your conference in February, then usually I'll thereby have expressed the intention to do this. (Only 'usually': because I *might* anticipate being brought there by force.) This statement is no less a way to express the *belief* that I am going to be at the conference. And an easy way to explain why the very same statement can express both the intention to do something and the belief that one will do it is by appeal to the hypothesis that intention is identical to belief.

Finally, Marušić and I argued that the cognitivist account of intention is supported by the role that intentions play in planning behavior. If I intend to go to the conference in

---

<sup>1</sup> For these arguments see Marušić and Schwenkler (2018), pp. 316-321.

<sup>2</sup> For more on how to understand this form of reasoning, see Schwenkler (forthcoming).

February then I'll make the necessary travel arrangements, book a hotel room, keep those days otherwise free on my calendar, and so on. All this behavior makes sense only on the assumption that I *believe* I am going to be there: for otherwise, why would I book a hotel room and make hotel reservations in that city rather than somewhere else? And the thesis of strong cognitivism makes good sense of why intention would play this belief-like role in planning.

Two points of clarification are in order. First, on Marušić and my view intentions are different from other beliefs in being distinctively 'practical': they are part of an agent's view of her future insofar as she means to *create* the truth that she envisions, which is why they do not have to *reflect* evidence about what is anyway going to happen. This aspect of our position is supposed to account for the possibility of believing you will do something without thereby intending to do it. If, for example, my belief that I am going to forget to stop by the bookstore is what Anscombe (1963, p. 2) calls an 'estimate of my chances', i.e., a belief based on evidence about my usual tendency to forget things, then it is not a *practical* belief and so not, on our view, an intention. That is why it is possible to hold this belief without thereby intending not to stop. Second, we also hold that intentions, like many of our other beliefs, often have an implicitly *conditional* character that's not always expressed in our stated descriptions of what we are going to do. For example, the intention I'd express by saying 'I'm going to stop by the bookstore this evening' is not an intention to do this *no matter what*: for example, I surely don't intend to do this even if I need to rush home because my house is on fire. And we appeal to this idea to explain why it's possible to intend to do something despite being unsure of whether you actually will actually do it. I might, for example, intend to speak at the conference that you've invited me to, but not be able to rule out the possibility that I'll have to stay home to care for a sick child or skip the trip because the airlines have been shut down due to a global pandemic. (A rare case where life keeps up with the philosophical imagination.) In this case, the belief with which Marušić and I would identify my intention to speak at the conference is the belief that I'll do this *if* the airlines are operating and *unless* I need to care for a sick child. And it's plausible that my intention to speak has this same conditional content.

Significantly, all the cognitivist arguments summarized above have the form of an inference to the best explanation. This means that none of them even purports to show that intention *must* be identical to belief: Marušić and I argued instead that in each case 'the best—because most elegant, intuitive, and unforced—explanation of the phenomena' in question is that intention is identical to belief, and that our cognitivist position 'should be favored for that reason' (Marušić and Schwenkler 2018, p. 317). This argumentative strategy means, however, that making a successful case against strong cognitivism doesn't require refuting any of these three arguments. Each could be perfectly good as it stands, and we would still have sufficient reason to adopt a different explanation of these

phenomena if we were shown on some other grounds that strong cognitivism should be rejected. That is what Clark aims to do in his paper.

Clark (2020) argues in two ways that intentions are not identical to beliefs about what one will do. The first argument is that intentions ‘are not shown to have been true or false by what you go on to do’:

If you prefer sleeping in the parlor to sleeping in the attic, and wind up sleeping in the attic, your preference for sleeping in the parlor is not thereby shown to have been false. Similarly, if you intend to sleep in the parlor, and wind up sleeping in the attic, your intention is not thereby shown to have been false. By contrast, if you think you will sleep in the parlor, and wind up sleeping in the attic, your thought is thereby shown to have been false. On the face of it, intentions to  $\phi$  fall in with desires to  $\phi$  and preferences for  $\phi$ -ing, rather than with thoughts and beliefs that you will  $\phi$ . (Clark 2020, p. 309)

I find this argument less than convincing. While it’s true that statements like ‘Your intention is false’ or ‘So-and-so falsely intended ...’ aren’t part of ordinary English parlance, the cognitivist can give a pragmatic explanation of this by appeal to the fact that intentions, and the statements that express them, aren’t supposed to be grounded in evidence for their truth. As Anscombe (1963, p. 55) observes, if I say, for example, that ‘(I think) I am going to go to bed at midnight’, the way for you to contradict my statement is not by saying ‘You won’t, for you never keep such resolutions’, but rather by saying something like ‘You won’t, for I am going to stop you’.<sup>3</sup> On the face of it, saying this last thing is a way of saying that my belief that I was going to go to bed at midnight was false, but not by providing *evidence* against it—which is what explicit talk of truth and falsity would seem to make the issue.

Another problem with Clark’s argument is that even if though sounds strange to call intentions themselves true or false, the same isn’t true of the agents who intend things or the statements by which intentions are expressed. If I decide (that I’m going) to do something which then I don’t do, then it’s quite natural to say that *I was mistaken* in thinking that I’d do this, and if I had told you that I’d do it then it would be appropriate to complain that *what I said* was false.<sup>4</sup> Yet *what I said* is nothing more than *what I intended*—viz., to do what I said I’d do. As Anscombe puts it in a related context, why

---

<sup>3</sup> For interpretation of this difficult passage, see Schwenkler (2019), pp. 106-113.

<sup>4</sup> Indeed, consider Clark’s own definition of belief: ‘To believe that  $p$  is to represent  $p$  as true, and to do so in such a way that if it is not the case that  $p$ , one is thereby shown to be in error about whether  $p$ ’ (2020, p. 311). Here it is the believer, and not the belief itself, that is described as the locus of error or correctness.

should we think there is more than ‘a dispensable usage’ reflected in the fact that we don’t *call* intentions true or false according as they are executed or not?<sup>5</sup>

Clark’s other objection to strong cognitivism is much more pressing, however. The idea behind it is one that I alluded to earlier, namely that it seems possible to *believe* you are going to do something that you want to do, or think you ought to do, without actually *intending* to do the thing in question. Clark finds a version of this idea in Harry Frankfurt’s writings:

a person may attempt to resolve his ambivalence by deciding to adhere unequivocally to one of his alternatives rather than the other; and he may believe that in thus making up his mind he has eliminated the division of his will and become wholehearted. Whether such changes have actually occurred, however, is another matter. When the chips are down he may discover that he is not, after all, decisively moved by the preference or motive he supposed he had adopted (Frankfurt 1998, p. 101; cited in Clark 2020, p. 312).

Clark takes Frankfurt’s remark to support the anti-cognitivist conclusion that believing that one will do something, even where this belief is of the distinctively practical variety, is *insufficient* for intending to do it. The state of intention requires that a person’s will is, if not entirely undivided, at least decisively on the side of doing the thing in question. And it seems to be possible to *believe* that you are so decided that you are going to do the thing you have decided on, when in fact this just isn’t true: you aren’t going to do it, and looking back we will be able to say that you *never were* going to do it at all. Intention requires something more than a mere belief about what one is going to do.

I’m going to put my reply to this objection on hold until the final section of this paper. For now I want to emphasize one thing that the objection doesn’t depend on, since this will bring out a distinction that’s going to be all-important in what follows. It’s central to Clark’s position, and to any adequate account of future-directed intention, that the question of what a person intends not be conflated with the question of what they actually go on to do. Even Anscombe, who holds that the expression and description of intention *are* ways of saying what is happening or is going to happen, also insists that it’s possible to intend to do what one never actually does:

a man can form an intention which he then does nothing to carry out, either because he is prevented or because he changes his mind: but the intention itself can be complete, though it remains a purely interior thing. (Anscombe 1963, p. 9)

---

<sup>5</sup> Here is Anscombe’s remark in its full context: ‘Execution-conditions for commands correspond to truth-conditions for propositions. What are the reasons other than a dispensable usage for not calling commands true or false according as they are obeyed or disobeyed?’ (1963, p. 3).

It is this phenomenon, that of ‘intention for the future’, that is the subject of the present discussion, and while all parties to the conversation agree with Anscombe that this topic isn’t separable from that of ‘what physically takes place, i.e. what a man actually does’ (Anscombe 1963, p. 9) in carrying out his intentions, we all agree on this as well: that a person can intend, in a full-blown and totally unqualified way, to do what he does not ever end up doing. That is, we all agree that a person may *have been*, in Frankfurt’s phrase, ‘decisively moved’ to do something, even if he never executed this decision for one reason or another. Below I’ll explore Clark’s attempt to account for this possibility, and then I’ll return to consider whether this second objection to strong cognitivism is truly fatal.

### (III)

Clark’s alternative to strong cognitivism centers on the idea that to intend to do something requires that one *be going* to do it, in a sense of the italicized phrase that allows that a person can *have been going* to do something that she actually didn’t do at all. Examples abound: ‘I was going to buy a Ford, but then I saw that the Chevrolet dealership was having a sale’; ‘I was going to pick up some milk at the store, but then I forgot’; ‘I was going to travel to the beach, but it turned out that the bridge to get there was closed for repairs’. Nor is this phenomenon limited to intentional action. One might say, for example, that a branch *was going* to fall on her house (but then it was taken down), that a car’s brakes *were going* to fail (if she hadn’t gone into the shop in time), that it *was going* to rain (and then the wind suddenly changed direction). In all these cases, we describe a past tendency that turned out not to be realized. (The way I’m using ‘tendency’ here and in what follows is supposed to express the very same concept as Clark’s ‘being going to’.<sup>6</sup>) And Clark’s proposal is that intention consists partly in just such a tendency. On his view, to *intend* to do something requires that one *be going* to do it—and the reality of such a tendency isn’t guaranteed by the presence of a belief about what one is going to do.

---

<sup>6</sup> This use of ‘tendency’ is admittedly somewhat technical. Ordinarily, when I say, e.g., that I have a tendency to forget things (as in my discussion above of Anscombe on making an ‘estimate of my chances’), I am attributing to myself a *general* or *standing* tendency to do this, rather than making a claim about what I *right now* am going to do. Indeed, in this ordinary sense of ‘tendency’ I can (right now) have a tendency to do something that I am (right now) *not* going to do at all: for example, I might usually forget my keys when I leave for the office, though I’ve really made quite sure that I won’t do so on this particular occasion. Likewise, I can (right now) be going to do something that I don’t have any general tendency to do: for example, perhaps I’m going to make that once-in-a-lifetime visit to Paris in June. (If only.) Still, I’m going to use ‘tendency’ in this somewhat artificial way just because I find it to provide a less awkward way to describe my topic than with Clark’s phrase ‘being going to’. (I thank Phil Clark for helpful comments that prompted this footnote.)

Of course, the *mere* fact that someone is (or was) going to do a certain thing doesn't mean that she has (or had) formed an intention to do it. Something more than a sheer tendency is required, and Clark argues that the further ingredient is the practical or 'directive' belief that Marušić and I identified with intention:

Suppose, for example, that you are Maggie Fitzgerald, the fictional prizefighter, heading into the ring. In previous bouts with this opponent, you've never lasted a full three rounds. Given your track record, you judge that you will go down in the third round. Alternatively, imagine that a gambler has offered you money to lose in the third round. Given the advantages of doing so, you decide that you will go down in the third round. (Clark 2020, pp. 318-319)

Clark wants us to suppose that in both of these cases, Maggie is indeed *going to go down* in the third round. And his proposal is that in the case where Maggie *intends* to go down in the third round she must also *believe* that she is going to this, and this belief must be a *directive* attitude that is grounded in reasoning that Maggie takes to show doing this to be somehow worthwhile. Clark uses the example of being a planet to illustrate how he understands the relationship between a person's intentions and her directive beliefs:

It follows from the fact that something is a planet that it orbits a star. But the planet itself is just the object that orbits the star. Its being a planet requires the presence of the star, but the star and the planet are distinct. Indeed, the star is not even part of the planet. Analogously, while intending to do something does entail thinking you will do it, the intention itself is not a thought about what you will do. It is not identical to the thought, nor does it have the thought as a proper part. (Clark 2020, p. 309)

Clark's idea, then, is that what I have called one's 'tendencies', or what Clark calls one's 'being going to do' certain things, are to directive beliefs about what one will do as planets are to stars, and that we should identify intentions with these tendencies rather than with directive beliefs. This position avoids the difficulties that I discussed in the preceding section while also accounting for the appearances that Marušić and I used to motivate strong cognitivism: it explains the close connection between a person's intentions and her view of her own future while allowing for the possibility of self-deception and eschewing the admittedly unnatural construal of intentions as propositional attitudes that can be true or false.

But it is tricky to get the details of this position just right. Suppose, for example, that Maggie the prizefighter has formed a directive belief that she'll go down in the third round, but in fact she isn't 'decisively moved' by this conviction: she's scared of throwing the fight, and when the third round comes she won't have the nerve to take a

fall. And now suppose further that Maggie *is* going to go down in the third round nonetheless, since she simply isn't a good enough fighter to stay on her feet. In this case Maggie seems to satisfy the two conditions on intending that are discussed just above: she is going to go down in the third round, and she has a directive belief that she's going to do this. Nevertheless, at least if the Frankfurtian objection discussed in the preceding section holds up, Maggie doesn't really intend to do this. Something more is needed.

The solution seems obvious. What's necessary for Maggie to count as intending to go down in the third round is for the reasons behind her directive belief to be the reasons that ground her tendency—i.e., that make it the case that she is going to do this. In the case we just imagined, Maggie is going to go down, not because she's been offered money to lose, but because she just isn't a good enough fighter. By contrast, in Clark's original case the reasons why Maggie is going to go down are the same as the reasons why she has decided to do this.

Yet even that is not quite sufficient. Imagine, for example, that thinking about why she's accepted the bribe is going to make Maggie very nervous, which means she will fight poorly and go down in the third round, even though this isn't something that she (*really*) intends to do. (A familiar-sounding example of a 'deviant causal chain'.) To get around this, we must suppose that Maggie's own thought about what she is going to do and why is *itself* the ground of her tendency to do this, and that this fact is no accident. (By analogy: being a planet requires not just the existence of a nearby star, but also that the orbiting relation hold between the two things.<sup>7</sup>) Maggie's intention, which according to Clark is identical to her tendency to go down, must be grounded in the practical reasoning that is also the ground of her directive belief that she'll do this.

#### (IV)

We are considering the possibility that intention is not merely a practical or directive belief, but rather what I will call a *self-conscious practical tendency* toward doing a certain thing. ('Self-conscious' because of its intimate relation to directive belief, 'practical' because it is grounded in practical reasoning.) According to this position, a person doesn't intend to do something unless she really is going to do it. However, the *mere* fact that someone is going to do something, even if she has a directive, non-observational belief that she will do this, doesn't suffice for intention either. What distinguishes an intention from other sorts of tendency is that it is a tendency grounded in the reasoning by which a person decides what she will do.

What reasons are there to favor this position over the cognitivist one, or vice versa? We have already considered one argument in favor of Clark's position: that it

---

<sup>7</sup> I thank Phil Clark for pointing out the relevance of this analogy to the point I am making here.



seems possible for a person's beliefs and her intentions to come apart, since you might believe that you are going to do something without really having committed yourself to doing it. But there is a different sort of case that cuts in the opposite direction. (I thank Marshall Bierson for showing this to me.) Suppose you've decided to go to the beach, and the only way there is by driving over a bridge that, unbeknownst to you, has long been closed for repairs. In this case there seems to be no question of whether you intend to go to the beach: your will isn't in any way conflicted, and you've made up your mind that this is what you are going to do. But are you *going to go* there? It is hard to see how this could be. After all, the way to the beach is closed, and so not only *won't* you go there (at least this time), but there will also be a sense in which you *never were* going to go, since doing this was never a possibility for you. Such a case suggests that Clark's demand that intention involve a real tendency is simply too strong.

I can think of two ways to respond to this objection. One is to bite the bullet and say that in the case I've just described, you really are going to go to the beach despite the fact that the bridge that you need to get there was closed.<sup>8</sup> There is something to this: after all, in a case like this you can certainly say, after the fact, 'I was going to go to the beach—except that the bridge turned out to be closed', and it seems appropriate enough for me to say the same about you. But this reply only works in the backward-looking case. If you've made up your mind to go to the beach but the bridge that you need to get there is closed, then while *you* will say, in your ignorance, that you are going to go to the beach, *I* cannot say the same of you if I know what I do about the bridge. (Nor is this only a matter of conversational implicature. For I also cannot say, 'NN is going to go to the beach—though by this I don't mean that she *will* go there, since in fact the only bridge from the mainland is closed.') In this case, the most that I can say is that you *think* you are going to do this, or that this is what you *intend* to do. Knowing that you have no way to get to the beach, I cannot seriously say that you are going to go there.

---

<sup>8</sup> I think that this is how Clark himself would respond. He writes in an endnote that it is important to 'distinguish the relevant notion of being going to from the idea that some event is imminent, or in the offing', and illustrates this with a twist on the case of Anna Karenina in which 'she was going to continue her affair with V' but in fact V has been killed in a duel. According to Clark, 'It remains true, in this case, that Anna was going to continue the affair, even though no such event was imminent. This was the direction in which she was headed ... The sense in which Anna was going to continue the affair, in this case, is not that this was imminent' (Clark 2020, p. 325 n.7). My response to this is twofold. First, as I say in the main text to which the present note is appended, this reply works only in the backward-looking case: if I am one of Anna's contemporaries and I know that V has been killed in a duel, I will not say of Anna that she *is* going to continue her affair with V. Second, the appeal to events that are going to happen without being 'imminent' doesn't help Clark's case either, since ordinary talk of 'imminence' just has to do with what is going to happen *soon*. (Saying that my death is not imminent doesn't call into question whether I *will die*, and saying that something—a terrorist attack, for example—*was* going to happen 'imminently' doesn't entail anything about whether it *did*; indeed, the usual implication is to the contrary.) The question remains whether there is any good sense of 'to be going to' in which a person can be said to be going to do something which we know that she is not going to do.

Another response, which appears at first to be more promising, turns on an idea that Marušić and I developed in our paper and that I touched on briefly in Section II, namely that the content of our intentions is usually *conditional* in certain respects that are often left implicit when we express them. Building on earlier work by Luca Ferrero (2009), Marušić and I (2018, pp. 322-328) identified two ways in which conditions can enter into the content of an intention: a ‘precondition’ describes a condition that must be satisfied in order for one to be *able* to act in a given way, and a ‘restrictive condition’ describes a condition under which one would *decide* to do something else. Thus, for example, the intention that you might express by saying ‘I am going to go to the beach’ will usually have as a precondition that there is a way for you to drive there, and as a restrictive condition that you’re not in the midst of a hurricane. Applied to the present objection, then, the idea would be that this conditionality also exists at the level of our practical tendencies. For example, in the case where the bridge to the beach happens to be closed, though it won’t be true that you are going to go to the beach *simpliciter*, perhaps you are nevertheless going to: *go if the bridge is open*. This conditional description of a practical tendency can remain in place even if one of the enabling conditions for the action in question is not satisfied.

I said that this response appears more promising than the first one. But I don’t believe it succeeds, for two reasons. First, in the case just described if I know that the bridge won’t be open on the day you are planning on going to the beach, then I still cannot really say of you, ‘NN is going to go to the beach if the bridge is open’, but only that this is what you *would* be going to do if it *were*. And that is something I could also say in a very different sort of case, in which you had learned about the closure and decided against the trip. However, in the case at issue your intention isn’t hypothetical or counterfactual—its *content* is, but not the fact of your having it.

This last observation reveals an important difference between the way we use the phrase ‘going to  $\phi$  if ...’ to express or attribute a conditional intention, and the way we use this same phrase to attribute what we might call ‘natural’ tendencies, like the one we describe in saying that a tree is going to fall over. If, for example, you say of a storm-damaged tree that ‘It is going to fall on the house if we have another very bad storm’, then the condition you describe is *external* to the tendency in question: you are saying that *if* a certain thing happens, *then* the tree is going to fall. That is, in these cases the phrase ‘going to  $\phi$ ’ is included within the scope of the conditional, and so the sentence describes something that *might be* going to happen, but isn’t going to happen as of now.<sup>9</sup> The conditional character of an intention is different from this, however, in that it is *internal* to the tendency itself. Thus, for example, if you’ve made up your mind to go to the beach on the condition that the weather is sunny, then in saying ‘I am going to go to

---

<sup>9</sup> This kind of description can also find a place in the description of human action, as when you say of a student heading to college that she is going to fail her classes if she spends four nights a week out drinking.

the beach if the weather is sunny' you describe something that ('as of now') you *are going* to do, namely: go to the beach *if* there is sunny weather. So the phrase 'going to  $\phi$ ' ranges over the conditional rather than the other way around. The conditionality of conditional intentions is different in this way from that of ordinary tendencies.

Above I claimed that I can't really say of you, in the case where your intention to go to the beach is conditional on the bridge being open and I know that in fact the bridge is closed, that you *are* going to go to the beach if the bridge is open, but only that you *would be* going to do this if it *were*. (This makes it different from the case where you are going to go if it's sunny, and it's so far undetermined whether that will be so.) This presented a problem for the response we are considering on Clark's behalf—the response that says it's possible to for someone to be going to do something even when the conditions required for doing it aren't satisfied, because practical tendencies are conditional in the same way as intentions, and so in this case you really are going to: go to the beach if the bridge is open. But the distinction I've just drawn between the conditionality of ordinary tendencies and that of conditional intentions suggests another reason why the envisioned response won't work, namely that we seem not to have an *independent* grasp on the relevant notion of 'being going to ( $\phi$  if ...)' to draw on in explaining what intentions are. According to Clark, an intention is a kind of conative state—but then so are desires, preferences, wishes, and hopes, and so we need an account of how intentions are different from these. The appeal to the category of 'tendency' or 'being going to' seemed to promise a way of doing this: for a person can desire, prefer, wish, or hope to do something and *not* be going to do that thing, and so if intention is different in this respect then that helps to specify the kind of conative state that it is. I have just argued, however, that it's only in the special sort of case that Clark is interested in—viz., the sort of case that involves the expression or attribution of conditional intentions—that the tendency described with the phrase 'to be going to ...' has the internally conditional structure in question. It seems, then, that we need a prior grasp of intention in order to understand what it is for someone to be going to:  $\phi$  *on the condition* that such-and-such.

Things are different for the strong cognitivist, however. According to the strong cognitivist, intention is a *cognitive* state like belief—and the conditionality of a practical belief is no different from that of a theoretical one. Further, the strong cognitivist holds that what distinguishes those beliefs about a person's future that are her intentions from those that are mere 'estimates' is that intentions are beliefs that are grounded in practical reasoning. It is, of course, perfectly legitimate to question whether any of this is plausible. But it avoids the kind of circularity that I have argued would come with any attempt to use the category of tendency or 'being going to', understood as something that can be internally conditional, to illuminate the difference between intentions and other kinds of conative state.

I confess that I am somewhat uncomfortable about leaning as much as I have in this section on intuitions about ‘what we would say’ in various hypothetical cases, such as the one in which the bridge leading to the beach is closed. Unfortunately, without these intuitions the ontological category of ‘tendency’ or ‘being going to’ is quite hard to get a handle on. Clark is trying to make good on the idea that, as Anscombe put it, a person ‘can form an intention which he then does nothing to carry out, either because he is prevented or because he changes his mind’ (Anscombe 1963, p. 9), and his proposal is that in such a case the person will have *been going* to do what he nevertheless didn’t. But what are we saying, when we say such a thing? A deep difficulty is that from the forward-looking perspective, the statement ‘I am going to do such-and-such’ entails that this is what I *will do*—as Anscombe notes very clearly:

Nor can we say: But in an expression of intention one isn’t saying anything is going to happen! Otherwise, when I had said ‘I’m just going to get up’, it would be unreasonable later to ask ‘Why didn’t you get up?’ I could reply: ‘I wasn’t talking about a future happening, so why do you mention such irrelevancies?’ (Anscombe 1963, p. 92)

This point applies in turn to statements that attribute an intention to someone else by describing their future actions, such as in the case where I say that you are going to go to the beach. As we have seen, however, from the backward-looking perspective it *is* possible to say of someone (or something) that she (or it) *was going to* do such-and-such, but never actually went on to do it. Repurposing a phrase that Anscombe borrows from Wittgenstein, we may say that when such a statement is used to attribute a past intention, its purpose is to reveal something about a person ‘which goes beyond what [actually] happened’ (*Philosophical Investigations* I, §659; quoted in Anscombe 1963, p. 45): it’s a statement about what the person *then* intended, what she *then* was going to do. And as I have emphasized throughout this section, Clark’s position requires us to treat these statements as picking out a ‘past future’: not just a way things *would have been* (except ...), but a way that they *really were going* to be. The challenge is to get a firm enough handle on this category that we can then use it to explain the nature of intention itself, and also to say whether, as Clark supposes, the truth of such a statement is necessary for a person to have intended to do the thing in question. In this section I have given reason for skepticism on both counts. Intending to do something might, indeed, require having a real tendency toward doing it, but that tendency need not be of the sort that’s expressed with the ordinary phrase ‘going to  $\phi$ ’.

(V)

I have argued, against Clark, that a person can have intended to do something that she *never was going* to do—not because her will was conflicted or liable to change, but because the world was so arranged that it never was possible for her to execute her intention. In such a case the person’s intention will have been ‘complete’ despite her never having been in what Clark calls the ‘state of being going to do’ the thing in question, at least if this phrase is understood as attributing a real tendency. And this possibility undermines Clark’s defense of noninferential weak cognitivism.

But we should be wary of overstating this conclusion. From the fact that it is *possible* to intend to do what one isn’t really going to do, it doesn’t follow that intention *never* consists in a real tendency—that it *never* reaches out to encompass the fact that one is, self-consciously and in a way that is grounded in one’s practical reasoning, going to do a certain thing. That’s the sort of slide that we see on display in the argument from illusion, where the fact that perceptual experience *sometimes* stops short of the world is used to show that it’s *never* a real relation to it. Such reasoning is invalid until further premises are added to it. The same point applies in connection with the present case.

So the mere fact that it is *possible* to intend to do something one isn’t going to do doesn’t show that intention doesn’t *ever* consist in a real practical tendency. Indeed, we can say more than that: it’s also compatible with this being what intentions are in the basic case, i.e., the case that should be central in our analysis of what intention is. And this stronger view should seem attractive when we reflect on the way that intention appears from within the perspective of first-person deliberation.<sup>10</sup> As I explained above, a central motivation for the cognitivist position is that from that deliberative perspective the superficially self-directed question ‘What do I intend?’ is transparent to the world-directed question ‘What am I going to do?’: the *only* way that my present intentions can show up for me, at least in a non-alienated way, is in deliberation about my future actions. And it’s this perspective that should be treated as central in understanding what intention itself is: not a mere desire to do something or a mere judgment that doing it would be reasonable or obligatory, but a *commitment* to carrying it out. If committing oneself in this way weren’t, in the fundamental case, a way of making it the case that one *is going to do* the thing one commits to, then practical deliberation would lose its rationale.

But is Clark right in thinking that the belief that one will do something, where this belief is grounded in one’s practical reasoning, isn’t itself sufficient for having the sort of commitment I’ve just described? In Section II I summarized what I see as Clark’s most powerful argument for this anti-cognitivist conclusion: it was that someone who reasons practically to the belief that she’s going to do something might nevertheless not intend to

---

<sup>10</sup> I develop this idea further in Schwenkler (2019), pp. 203-210.

do it, since it may be that ‘when the chips are down’ the person is not going to follow through—and, indeed, that this is what *was going to happen* all along, notwithstanding the agent’s belief to the contrary. Clark concludes from this that while intending to do something might *entail* having a directive belief that you are going to do it, nevertheless having this belief isn’t *sufficient* for intending, and so the belief itself can’t be what intention consists in.

I think we can resist this conclusion by reflecting on the way that beliefs can differ in what philosophers have called *strength*. Consider, for example, the attitude you might take toward the question of whether a friend is trustworthy, or whether your favorite team is going to win the big game. No matter what you say to others or think to yourself, in cases like this if ‘when the chips are down’ you don’t actually trust your friend or place your bet on your team, this will be evidence that the belief you had—that is, the belief you had *that your friend was trustworthy*, or *that your team was going to win*—was a *weak* belief, since you were prepared to act on it only to the extent that there wasn’t much hanging on the matter. And this point is sometimes put by saying that your eventual behavior reveals that you didn’t *really* believe what you said after all. That’s because if you ‘really had’ believed it, then you’d have been willing to stick by your friend, or your team, despite the significant risks incurred by doing so.

Consider now the case of the person who takes it that she is going to do a certain thing, say break off a longstanding affair, though in fact she isn’t really committed to doing this. We can easily imagine this person *saying* she is going to break things off, and also *thinking to herself* that this is what she is going to do, but as we have just seen all this isn’t sufficient for her to have a strong belief that she is going to do it. For we have to consider further what will happen when it comes time for the person to *act* her supposed commitment, either by breaking things off directly or doing something else whose rationality is contingent on her eventually doing this, such as declining to buy expensive tickets for a getaway. It is in situations like these that ‘the chips are down’: and if we imagine that in such a situation the person comes to the realization that she (perhaps) isn’t going to break off the affair after all (so that, say, she goes ahead and buys the tickets), how are we to understand this realization?

In principle there are two possibilities. First, it might be that the person *really did believe* that she was going to break it off, right up until the crucial moment—and then came to believe instead that she was going to continue the affair. But how is this transition supposed to have happened? It is not as if the person changed her mind because of *evidence* that showed that she wasn’t going to do what she thought she would—as when I think I am going to be somewhere at 5:15 and then come to see that the traffic is too heavy for that. Nor did she *decide* not to follow through after recognizing apparent *reasons* in favor of continuing the affair, or against breaking it off—for as we are imagining the case it was true all along that this is what she was going to do, and so the crucial moment was not one that brought about a change in her intentions. On the present

interpretation what seems instead to happen is that, as of the crucial moment, the person *simply* thinks, and believes, that ‘I am not going to break this off’. And the fact that she is able *simply* to think this gives us reason to say that she didn’t really (strongly) believe anything different beforehand, and that the truth about herself was just too uncomfortable for her to admit.

So here is the better interpretation of the case. A person who takes herself to have decided to do a certain thing that she isn’t really going to do—not because she’s going to forget or be unable to do it or because something will happen to change her mind, but because she isn’t really (strongly) *committed* to doing it after all—also doesn’t really (strongly) *believe* that she is going to do this thing, even as she may tell herself inwardly something different. In addition to helping us understand what *changes* in such a person when the chips are finally down (that is, that the change is a matter of coming to *realize* what she is (and always was) going to do, in the face of circumstances that make this situation clear), this interpretation of the case also makes good sense of the way that, on Clark’s view as well as Marušić and my own, directive beliefs about what one is going to do are grounded in *practical* reasoning rather than in evidence about one’s future behavior. The person who judges, say, that she is going to break off a longstanding affair, will hold this belief in light of considerations that favor *doing* just that, rather than considerations that show it likely that this will happen. And what happens when the chips are down is *not* that this person recognizes new reasons to continue the affair, or sees that her reasons for deciding to break it off were insufficient. Rather, what happens is that she *continues* to see breaking off the affair as favored in just the same way, and to just the same degree, as she had seen it as favored before, only now she realizes that she is not, in fact, actually going to break it off for these reasons. This suggests that the belief she previously had, on those grounds, that she was going to break the affair off was a *weak* belief. If the belief had been stronger, she wouldn’t have given it up so easily when the time came to put it into action. Her belief was precisely as weak as her intention was—just as the strong cognitivist would have it.

This response to Clark’s argument recalls the one that Marušić and I put forward in response to a similar objection in our original paper. There, we criticized the idea ‘that belief is purely intellectual and dispassionate, that the criteria for belief cannot require commitment and motivation, and that what a person believes floats free of how things are with her “deep self”’ (Marušić and Schwenkler 2018, p. 334). What we failed to emphasize, however, is that matters like these are relevant not just to the *presence* of a belief but also to what we called its *strength*. The argument I’ve just given hinges on the idea that having a strong belief that *p* requires having a stable, rationally grounded commitment to act on the truth of *p* even ‘when the chips are down’ and there are temptations to do otherwise. A person who is committed in this way to the truth of the judgment that she’ll do a certain thing is committed to acting in the way that she believes she will. To the extent that her commitment is strong and stable, this is exactly how she is

going to act. Unless she is going to fail, be prevented, or have something happen that changes her mind, this person *is going to do* the thing that she believes she will.

#### REFERENCES

Anscombe, G.E.M. (1963). *Intention*, 2<sup>nd</sup> ed. Repr. 2001, Cambridge: Harvard University Press.

Bratman, M. (1999). *Intention, Plans, and Practical Reason*, rev. ed. Stanford: CSLI Publications.

Clark, P. (2020). 'Intentions, Intending, and Belief: Noninferential Weak Cognitivism'. *Pacific Philosophical Quarterly* 101: 308-327.

Ferrero, L. (2009). 'Conditional Intentions'. *Noûs* 43: 700-741.

Frankfurt, H. (1998). 'The Faintest Passion'. In *Necessity, Volition, and Love* (New York: Cambridge University Press), pp. 95-107.

Marušić, B., and J. Schwenkler. (2018). 'Intending Is Believing: A Defense of Strong Cognitivism'. *Analytic Philosophy* 59: 309-340.

Schwenkler, J. (2019). *Anscombe's 'Intention': A Guide*. New York: Oxford University Press.

Schwenkler, J. (Forthcoming). 'How Can There Be Reasoning to Action?' *Analytic Philosophy*.