

The final and definitive version of this paper will be published in *Philosophical Studies*.

AGAINST COGNITIVISM ABOUT PRACTICAL RATIONALITY

John Brunero

Cognitivists about Practical Rationality argue that we can explain some of the (apparent) requirements of *practical* rationality by appealing to the requirements of *theoretical* rationality. First, they argue that intentions involve beliefs, and, second, they show how the theoretical requirements governing those involved beliefs can explain some of the practical requirements governing those intentions (or they show how these apparently practical requirements are actually theoretical requirements). This paper avoids the ongoing controversy about whether and how intentions involve beliefs and focuses instead on this second part of the Cognitivist approach, where I think Cognitivism faces significant difficulties. I proceed by considering two attempts by Cognitivists to explain (apparent) requirements of practical rationality and I argue that neither of them succeed.

Reflection on the relationship between intention and belief, specifically on the way in which intention involves belief, has recently led some philosophers to think that we can appeal to theoretical rationality to explain some apparent requirements of practical rationality. This approach has come to be known as "Cognitivism about Practical Rationality", or simply "Cognitivism" (though it bears no relation to the view to which this term traditionally refers in meta-ethics).¹ (I here say "apparent requirements of practical rationality" because some Cognitivists, like Kieran Setiya, have argued that some rational requirements that are often supposed to be practical, such as the instrumental principle, are actually theoretical rather than practical, while other Cognitivists, like R. J. Wallace, have argued that the instrumental principle is a genuinely practical requirement of rationality which is explained by appealing to theoretical rationality.² To cover all such views, we'll say that Cognitivists appeal to theoretical rationality to explain some apparent requirements of practical rationality.) In this paper, we'll be concerned with two (apparent) requirements of practical rationality in particular: the requirement to have consistent intentions and the requirement to be means-ends

coherent (also sometimes called "the instrumental principle") -- since these are the two requirements to which Cognitivists have paid the most attention so far.³

The Cognitivist approach to practical rationality divides into two parts. First, Cognitivists defend some thesis about the way in which *intention involves belief*. Perhaps they defend the view that intending to x involves believing that one will x , or the view that intending to x involves believing that it is possible to x . Second, Cognitivists argue that some of the requirements of theoretical rationality which govern those involved beliefs (such as consistency and closure requirements) explain some of the (apparent) requirements of practical rationality governing those intentions. As we'll see below, Cognitivists differ on the details of the explanation they provide.

There has been a great deal of attention paid to the first part of the Cognitivist approach, and a great deal of controversy. A number of philosophers have defended the view, sometimes called the "Strong Belief Thesis", that intending to x involves believing that one will x .⁴ Some proponents of the Strong Belief Thesis have argued that the belief here is self-referential: intending to x involves believing that one will x because of that very intention.⁵ And others have *identified* an intention to x with a desire-like, self-referential belief of this sort.⁶ Yet other philosophers have rejected the Strong Belief Thesis, including Michael Bratman, who has presented some famous counterexamples to it, and Richard Holton, who has recently argued that intending to x involves neither full belief nor even *partial* belief that one will x .⁷

In light of this controversy, some Cognitivists have made use of a weaker thesis about how intention involves belief, which we could call the "Weak Belief Thesis," according to which intending to x involves believing that x is possible. R. Jay Wallace,

for instance, has developed a Cognitivist account of means-ends coherence working from the Weak Belief Thesis.⁸ While the Weak Belief Thesis is certainly less controversial, significant doubts have been raised about whether it is strong enough to do the work that it must do for the Cognitivist -- for instance, doubts have been raised about whether one could account for the consistency requirement on intentions using the Weak Belief Thesis -- suggesting that Cognitivists would need to appeal instead to the more controversial Strong Belief Thesis.⁹

The strategy of this paper is to sidestep the controversy about whether and how intentions involve beliefs. We'll just assume, for the purposes of this paper, that the Strong Belief Thesis is true. We'll focus instead on the second part of the Cognitivist approach: the claim that the requirements of theoretical rationality which govern those involved beliefs about what one will do can explain some of the (apparent) requirements of practical rationality governing those intentions. I'll argue in this paper that there are significant difficulties with this part of the Cognitivist approach; specifically, I'll argue that the requirements of theoretical rationality cannot, at least in the ways that have been proposed, explain either the requirement to have consistent intentions or the requirement to be means-ends coherent.

Before considering how the Cognitivist explanations would go, we should first introduce the (apparent) requirements of practical rationality that Cognitivists seek to explain, as well as the requirements of theoretical rationality that they use to explain them. There is, of course, some disagreement about how to precisely formulate the requirements of rationality I mention here, but I do not think any such disagreement will affect the lines of argument presented in this paper.

Let's first consider some requirements of practical rationality. Suppose you intend to travel to Prague this afternoon, and intend to travel to New York this afternoon, and believe that if you travel to Prague this afternoon, you will not travel to New York this afternoon. You are in an irrational state: your intentions are, given what you believe, inconsistent with one another. You could remove yourself from this irrational state by either abandoning your intention to travel to Prague this afternoon, abandoning your intention to travel to New York this afternoon, or by abandoning your belief that if you travel to Prague this afternoon, you will not travel to New York this afternoon (though your abandoning this belief may be prohibited by some other requirement of theoretical rationality). Since there are three ways to remove yourself from this irrational state, the requirement to have consistent intentions is usually understood as a "wide-scope"¹⁰ requirement in that the scope of "requires" does not attach to a specific attitude, but ranges over a disjunction of attitudes:

Intention Consistency (IC): Rationality requires that [if one intends to x , and believes that if one will x then one will not y , then one not intend to y].¹¹

IC does not require that you abandon some specific attitude; rather, it requires that you *either* abandon your belief that if you will x then you will not y , *or* abandon your intention to x , *or* abandon your intention to y .

Practical rationality also requires that we be *instrumentally* rational, or, as it is sometimes put, that we be *means-ends coherent*. Suppose you intend to travel to Prague this afternoon, and you believe that to do so you must buy an airline ticket, but you don't intend to buy an airline ticket. You are in an irrational state: you do not intend to do what you believe to be a necessary means for carrying out your intention. A requirement of

means-ends coherence requires that you *either* form an intention to buy the airline ticket, *or* abandon your intention to travel to Prague this afternoon *or* abandon your belief that to do so you must buy an airline ticket.¹² We could formulate this as a wide-scope requirement:

Means-Ends Coherence (ME): Rationality requires that [if one intends to x , and believes if one will x , then one will intend to y , then one intend to y].

Note that I've used "if one will x , then one will intend to y " instead of "if one will x , then one will y " in formulating *ME*. I've done so because we are not rationally required to form the intentions to do *whatever* we believe is a necessary condition for carrying out an intention to x . For instance, suppose I believe I will travel to Prague only if I make use of a mode of transportation that emits carbon dioxide into the atmosphere. Surely I am not rationally required to *intend* to make use of a mode of transportation that emits carbon dioxide into the atmosphere. Rationality does not require that we *intend* to bring about the side effects that we *foresee* will come about as a result of other intentions we have. Our formulation of *ME* takes this into account by limiting the applicability of *ME* to cases in which one believes that the formation of an intention to y is necessary for x -ing. Since the agent in our example does not believe that his *formation of an intention* to make use of a carbon-dioxide-emitting mode of transportation is necessary for traveling to Prague, *ME* does not require him to form such an intention. But *ME* does require him to intend to buy an airline ticket since (we are supposing) he believes he must form this intention in order to travel to Prague. Now there might be other, and better, ways to formulate *ME* that take into account this problem of necessary side effects. But, again, I

do not suspect that the precise formulation of these requirements of rationality will affect the lines of argument to follow.

Let's now consider some requirements of theoretical rationality, which, like the requirements of practical rationality above, are wide-scope requirements. First, there is the wide-scope requirement:

Belief Consistency (BC): Rationality requires that [if one believes p , and believes $p \rightarrow \neg q$, then one not believe q .]

Suppose you believe that it's raining, and that if it's raining, the streets are not dry, but you also believe that the streets are dry. You have a set of beliefs that violates *BC*. You could come to comply with *BC* by revising any one of the beliefs in this set. In addition to not having inconsistent beliefs, theoretical rationality requires us to come to believe the logical consequences of our beliefs, as required by

Belief Closure (CL): Rationality requires that [if one believes p , and believes $p \rightarrow \neg q$, then one believe $\neg q$].

Like the others, *CL* is a wide-scope requirement; someone in violation of *CL* could comply with it by abandoning the belief that p , abandoning the belief that $p \rightarrow \neg q$, or coming to believe $\neg q$. There might be some controversy over *CL*, but since *CL* is employed in Setiya's version of Cognitivism, considered below, we'll grant it here for the sake of argument.¹³

In this paper, I proceed by considering two Cognitivist explanations. In §1, I consider an account which takes the Strong Belief Thesis to be true and argues that *BC* explains *IC*. In §2, I consider an account which again takes the Strong Belief Thesis to be true and argues that *CL* (along with another requirement of theoretical rationality)

explains *ME*. I argue that neither of these explanations work. It is of course possible that there is a better Cognitivist account yet to come. I hope this paper will show some of the difficulties that such an account would have to resolve or avoid.

§1. *Belief Consistency and Intention Consistency*

If we grant the truth of the Strong Belief Thesis, then it seems that we could provide a rather straightforward explanation of the consistency requirement on intentions in terms of the consistency requirement on beliefs. Someone who violates *IC* -- that is, someone who intends to *x*, believes that if he will *x*, then he will not *y*, and intends to *y* -- will, assuming the truth of the Strong Belief Thesis, hold the following beliefs:

- (1) I will *x*
- (2) If I will *x*, I will not *y*.
- (3) I will *y*.

And these beliefs violate *BC*. One could come to comply with *BC* by ceasing to believe (2), which would also be a way of complying with *IC*, or by ceasing to believe (1), which, according to the Strong Belief Thesis, would constitute a way of ceasing to intend to *x*, and so would constitute a way of complying with *IC*, or by ceasing to believe (3), which, likewise, would constitute a way of ceasing to intend to *y*, and so would constitute a way of complying with *IC*. In short, once we grant the truth of the Strong Belief Thesis, it seems quite easy to account for the consistency requirement on intentions in terms of the consistency requirement on the "involved" beliefs about what one will do.

I want to suggest here that matters are not so simple. According to the Strong Belief Thesis, believing that one will *x* is *necessary* for intending to *x*. But problems for

the above attempt to use *BC* to explain *IC* arise from the fact that believing one will *x* is *not sufficient* for intending to *x*. I'll argue here that (a) believing one will *x* is not sufficient for intending to *x*, and (b) this fact generates a problem for this attempt to use *BC* to explain *IC*.

The thesis that believing one will *x* is sufficient for intending to *x* is vulnerable to some obvious counterexamples. Suppose a gangster is now kidnapping me and I come to believe that I will go to the gangster's hideout. Surely I do not *intend* to go to the gangster's hideout, even though I believe I will.¹⁴ We could circumvent this and similar counterexamples by instead claiming that *believing one will x intentionally* is sufficient for intending to *x*. But this does not seem true either, for two reasons. First of all, one's belief could be false. Perhaps I am led by wishful thinking to believe that I'll (intentionally) fight and escape from the gangsters who are holding me hostage, but in fact I have no intention of fighting them. Second, even if one's belief is true, it doesn't follow that one *now* intends to *x*. Consider the example from Anscombe of the individual who believes that he'll give up some piece of information after being tortured by the gangsters but does not now intend to (and now intends not to) give up the information.¹⁵ He believes he will give up the information intentionally but does not now intend to give up the information. (Perhaps such an individual is irrational in not now intending to give up the information -- after all, if you know you're going to give it up eventually, why not just give it up now and avoid the torture? But surely it is *possible* for someone to be irrational in this way.)

So, what does this have to do with the attempt to use *BC* to explain *IC*? Well, if it is possible for one to believe one will *x* without intending to *x*, then it is possible for one

to have inconsistent beliefs about what one will do without having inconsistent intentions. Specifically, it is possible for one to believe (1), (2) and (3) without it being the case that one intends to x and intends to y . Perhaps one merely intends to x and does not intend to y . But here's the problem for Cognitivism: we want to be able to say that if such an individual (someone who intends to x and does not intend to y) were to form the intention to y (without changing any other attitudes) he would be introducing some *additional* irrationality into his attitudes -- specifically, in addition to the theoretical conflict between beliefs (1), (2) and (3), he would now be coming to have inconsistent intentions as well, and this would violate a requirement of rationality, specifically *IC*. Rationality prohibits him from forming the intention to y (without changing any other attitudes). But how can the Cognitivist explain why this is so? For the Cognitivist, the problem with coming to have inconsistent intentions is that doing so involves coming to have inconsistent beliefs about what one will do. But the individual in our example *already* has inconsistent beliefs about what he will do. And so the need to avoid coming to have inconsistent beliefs about what he will do cannot explain why he should avoid having inconsistent intentions in this case! Some other explanation is needed.¹⁶

We could illustrate the difficulty with another example borrowed from Anscombe: St. Peter believes that he will deny Jesus three times (based on the testimony of a highly reliable predictor: Jesus himself) and yet he intends not to deny Jesus.¹⁷ Assume the Strong Belief Thesis is true. According to that thesis, if one intends not to deny Jesus, then one believes one will not deny Jesus. So, St. Peter believes he will not deny Jesus *and* he believes that he will deny Jesus -- he has inconsistent beliefs. Now, if St. Peter were to form an additional intention to deny Jesus, without abandoning his

intention not to deny Jesus, he would introduce an *additional* conflict of attitudes. But, again, according to the Cognitivist account, the problem with coming to have inconsistent intentions is that in doing so one comes to have inconsistent beliefs about what one will do. But St. Peter *already* has inconsistent beliefs about what he will do. And so the need to avoid having inconsistent beliefs about what he will do cannot explain why St. Peter should here avoid having inconsistent intentions.

How might a Cognitivist reply to this objection? Perhaps the Cognitivist could argue that forming the intention to *y* (without changing any other attitudes) would introduce *some other* conflict prohibited by theoretical rationality -- that is, some conflict other than the conflict of believing (1), (2) and (3) -- and the need to avoid *this other conflict* explains why one should avoid forming the intention to *y* (without changing any other attitudes). But it is hard to see what this other conflict could be. Perhaps it's true that the formation of an intention to *y* would involve a change in belief. As many advocates of the Strong Belief Thesis point out, intentions involve *self-referential* beliefs: intending to *y* involves believing that one will *y because of that very intention*. For example, whereas St. Peter currently believes that he will deny Christ, but he doesn't know *how* he will (since he's simply trusting the testimony of a highly reliable predictor) were he to form an *intention* to deny Christ, he would believe that *he will deny Christ because he intends to deny Christ*. He would have an explanation he did not have before. But *that* change in belief does not introduce any new conflict prohibited by theoretical rationality. It's true that there remains the original theoretical conflict -- St. Peter believes that *he will deny Christ* and that *he will not deny Christ* -- but there is no additional theoretical conflict generated by the self-referential belief; it simply provides an

explanation where there was none before. But if there is no additional theoretical conflict that would be generated by his forming an intention to deny Christ, then we cannot present a Cognitivist explanation for why, to be rational, St. Peter should not form this intention.

Perhaps there is another line of reply available to the Cognitivist. The Cognitivist could argue that we think that St. Peter's forming the intention to deny Christ is irrational because his forming the intention provides a *second* sufficient reason for his belief that he will deny Christ -- the belief which is part of the theoretical conflict. His believing that he will deny Christ is now overdetermined by reasons: the fact that he intends to deny Christ *and* the fact that a highly reliable predictor said he would do so are both sufficient reasons for him to believe that he will deny Christ. So, according to this reply, St. Peter's forming the intention would further entrench (by providing a second sufficient reason) a belief involved in the theoretical conflict, and this explains why it is irrational for him to come to intend to deny Christ. But this reply will not work for the simple reason that there is nothing irrational about coming to have a second sufficient reason for a belief that is involved in a theoretical conflict. If I believe p and $\neg p$, there is nothing irrational about my coming to believe in some additional piece of evidence that provides a sufficient reason for believing $\neg p$. Of course, I'm irrational insofar as I continue to believe, at the same time, p and $\neg p$, but there is no additional theoretical irrationality involved in my coming to believe in the additional piece of evidence.

Perhaps there is another, more radical, line of reply available to the Cognitivist. Perhaps the Cognitivist could claim that an intention *just is* a belief about what one will do.¹⁸ Call this the "Identity Thesis." If the Identity Thesis is correct, then a conflict

between inconsistent intentions *just is* a conflict between inconsistent beliefs about what one will do. And so, when someone who intends to x and believes (1), (2) and (3) comes to intend to y , he does not introduce a *second* conflict of attitudes at all. There is only one conflict of attitudes throughout: the conflict in believing (1), (2) and (3). There is no further conflict of attitudes *distinguishable from* this theoretical conflict.

The reply is a radical one since it would require that we revise our understanding of *IC*. It is clear that here intending to y (without changing any other attitudes) would introduce a violation of *IC* where there was no violation of *IC* before. So, to say that intending to y does not introduce a second, distinct conflict of attitudes requires that we *revise IC* -- the very requirement that the Cognitivist is supposed to be *explaining*.¹⁹

But I think we can reject the radical reply quite easily since the Identity Thesis is false -- for reasons originally pointed out by Donald Davidson. As Davidson argued, we cannot identify an intention with a belief about what one will do because "reasons for intending to do something are in general quite different from reasons for believing one will do it."²⁰ Consider St. Peter once more. The fact that loyalty is a virtue provides St. Peter with a reason to *intend* not to deny Christ, but this consideration by itself provides him with no reason to *believe* that he will not do it. The qualification "by itself" is there because there could be other considerations that allow for this consideration to also be a reason for him to believe that he will not do it -- perhaps St. Peter knows he is reliably motivated to do what is virtuous. But it is nonetheless true that the consideration *by itself* is a reason to *intend* to act without by itself being a reason to *believe* that one will so act. (Likewise, the fact that a highly reliable predictor predicted that St. Peter would deny Christ provides St. Peter with a reason to *believe* he will deny Christ but this

consideration by itself provides him with no reason to *intend* to deny Christ.) Since something can be a reason to intend to x without being a reason to believe that one will do x , we cannot identify intending to x with believing that one will do x .

It follows that we cannot identify a conflict between inconsistent intentions with a conflict between inconsistent beliefs about what one will do. And so it is clear that when the person who intends to x and believes (1), (2) and (3) comes to intend to y , he does indeed introduce a conflict of attitudes *distinguishable from the conflict between his beliefs (1), (2) and (3)*. Some consideration may count in favor of a certain way of resolving the conflict between his intending to x , intending to y and believing (2) *without* by itself counting in favor of a certain way of resolving a conflict between his believing (1), (2), and (3). The conflicts are therefore not identical.

Let's sum up. I have argued in this section that even if we grant the Strong Belief Thesis, we need to pay attention to the fact that believing one will x is not sufficient for intending to x . This fact makes it possible for a person to have inconsistent beliefs about what he will do without having inconsistent intentions. But some explanation is needed for why such a person should not go on to form inconsistent intentions, and the Cognitivist, I've argued, is not well positioned to provide such an explanation. I've considered three possible replies a Cognitivist could make to circumvent this difficulty and I've argued that none of them succeed. So, we can conclude that even if we suppose the Strong Belief Thesis is true, we cannot use *BC* to explain *IC*.

§2. *Belief Closure and Means-Ends Coherence*

Kieran Setiya has presented an interesting Cognitivist account which defends the Strong Belief Thesis and argues that *CL* (along with another requirement of theoretical rationality) explains *ME*.²¹ As I mentioned at the start of the paper, other formulations of the requirement to be means-ends coherent are possible. Here is Setiya's formulation:

(*ME**) You should [if you intend to do *E* and believe that you will do *E* only if you do-*M*-because-you-now-intend-to-do-*M*, intend to do *M*]²²

Like the my original *ME*, *ME** avoids the problem of necessary side effects since it doesn't apply simply when one believes that *M*-ing is necessary; one must believe that doing-*M*-because-one-now-intends-to-do-*M* is necessary. (Returning to the example I gave earlier of the person who intends to travel to Prague, since he does *not* believe that making use of a carbon-dioxide-emitting mode of transportation-*because-one-now-intends-to-do-so* is necessary for traveling to Prague, he is not required by *ME** to form the intention to do so.) But unlike my original *ME*, *ME** applies only when one believes doing-*M*-because-one-now-intends-to-do-*M* is necessary. If I intend to travel to Prague and believe that I will travel to Prague only if I intend to buy an airline ticket, but I think I could form the intention to buy the ticket *five minutes from now* and still travel to Prague, then *ME** does not apply to me. For *ME** to apply, I must believe that my *now intending* to buy the ticket is necessary. As Setiya puts it, *ME** applies only when one believes that a decision is necessary.²³ It is not my purpose here to argue either for or against *ME** as a formulation of the requirement of means-ends coherence; I'm simply explaining Setiya's view so that we can turn to his Cognitivist explanation of *ME**.

As I mentioned earlier, Setiya thinks that intention involves belief. More precisely, he defends the view that intention is a matter of a motivating, or desire-like,

self-referential belief. As he puts it, "when I intend to ϕ , the content of my attitude is that I am going to ϕ because of that very intention: intention represents itself as motivating action."²⁴ He argues that someone who intends some end E will believe:

(4) I will do E .

And when ME^* is applicable to him, he will believe:

(5) I will do E only if I will do- M -because-I-now-intend-to-do- M .

Applying CL to these beliefs, we can say that someone who believes (4) and (5) is rationally required to either abandon (4), abandon (5), or come to believe:

(6) I will do- M -because-I-now-intend-to-do- M .²⁵

Now, *one* way for one to come to believe (6) is by coming to intend to do M . (According to Setiya's theory of intention, which we are assuming is correct, coming to intend to do M involves believing (6).) But, of course, one could also come to believe (6) *without intending to do M* . But the important question is whether one could come to believe (6) without intending to do M and *without violating any other requirement of theoretical rationality*. If there is some other requirement of theoretical rationality that would prohibit believing (6) without intending to do M , then we can present a Cognitivist explanation of ME^* by appealing to CL and *that other requirement of theoretical rationality*. These requirements of theoretical rationality would then *together* explain means-ends coherence. More precisely, according to Setiya, we could then see how ME^* is actually a requirement of theoretical reason rather than practical reason.²⁶

So, the important question, as I mentioned, is whether one could come to believe (6) without intending to do M and *without violating any other requirement of theoretical rationality*. Setiya argues that one could not. Let's consider his argument.

Setiya begins his argument by discussing a certain kind of pathological inference: the inference of p from one's belief that p . He argues that one reason this inference is pathological is that

an inference of this kind could never be both sound and ampliative. Whenever I would form a new belief by inferring that p from the belief that I believe that p , the premise of my inference is false: I did not, in fact, believe that p . Whenever the premise is true, however, the inference is redundant: I already believe that p . In the epistemic sense of should, one should never make an inference of this pathological kind.²⁷

He immediately goes on to argue that similar considerations show why it is incoherent to believe (6) without intending to do M :

What does this have to do with intention and instrumental reason? It is a curious fact that the belief that one intends to ϕ can only be acquired, on the basis of evidence, by an inference that is defective in just this way. An inference to the conclusion that I intend to ϕ , from *any* premise, will instantiate a pattern that cannot be both sound and ampliative. For suppose that I infer that I intend to ϕ on the ground that p . If the conclusion is false, the inference is unsound. But if the conclusion is true, the self-reference of intention ensures that the inference is redundant. If I intend to ϕ , I already believe that I am going to ϕ because I so intend. Theoretical inference cannot support the belief that I intend to ϕ without involving a defect that amounts to incoherence -- like inferring the facts from my own beliefs. It follows, I think, that there is something incoherent about the belief that I intend to ϕ , unless it is constituted by the intention to ϕ .²⁸

If Setiya's argument is successful, then we can provide a Cognitivist explanation of *ME** by appealing to *CL* and *the requirement to avoid this kind of incoherence*.

However, I think there is a significant flaw in this argument that I'll explain below. To put it quickly, the flaw is that the question of whether there could be an inference to *p* that is both *sound* and ampliative is not relevant to whether an agent is *rational* (or *coherent*, to use Setiya's phrasing) in believing, or coming to believe, *p*. Even if it is true that any ampliative inference to *p* would have to be *unsound*, it does not follow that any ampliative inference to *p* would have to involve *theoretical irrationality*. Rationality, to use a slogan I like, is a matter of whether our beliefs *cohere* with one another, not whether they *correspond to the world* (or are based on other beliefs which correspond to the world).²⁹ And so the worry about the unavailability of a *sound* ampliative inference to *p* is out of place in this context. This sketch of an objection will be presented in more detail below.

An example will help here. Let's suppose that I falsely believe that I intend to leave my wife for some other woman. (Perhaps I later come to see that my actions better supported the hypothesis that I merely desired, and did not intend, to leave her for some other woman.³⁰) And let's suppose that I come to have this belief because my psychiatrist, who I reasonably think knows me much better than I know myself, has told me that this is what I intend to do. There are several ways we could model my reasoning leading to this false belief. Perhaps we could model my reasoning as follows:

- (7) My trusted psychiatrist thinks I intend to leave my wife.
- (8) If my trusted psychiatrist thinks I intend something, then I intend it.
- (10) I intend to leave my wife.

I here come to believe (10) by reasoning from (7) and (8). Or perhaps we could model my reasoning more simply as follows:

(9) There is conclusive evidence that I intend to leave my wife.

(10) I intend to leave my wife.

In coming to believe (10) on the basis of (9), I would be coming to comply with a requirement of rationality, specifically the requirement that Niko Kolodny has formulated as:

(B+) Rationality requires one to believe that p , if one believes that there is conclusive evidence that p .³¹

And, as Kolodny argues, in proceeding from the content of (9) to formation of the belief that (10), one would be (properly) *reasoning* one's way into compliance with (B+).³²

How we choose to model my reasoning in this case is not really that important. What is important is that my belief in (10), which is false, is one that it is rational for me to hold given the other beliefs I have, and one which I've arrived at in a rational way. Now, as Setiya points out, since (10) is false, there is no *sound* inference to (10). Since (10) is false, one of my premises on the first model, (7) or (8), would also have to be false. But it does not follow that in coming to believe (10), and in continuing to believe (10), I display any theoretical irrationality. My belief in (10) is based on what I (reasonably) believe the evidence conclusively supports! And so even though it's true that any ampliative inference to (10) would have to be *unsound*, it does not follow that any ampliative inference to (10) would have to involve *theoretical irrationality*. Nor does it follow that my believing (10) or my coming to believe (10) would have to involve theoretical irrationality.

Let's return to Setiya's argument. Setiya is right that if I were *already* to intend to leave my wife, then, according to his theory of intention, I would already believe (10). And, if I were to already believe (10), then an inference to (10) would not be ampliative. In order for an inference to (10) to be ampliative, I would have to *not already believe* (10), which requires that I *not already intend* to leave my wife, which means that (10) would have to be *false*. But if (10) is false, he argues, any inference to (10) would have to be unsound. But here is where my objection comes in: as I argued above, even if it's true that any ampliative inference to (10) would have to be unsound, it does not follow that any ampliative inference to (10) would have to involve theoretical irrationality, nor does it follow that my believing (10) or my coming to believe (10) would have to involve theoretical irrationality. But if there need not be any theoretical irrationality here, then we will be unable to explain *ME** in the way that Setiya suggests -- that is, by appealing to *CL* and the requirement to avoid the supposed theoretical incoherence involved in having false beliefs about what one intends.

What about the analogy that motivates Setiya's argument? Setiya argues that just as there is something incoherent in (a) inferring that *p* from one's belief that *p*, there is something incoherent, in just the same way, in (b) inferring that one intends to ϕ from *any* premise whatsoever. I've argued above that (b) need not involve any incoherence. Am I now also committed to thinking that (a) need not involve any incoherence as well?

I'm actually not sure whether (a) involves incoherence or, alternatively, there is just something peculiar about (a) and this peculiarity falls short of incoherence. But let's consider what might be the source of that peculiarity or incoherence in (a). Let's suppose that I reason from

(11) I believe Princeton is in New Jersey.

to

(12) Princeton is in New Jersey.

What's peculiar about such reasoning (if that's the right word to describe it), to my mind, is that normally one should not be looking *inward towards one's own beliefs* in order to draw a conclusion about the location of Princeton. Rather, one should be looking *outward at the world*: does the evidence (a map, perhaps) point to Princeton being in New Jersey? That's what strikes me as peculiar about such reasoning. But note that this peculiarity is not involved in the lines of reasoning I modeled above: there one is looking to *what one's trusted psychiatrist has said*, or to *what the evidence conclusively supports* in order to form a belief about what one intends. So, it seems that we might be able to explain why even if I'm right that (b) involves no incoherence, there is still something peculiar, and perhaps incoherent, about (a). Of course, this is merely a rough sketch; much more could be said about the peculiarity or incoherence involved in (a). I merely want to suggest that Setiya's analogy between (a) and (b) might not be apt, and that there may be a way of analyzing the alleged incoherence involved in (a) that doesn't commit us to thinking that the *same* incoherence is involved in (b).

Let's sum up. Setiya argues that we can explain means-ends coherence by appealing to *CL* and a requirement to avoid the incoherence involved in believing that one intends to *M* without actually intending to *M*. He argues that believing one intends to *M* without actually intending to *M* is incoherent because any ampliative inference to this belief would have to be unsound. This same kind of incoherence, he argues, is involved when one infers the facts from one's own beliefs. I've argued that even if any ampliative

inference to this belief would have to be *unsound*, it doesn't follow that there is any theoretical *incoherence* involved in believing it. And the kind of incoherence (or perhaps mere peculiarity) involved in inferring the facts from one's own beliefs might be due to something else entirely.

§4. *Conclusion*

I've argued in this paper that even if Cognitivists can successfully establish the truth of the Strong Belief Thesis, there are difficulties involved in their accounts of how the requirements of theoretical rationality explain the (apparent) requirements of practical rationality. I've considered two Cognitivist explanations -- one which takes the Strong Belief Thesis to be true and argues that *BC* explains *IC*, and one which takes the Strong Belief Thesis to be true and argues that *CL* (along with another requirement of theoretical rationality) explains *ME* -- and I've argued that neither of them succeed. I consider these to be the most plausible Cognitivist explanations on offer, but of course there could be another, better Cognitivist account yet to come. I hope this paper points to some difficulties that such an account would have to resolve or avoid.³³

References

- Anscombe, G.E.M. (1963) *Intention*. Cambridge, MA, Harvard University Press.
- Armstrong, David (1963) "Is Introspective Knowledge Incorrigible?" *The Philosophical Review*, 72: 417-432.
- Bratman, Michael (1981) "Intention and Means-Ends Reasoning" *The Philosophical Review*, 90: 252-265.
- (1987) *Intentions, Plans and Practical Reason*. Cambridge, MA, Harvard University Press.
- (1999) "Cognitivism about Practical Reason" reprinted in *Faces of Intention*. Cambridge, Cambridge University Press.
- (forthcoming-a) "Intention, Belief, Practical, Theoretical" forthcoming in Jens Timmerman, John Skorupski, and Simon Robertson, (eds.) *Spheres of Reason*.
- (forthcoming-b) "Intention, Belief, and Instrumental Rationality" forthcoming in David Sobel and Steven Wall (eds.) *Reasons for Action*. Cambridge, Cambridge University Press.
- Bridges, Scott (forthcoming) "Rationality, Normativity and Transparency" *Mind*.
- Broome, John (1999) "Normative Requirements" *Ratio*, 12: 398-419.
- , (2007) "Wide or Narrow Scope?" *Mind* 116: 359-370.
- , "The Unity of Reasoning" forthcoming in Jens Timmerman, John Skorupski, and Simon Robertson, (eds.) *Spheres of Reason*.
- Brunero, John (2005) "Two Approaches to Instrumental Rationality and Belief Consistency" *Journal of Ethics and Social Philosophy*, 1: 1-20.
- Davidson, Donald (2001) "Intending" reprinted in *Essays on Actions and Events*, Second Edition. Oxford, Oxford University Press.
- Hampshire, Stuart and H.L.A. Hart (1958). "Decision, Intention, and Uncertainty" *Mind*, 67:1-12.
- Harman, Gilbert (1976) "Practical Reasoning" *Review of Metaphysics*, 29: 431-463.
- (1988) *Change in View*. Cambridge, MA, MIT Press.
- Holton, Richard (2008) "Partial Belief, Partial Intention" *Mind*, 117: 27-58.
- Kolodny, Niko (2005) "Why be Rational?" *Mind*, 114: 509-563.
- (forthcoming) "The Myth of Practical Consistency" *European Journal of Philosophy*.
- Raz, Joseph. "The Myth of Instrumental Rationality" *Journal of Ethics and Social Philosophy*, 1: 1-28.
- Setiya, Kieran (2007a) "Cognitivism about Instrumental Reason" *Ethics*, 117: 649- 673.
- , (2007b) *Reasons without Rationalism*. Princeton, Princeton University Press.
- Velleman, J. David (2007) *Practical Reflection*. Stanford, CA, CSLI Publications, David Hume Series Edition.
- (unpublished) "What Good is a Will?"
- Wallace, R. J. (2001) "Normativity, Commitment and Instrumental Reason" *Philosopher's Imprint*, 1: 1-26.
- (2006) "Postscript to 'Normativity, Commitment and Instrumental Reason'" in *Normativity and the Will: Selected Essays on Moral Psychology and Practical Reason*. Oxford: Oxford University Press.

Notes

¹ This usage of the term was coined by Michael Bratman, who has criticized versions of Cognitivism in the course of arguing for his influential Planning Theory of intention and agency, presented in Bratman 1987. For his critiques of Cognitivism, see Bratman 1999, forthcoming-a, forthcoming-b.

² See Setiya 2007a: 671-673 and Wallace 2001: 21.

³ This paper will narrowly focus on Cognitivism about *requirements of practical rationality*, and will not have space to consider some interesting Cognitivist approaches to agency and practical reason, including the important arguments in Broome forthcoming, and Velleman unpublished.

⁴ See Hampshire and Hart 1958, Grice 1971, Velleman 2007, Harman 1976, 1986. It is no objection to the Strong Belief Thesis that we are well aware of the possibility that we might not carry out our intentions because of some interference or because we might change our minds. According to the thesis, intending to *x* involves believing one will *x*, but one could also believe that one's beliefs, including this one, are fallible. For a more thorough response to this objection see especially Harman 1976: 432-433 and also Anscombe 1963: §52.

⁵ See Harman 1976, 1988.

⁶ See 2007a: 663-664 and 2007b: Part I, especially 48-56.

⁷ See Bratman 1987: 37-38, and Holton 2008. See also Davidson 2001: 91-96.

⁸ See Wallace 2001 and Wallace 2006. For critical discussion of Wallace's view, see Raz 2005, Brunero 2005: 4-6, and Bratman forthcoming-a: §3 and §4.

⁹ Michael Bratman has pointed out that the Weak Belief Thesis cannot, by itself, explain the consistency requirement on intentions. Suppose it is inconsistent to intend to x and intend to y . Even if it's true that, as the Weak Belief Thesis contends, intending to x involves believing that x is possible and intending to y involves believing that y is possible, the conjunction these beliefs does not support the belief that x and y is possible. But now we have no explanation for why it is rationally criticizable to intend to x and intend to y . An appeal to the Strong Belief Thesis can avoid this problem since, as Bratman notes, "a conjunction of beliefs in the success of each intention does support a belief in conjunctive success." See Bratman forthcoming-a: §3. So, if an account of Cognitivism is to provide an explanation of the consistency requirement on intentions, it would do better to work from the Strong Belief Thesis instead. Such an account is considered in §1 below. (Bratman also provides a critique of Wallace's account of means-ends coherence. See Bratman forthcoming-a: §4.)

¹⁰ See Broome 1999, 2007 on wide-scope requirements. Broome 2007 defends, against objections from Kolodny 2005, the view that all requirements of rationality are wide-scope requirements.

¹¹ A more complete specification of the requirement would include temporal markers. See, for instance, Kolodny's *Simple Intention Consistency (C)*, according to which, "one is rationally required (either not to believe at t (that if one X 's, then one does not Y), or not to intend at t to X , or not to intend at t to Y). See Kolodny, forthcoming.

¹² Note that while *IC* requires of you, in this case, merely that you *not intend not to buy* an airline ticket (or revise one of the other relevant attitudes), *means-ends coherence*

requires that you *intend to* buy an airline ticket (or revise one of the other relevant attitudes).

¹³ See Setiya 2007a: 665-666. There is one minor difference: where my formulation has "Rationality requires", Setiya's has "You should", which, he explains, is "the epistemic 'should'." This difference does not matter for my argument against Setiya's Cognitivism below.

¹⁴ A slight variation on an example from Bratman 1981: 255.

¹⁵ Anscombe 1963: 94. I've varied the example slightly for stylistic reasons.

¹⁶ It is true that Cognitivists need not explain the *whole* of practical rationality in theoretical terms. They may extend the scope of their Cognitivism to some requirements of practical rationality (such as *IC* and *ME*) but not to others. But to provide an adequate Cognitivist explanation for any single requirement of practical rationality, the Cognitivist must show how in every case in which an intention makes applicable that requirement of practical rationality, there is an explanation in terms of theoretical rationality for why this is so. As Bratman notes, the Cognitivist seeks to explain why intentions *quite generally* are subject to that requirement of practical rationality, not merely why a certain subset of intentions are subject to it. See Bratman forthcoming-a: §2, paragraph 3.

¹⁷ Anscombe 1963: 94.

¹⁸ David Velleman has defended a view along these lines in his *Practical Reflection* but he has since distanced himself from that view. See Velleman 2007: xix-xx.

¹⁹ We should note that the requirement of intention consistency is controversial, and is challenged in Kolodny's forthcoming paper "The Myth of Practical Consistency". I have been arguing that if we accept *IC*, then we should reject the Cognitivist attempt, outlined

above, to use *BC* to explain *IC*. But let's now suppose we follow Kolodny in rejecting *IC*. As Kolodny notes, those who reject *IC* need to provide an alternative explanation of the *Violation Intuition (V)*, according to which "If one believes at *t* (that if one *X*'s, then one does not *Y*), intends at *t* to *X*, and intends at *t* to *Y*, then one violates some requirement." Kolodny goes on to present two explanations of *V*: a Cognitivist explanation (§2), and (because he is skeptical about the Strong Belief Thesis presupposed in the Cognitivist explanation) a non-Cognitivist explanation (§3). I find the latter quite persuasive, but find the former inadequate because of the problem posed by the case of St. Peter described above. According to Kolodny's Cognitivist explanation (leaving aside the details), "if one has inconsistent intentions, then one has some belief that violates a narrow-scope requirement of reason." Specifically, he claims that "If one believes (that if one *X*'s, then one does not *Y*), intends to *X*, and intends to *Y*, then either one believes without sufficient reason that one will *X*, or one believes without sufficient reason that one will *Y*, or one believes without sufficient reason that (if one *X*'s, then one does not *Y*)." Thus, he argues, we have a Cognitivist explanation of *V*: in short, if one has inconsistent intentions, then one has a set of beliefs such that some narrow-scope requirement of theoretical reason is violated. But here's the problem: St. Peter *already has* inconsistent beliefs about what he will do -- he already has a set of beliefs such that some narrow-scope requirement of theoretical reason is violated. And so the need to avoid having such beliefs cannot explain why he should here not form the additional intention to deny Jesus. And yet we have the intuition that his forming this intention would violate some requirement of rationality. And so this Cognitivist explanation of *V* seems inadequate.

²⁰ Davidson 2001: 95.

²¹ Setiya 2007a.

²² Setiya 2007a: 668. *ME** is my label, not his.

²³ Setiya 2007a: 667-668. Setiya argues that when one doesn't believe that a decision is necessary, one could simply, without irrationality, trust oneself to form the intention at some later time. I could, for example, without irrationality, trust that I will at some later time form the intention to buy the airline ticket to Prague. That's why *ME** applies only when one believes a decision is necessary.

²⁴ Setiya 2007a: 664.

²⁵ Note that abandoning (4) would constitute a way of ceasing to intend to do *E*, and so would be a way of coming to be means-ends coherent. And abandoning (5) would be a way of complying with both *CL* and *ME**.

²⁶ Setiya 2007a: 671-673.

²⁷ Setiya 2007a: 670.

²⁸ Setiya 2007a: 670-671.

²⁹ I've borrowed the slogan from Scott Bridges, forthcoming. The argument below does not depend on this slogan; I just find it to be a nice way to put the point.

³⁰ The possibility of one's observable behavior being used to evaluate the truth of one's beliefs about one's own attitudes is discussed by David Armstrong 1963: 427.

³¹ Kolodny 2005: 521. Kolodny argues that this is a narrow-scope requirement. But this point does not matter for the argument in the text here.

³² Kolodny 2005: 520.

³³ Acknowledgements.