Guest Editor: Julian Fink

Contents · Obsah

Within the debate concerning reason and rationality, instrumental *incoherence was for a long time conceived of as the paradigm of irrationality. It was also conceived of as a metaphysically and ontologically innocuous source of normativity (Mackie 1977, 27-28; Dreier 1997, 93; cf. also Raz 2005, 26): if, for example, you intend to write a preface and you think that a necessary means to do so is to ignore your incoming emails, yet you have no motivation whatsoever to ignore your incoming emails, then your attitudinal coherence displays a normative failure. You are not as coherent as you ought to be.*

However, with the emergence of the so-called 'bootstrapping objection' (Bratman 1997; Broome 2001) and the debate concerning the 'scope' of rational requirements (e.g. Broome 2007; 2013; Brunero 2010; Kolodny 2005; Rippon 2011; Schroeder 2004), the innocuous status of the normative significance of (instrumental) coherence became subject to debate (Broome 2005; 2008; Raz 2005; Kolodny 2005; 2007). This led to a paradigmatic shift in how to understand the relationship between rational requirements and normativity. While there now exists considerable doubt that rational requirements are normative, it is commonly agreed that one's normative *point of view is a key feature of one's rationality. Here the question is not only if one can hold a particular normative judgement and still be rational; what is significant too is whether your normative outlook coheres appropriately with your motivation. In fact, it is now commonly agreed that rationality requires us to intend to make the world fit with our first-personal ought beliefs. Enkrasia (i.e. coherence between your normative views and your motivation) is thus seen as a rational ideal and as a source of rational requirements.*

Nevertheless, many elementary questions regarding the application, content, and significance of an enkrasia-requirement *remain unanswered, or subject to debate. For example, why does rationality require us to be* enkratic? *Can an irrational ought-belief issue requirements upon us? What is the logical form (i.e. 'scope') of the enkrasia-requirement? Is the enkrasia-requirement best formulated as a state- or as a process-requirement? Should we formulate the enkrasia-requirement as a synchronic or diachronic requirement? Are there normative reasons to satisfy the enkrasia requirement? Does the enkrasia-requirement constitute a standard of correct reasoning? Do ought-beliefs rationally cause intentions with or without the help of an external motivational attitude?*

This special issue on 'The Nature of the Enkratic Requirement of Rationality' aims to answer some of these fundamental questions. The present papers take direct issue with the plausibility of an enkratic rationality, how to formulate a requirement of enkratic ra-

tionality correctly, and whether enkratic reasoning represents correct reasoning. Furthermore, the assembled papers explore the relationship between enkratic rationality and normative uncertainty, the potential conflict between local enkratic coherence and one's overall-degree of rationality, the authority of normative beliefs with regard to other fundamental principles of rationality, the possibility of rational akrasia, the 'bootstrapping objection' and its potential to thwart the view that attitudes entail normative reasons, and whether enkratic coherence really goes as far as requiring you to have an intention to A whenever you believe that you ought to A. This special issue represents an important step forward in elucidating these elementary issues pertaining to a general theory of rationality.

In closing, I would like to express my gratitude to those who contributed to the genesis of this special issue. First, and foremost, I am very grateful to Marián Zouhar, the chief editor of Organon F, not only for inviting me to edit this special issue, but also for advising and helping me during the editorial process. I would also like to express my gratitude to Tibor Pichler, the director of the Institute of Philosophy, for facilitating my fruitful cooperation with the Slovak Academy of Sciences and for making many things possible. Many thanks also to Lukáš Bielik for being extremely helpful during the entire editorial process.

I would also like to thank the authors of this special issue for their insightful contributions. I am very grateful that almost all authors took up my invitation to present their papers at a workshop at the University of Vienna (May 2013) sponsored by the ERC Advanced Grant 'Distortions of Normativity'. The discussions at the workshop shaped, and sharpened, the arguments presented in this special issue. I am also very grateful to the numerous reviewers for their constructive suggestions and criticism.

Last, I would like to express my pride in the fact that this collection of papers from eminent and aspiring 'Western philosophers' is published in a Central European journal. I feel that an intensive cooperation and exchange between academics and institutions in the so-called 'West' and 'Central and Eastern Europe' is already long overdue. Both sides should intensify their efforts in facilitating such cooperation in the future.

Julian Fink
julian.fink@univie.ac.at

References

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

Broome, J. (2001): Are Intentions Reasons? And How Should We Cope with Incommensurable Values? In: Morris, C. – Ripstein, A. (eds.): *Practical Rationality and*

Preference: Essays for David Gauthier. Cambridge: Cambridge University Press, 98-120.

Broome, J. (2007): Wide or Narrow Scope? *Mind* 116, 359-370.

Broome, J. (2005): Have We Reason to Do as Rationality Requires? A Comment on Raz. *Journal of Ethics and Social Philosophy* 1, 1-9.

Broome, J. (2008): Is Rationality Normative? *Disputatio* 11, 153-171.

Broome, J. (2013): *Rationality through Reasoning.* Oxford: Wiley—Blackwell.

Brunero, J. (2010): The Scope of Rational Requirements. *Philosophical Quarterly* 60, 28-49.

Dreier, J. (1997): Humean Doubts about the Practical Justification of Morality. In: Cullity, G. – Gaut, B. (eds.): *Ethics and Practical Reason.* Oxford: Clarendon Press, 81-99.

Kolodny, N. (2005): Why be Rational? *Mind* 114, 509-563.

Kolodny, N. (2007): How Does Coherence Matter? *Proceedings of the Aristotelian Society* 107, 229-263.

Mackie, J.L. (1977): *Ethics: Inventing Right and Wrong.* London: Penguin.

Raz, J. (2005): The Myth of Instrumental Rationality. *Journal of Ethics & Social Philosophy* 1, 2-28.

Rippon, S. (2011): In Defense of the Wide-Scope Instrumental Principle. *Journal of Ethics & Social Philosophy* 5, 1-21.

Schroeder, M. (2004): The Scope of Instrumental Reason. *Philosophical Perspectives* 18, 337-364.

Enkrasia

JOHN BROOME

Corpus Christi College. Oxford OX1 4JF. United Kingdom
john.broome@philosophy.ox.ac.uk

ABSTRACT: Rationality requires you to intend to do what you believe you ought to do. This is a rough formulation of the requirement I call 'Enkrasia'. This article presents a precise formulation. It turns out to be complicated, and the article also explains the need for each complication.

KEYWORDS: Enkrasia – intention – ought – practical reason – rationality – requirement.

Necessarily, if N is within the domain of rationality, rationality requires of N that

if	(1)	N believes at t that she herself ought that p
and if	(2)	N believes at t that, if she herself were then to intend that p, because of that, p would be so
and if	(3)	N believes at t that, if she herself were not then to intend that p, because of that, p would not be so
then	(4)	N intends at t that p.

This is the requirement of rationality I call 'Enkrasia'. To put it very roughly, ignoring the many qualifications that are embedded in the formulation: rationality requires you to intend to do what you believe you ought to do.

Enkrasia is a central feature of our rationality; it is at the heart of how a rational person conducts her life. It constitutes one of the main bridges between theoretical and practical rationality. We often spend time thinking

about what we ought to do. This is an exercise of our theoretical rationality; theoretical rationality is concerned with beliefs, and this process aims at forming a belief about what we ought to do. But it generally has a practical purpose. Thinking about what we ought to do is generally part of a longer process of deliberation that ultimately aims at deciding what to do.

When you think about what you ought to do and reach a conclusion, the belief you emerge with – the belief that you ought to do some particular thing – engages your practical rationality. Specifically, it engages with this requirement Enkrasia. If you are practically rational to the extent of satisfying Enkrasia on this occasion, you end up intending to do what your theoretical deliberation brings you to believe you ought to do.

The precise formulation of Enkrasia is complicated. This paper explains and justifies the complications, one fragment at a time.[1]

Necessarily

I assume that enkrasia, if it is a feature of rationality, is a necessary feature. It could not be a contingent matter whether or not rationality requires you to intend to do what you believe you ought to do.

if N is within the domain of rationality

Rationality is one of a class of things that, for want of a better name, I call 'sources of requirements'. Other sources of requirements are morality, fashion, the rules of chess and the law of each country. Sources issue requirements, but not universally; each source has a domain of application. For example, UK law requires people to drive on the left, but UK law has a limited jurisdiction that determines its domain of application. With some exceptions, it applies only to people in the UK. So, if you are within the UK, UK law requires you to drive on the left.

Similarly, rationality has a limited domain of application and issues requirements only to things within that domain. It does not apply to stones, trees, sheep or babies. I shall not try to specify the limits of the domain except to say that presumably only things that have some rational capacity fall within it. I assume that, as a matter of contingent fact, only people have this capacity, so rationality applies only to people.[2]

---

[1]   It is explained in more detail in my book (2013), particularly in section 9.5.

[2]   This issue is more thoroughly explored in Julian Fink's (ms).

Since rationality has a limited domain, we may say that it is conditional in its application.

rationality

Even someone who agrees that you are required to intend to do what you believe you ought to do might doubt it is rationality that requires it. She might think it is morality, for example. She might think that conscientiousness consists in intending to do what you believe you ought to do, and that conscientiousness is a part of morality.

However, it cannot always be morality that requires you to do what you believe you ought to do, because you often believe you ought to do something on grounds that have nothing to do with morality. For example, you might believe you ought to drink some water because otherwise you will get a headache. Your belief stems from self-interest rather than morality. Enkrasia nevertheless applies to you, and it could not be a requirement of morality in this case.

Could it be that in this case it is self-interest that requires you to intend to do what you believe you ought to do? In general, could it be that the requirement to do what you believe you ought to do issues from whatever source you believe the ought derives from? That would be odd. You may believe the ought mentioned in Enkrasia derives from various sources, but what is derived from these sources is always the same, all-things-considered ought that is described below. So it would be odd if Enkrasia issued from different sources in different cases.

Moreover, you may have no belief about what is the source of the ought. Perhaps your friend tells you that you ought to do something, and you believe her, but she does not tell you why. On other occasions too, you may believe your ought derives from more than one source. For instance, if you are a parent, you may believe that prudence and morality together determine that you ought not to climb dangerous mountains, although neither would do so on its own. Then the requirement not to be akratic in this case cannot issue from either prudence or morality.

Intuitively, rationality is concerned with coherence in your mind. Since Enkrasia requires a sort of coherence, this is a reason for thinking it is a requirement of rationality.

One argument may suggest it is not. Satisfying Enkrasia may demand an effort. It often involves overcoming a strong pull in the opposite direction. You may believe you ought to do something that you hate doing. We

might think rationality should not demand that sort of effort. Rationality is a matter of getting your attitudes into coherent order. That may demand a certain sort of ability, but we might think it should not demand effort.

However, we have to recognize that rationality does sometimes require effort. Suppose you believe you are a man and that all men are mortal. You may find it very hard to bring yourself to believe you are mortal, or to give up either the belief that you are a man or the belief that all men are mortal. Nevertheless, you are definitely irrational if you do not do one of these things. So the difficulty of satisfying Enkrasia is a weak argument against the view that it is a requirement of rationality.

requires

Compare 'Survival requires you to have a functioning liver' with 'The law requires you to wear a seat-belt'. 'Requires' has different meanings in these two sentences.

In the first, the subject of 'requires' is the name of a property, survival. 'Requires' has what I call 'the property sense'. The sentence means that having a function liver is a necessary condition for having the property of survival.

In the second sentence, the subject of 'requires' is the name of a source of requirements, the law. 'Requires' has what I call 'the source sense'. The law places requirements on you, which are specified in the legal code. True, the law's requirements are connected with a property; satisfying each of them is a necessary condition for having the property of being law-abiding. Wearing a seat-belt is a necessary condition for being law-abiding. But 'The law requires you to wear a seat-belt' says more than this; it implies also that not wearing a seat-belt is against the law. Not every necessary condition for being law-abiding is required by the law. A necessary condition for being law-abiding is to have a level of competence that allows you to understand the law. But the law does not require this level of competence; being legally incompetent is not against the law.

The names of some properties have, by the process of reification, become also the names of sources of requirements. 'Morality', 'prudence' and 'rationality' are examples. So 'rationality' has two meanings; it is the name of a property and also the name of a source of requirements. Consequently 'rationality requires' is ambiguous; 'requires' might have either the property sense or the source sense.

We might acceptably say that rationality requires you to have a functioning liver, since having a function liver is a necessary condition for being alive and being alive is a necessary condition for being rational. This is to use the property sense of 'requires'. However, the property sense is not very natural for rationality, and I do not use it.

As I use 'rationality requires', in Enkrasia and elsewhere, 'rationality' is the name of a source of requirements and 'requires' has the source sense. Rationality as a source places requirements on us, which we may call the 'code' of rationality. Not every necessary condition for being rational is required by rationality in this sense. For example, a necessary condition for being rational is to have some mental capacity: stones and snails are not rational. But having mental capacity is not in the code of rationality.

Enkrasia does not merely say that intending to do what you believe you ought to do is a necessary condition for being rational. It says further that this requirement is part of the code of rationality.

of

Rationality places the requirement specifically on the person N. N herself and no one else has the responsibility of intending to do what she believes she ought to do. The requirement is 'owned' by her, as I put it. It is a satisfactory feature of the locution 'required of' that it makes the ownership of the requirement explicit. Conversely, it is an unsatisfactory feature of 'ought' that it lacks this feature, as I shall explain below.

N

Strictly Enkrasia is a requirement schema rather than a requirement. To get a requirement, take the schema and make appropriate substitutions for the schematic letters. An appropriate substitution for 'N' is a term. In practice, this term will always refer to a person, since only people are within the domain of rationality. Informally, I often use the generic 'you' instead of 'N'.

if

What rationality requires of N is a conditional proposition. The requirement stated in Enkrasia is conditional in two different ways. First, it is conditional in its application, as I have explained: it applies only to things that are within the domain of rationality. Second, it is conditional in its content: what is required is a conditional proposition.

Some authors formulate Enkrasia as conditional only in its application and not in its content. They deny that rationality requires of you that, if you believe you ought to do something, you intend to do it. Instead, they assert that, if you believe you ought to do something, rationality requires of you that you intend to do it. I call their claim 'Narrow-Scope Enkrasia', because it gives 'requires' a narrower scope than Enkrasia does.

Narrow-Scope Enkrasia has the consequence that rationality is not normative in a particular sense. When I say that a source of requirements is normative, I mean that, when the source requires something of you, that fact constitutes a reason for you to do what it requires. Narrow-Scope Enkrasia implies that rationality is not normative in that sense.

Here is why. Suppose you believe you ought to do something. According to Narrow-Scope Enkrasia, it follows that rationality requires you to do it. If rationality is normative, it follows that you have a reason to do it. So from merely believing that you ought to do something it follows that you have a reason to do it. But it is implausible that that could be so. Suppose your belief is mistaken, and actually you had no reason at all to do this thing before you formed the belief. How could your mistaken belief bootstrap into existence a reason when there was none? It is implausible that it could. On the basis of an argument like this, and assuming that Narrow-Scope Enkrasia is true, Niko Kolodny draws the conclusion that rationality is not normative (cf. Kolodny 2005).

But that is an undesirable conclusion, and it is certainly undesirable to rest it on the assumption that Narrow-Scope Enkrasia is true when the alternative of Enkrasia is available. Enkrasia itself implies no such conclusion.[3] This is my main reason for adopting it in preference to Narrow-Scope Enkrasia.

In my formulation of Enkrasia, 'if' denotes material implication.

believes

N's belief has a normative content. Noncognitivists claim that to believe you ought to do something is to have some noncognitive attitude such as favouring towards doing it. Nevertheless, modern noncognitivists do not deny that you may have normative beliefs (for instance, Allan Gibbard 2003). They think it may be true that you ought to do something, and that therefore you may believe it. So a noncognitivist can accept Enkrasia.

---

[3]   Setiya (2007) claims it does, but I reject this claim in Broome (2013, sec. 5.7).

at *t*

Appropriate substitutions for '*t*' are terms that refer to times. The same letter '*t*' occurs in each clause of Enkrasia. This indicates that this is a synchronic requirement; it connects together contemporaneous attitudes. This means that the requirement does not directly regulate processes of deliberation, which take time. Suppose you believe you ought to *F* and then you come to intend to *F* as a result of deliberation. Your deliberation may not bring you to satisfy Enkrasia, because by the time the deliberation is competed you may have ceased to believe you ought to *F*.

she herself

'She herself' is a reflexive pronoun. It represents in indirect speech the first-person pronoun 'I' in direct speech. *N* would herself express the content of her belief by saying 'I ought that *p*'. (I assume she uses my grammatical innovation 'ought that', explained below).

Enkrasia would be false if it did not have a reflexive pronoun in this position. Suppose a donkey lies dead, and you believe that its owner ought to bury it. Suppose that actually you are its owner. Then you believe in a sense that you ought to bury the dead donkey. But suppose you do not realize the donkey is yours, so you could not express your belief by saying 'I ought to bury the dead donkey'. You do not believe that you yourself ought to bury the dead donkey; you do not self-ascribe the ought. Given that, you might not intend to bury the dead donkey, and nevertheless be perfectly rational.

It is often said that practical rationality is 'first personal'. This is one sense in which that is so. The normative belief must be self-ascribed.

ought

Some philosophers use the expressions 'morally ought', 'rationally ought', 'prudentially ought' and so on. They also use 'all-things-considered ought'. By 'ought' I mean what they mean by 'all-things-considered ought'.

I do not use 'morally ought' or other adverbially qualified 'ought's. Instead of 'You morally ought to *F*' I say 'Morality requires you to *F*'.

There are various sources of requirements, such as morality, self-interest and the law. Some of these sources are normative in the sense I explained: you have a reason to satisfy their requirements. For some *F*, you may be under one or more normative requirements to *F* and one or more

not to *F*. You then have some reasons to *F* and some not to *F*. Whether or not you ought to *F* is determined in some way by the combination of these reasons. It does not matter here how it is determined. It might be that the reasons weigh against each other in some way, or it might be that some reasons override others.

So even granted that morality is normative, if morality requires you to *F* it need not follow that you ought to *F*. Morality might be overridden by a stronger normative requirement not to *F*.

ought that *p*

An appropriate substitution for '*p*' is a sentence, which expresses a proposition. '*p*' does not necessarily describe an act, either of *N*'s or of anyone else's. For example, it might describe *N*'s having some intentional attitude such as a belief.

My expression 'ought that' is ungrammatical. I use it for a good reason. Enkrasia would be false if the ought was not *owned*, as I put it, by *N*. Suppose Alex believes that he himself ought to get a medal. He believe that he himself has done a heroic act, for which he deserves a medal. However, he does not think it is it down to him – it is not his responsibility – to get a medal; he thinks the authorities have the responsibility of awarding him one. Clearly he might be entirely rational even if he does not intend to get a medal.

So if Enkrasia is to be correct, *N* must believe that the ought mentioned in Enkrasia is owned by her herself – it is her responsibility. But there is no good way of expressing ownership of an ought in grammatical English. Compare the sentence 'Alex ought to get a medal' with 'Alice ought to get a haircut'. 'Ought' has different meanings in these sentences as they would most commonly be used. The second ascribes ownership of the ought to Alice; the first does not ascribe ownership of the ought to Alex. But these sentences have identical grammar. Ownership is therefore not grammatically represented.

I use my grammatical innovation 'ought that' as a means to specify ownership. The grammatical problem with 'ought' is that it is an auxiliary verb. It is always attached in use to a lexical verb, to form a single compound verb that has only one subject. In Alex's case, the single compound verb is 'ought to get'. 'Alex' is its subject. But we need room for two subjects: the subject of 'ought' should be the name of the owner of the ought, and the lexical verb should have a separate subject. In Alex's case, the own-

er of the ought is the authorities, and they should be named as the subject of 'ought', whereas the subject of the lexical verb 'get' should be 'Alex'. My construction allows this. I say 'The authorities ought that Alex gets a medal'.

Suppose Alex believes that the authorities ought that he himself gets a medal. However, suppose he does not believe that he himself ought that he himself gets a medal. Then he does not satisfy condition (1) of Enkrasia. So Enkrasia does not imply that Alex is not rational if he does not intend that he himself gets a medal. This is just as it should be.

By contrast, suppose Alice believes that she herself ought to get a haircut. More exactly, as I put it, Alice believes that she herself ought that she herself gets a haircut. Alice satisfies condition (1) of Enkrasia, so Enkrasia implies that she is not rational if she does not intend that she herself gets a haircut. This too is just as it should be.

Notice it is the ought that has to be self-ascribed. It does not matter whether or not there is any self-ascription in the proposition p. Suppose Alice believes her daughter Zuleika ought to get a bicycle, but that this is her own – Alice's – responsibility. In my language, Alice believes that she herself ought that Zuleika gets a bicycle. Enkrasia correctly implies that Alice is not rational if she does not intend that Zuleika gets a bicycle.

(2) N believes at t that, if she herself were then to intend that p, because of that, p would be so

Briefly, using an expression of Kant's, N believes that p is 'in her power'.

Suppose you believe that, for prudential reasons, you ought to believe God exists. You believe there is a possibility that God exists and that, if he does, things will go badly for you unless you believe he exists. However, you do not believe you could now bring yourself to believe God exists, in view of the lack of evidence for his existence. You do not believe that, if you were now to intend to believe God exists, you would believe God exists. In your position you might rationally not intend to believe God exists.

But then, is it not irrational for you to continue to believe you ought to believe God exists, given that you do not believe you can now achieve this result? The contents of your beliefs are apparently inconsistent with the principle that ought implies can, so are you not irrational on that account? Not necessarily. You may not believe that ought implies can. Even if ought actually does imply can in this case, it does not do so obviously. You might rationally not believe it. Alternatively, you might believe that, though in-

tending *now* would not bring you to believe God exists, you might later come to believe God exists. Then your belief is not inconsistent with the principle that ought implies can.

Clause (2) in Enkrasia is designed for people like you. You do not satisfy the condition stated in clause (2), so Enkrasia does not imply you are not rational if you believe you ought to believe God exists but do not intend to do so. This is as it should be.

(3) N believes at t that, if she herself were not then to intend that p, because of that, p would not be so

Briefly, N believes that p is 'up to her'. This is a converse of condition (2). Suppose you believe you ought to continue breathing, but you do not believe that your continuing to breath depends on an intention of yours. Then you can be rational even if you do not intend to continue breathing. More generally, rationality does not require you to intend anything you believe you ought to do, if you do not believe an intention of yours is needed to achieve your doing it.

she herself within (2) and (3)

Clauses (2) and (3) require self-ascription, for much the same reason as (1) does. Suppose you have found a dead donkey and you believe that the finder of any dead donkey has a responsibility for making sure it is buried (but not necessarily for burying it herself). That is to say, you believe that you yourself ought that the dead donkey is buried. However, you believe the owner will bury it. You are the owner, but you do not realize it. So although it is actually true, you do not believe that, were you yourself not then to intend that the dead donkey is buried, because of that the dead donkey will not be buried. You do not satisfy condition (3). In your situation, you may fail to intend that the dead donkey is buried, without violating Enkrasia.

That explains why (3) contains the reflexive pronoun. (2) contains it for a similar reason.

Within (2) and (3): then

'Then' in (2) and (3) represents in indirect speech the indexical 'now' in direct speech. N would express to herself the belief described in (3) using the sentence: 'If I were not now to intend that p, because of that p would not be so'.

Suppose you believe you ought to repent of your sins before you die. However, you believe you have plenty of time to make up your mind to do so, so you do not believe that, were you not now to intend to do so, you would not do so. Given that, you may rationally not now intend to do so. That is consistent with Enkrasia.

That explains 'then' in (3); 'then' appears in (2) for a similar reason.

Within (2) and (3): **were**, **because of that**, and **would**

The 'because of that' clauses in (2) and (3) are an abridgement. Spelt out more fully, the conditions are:

(2) N believes at t that, if she herself were then to intend that p, p would be so because she herself then intends that p.

(3) N believes at t that, if she herself were not then to intend that p, p would not be so because she herself does not then intend that p.

The embedded conditionals in these clauses are subjunctive and causal. They have to be both, because no weaker sort of conditional will do.

For example, Enkrasia would be false if we omitted the 'because' clauses. Suppose you believe that God decides who will go to heaven and who will not, and that he implants an intention to go to heaven in just those people who will go there. You believe that, were you to intend to go to heaven you would go there, and were you not to intend to go to heaven you would not go there, but you believe that your intention has no effect on whether or not you go there. Suppose you believe you ought to go to heaven. Then Enkrasia without the because clauses would imply you are not rational if you do not intend to go to heaven. But that is false. You could fail to intend to go to heaven and yet be rational. You believe intending to go to heaven is pointless, since it has no effect. So the because clauses are required.

intends

Enkrasia can be understood as the principle that rationality requires you not to be akratic. Traditionally, to be akratic is to fail to do something you believe you ought to do. But rationality cannot require you not to be akratic in this sense. Even a perfectly rational person might fail to do something she believes she ought to do. For example, suppose you believe you ought to catch the 12.35 train to London, but suppose the platforms are switched

without any announcements, and the result is that you actually catch the 12.35 train to Southampton instead. You fail to do what you believe you ought to do, but nevertheless you may be entirely rational.

This can happen because rationality supervenes on the mind, as Ralph Wedgwood puts it in (2002). Rationality requires your mind to be in good order. The example shows that, even when your mind is in good order, you may fail to do what you believe you ought to do. So rationality cannot require you to do what you believe you ought to do. However, intending to do what you believe you ought to do is a matter of proper order in your mind; your intention, which is a mental state, properly matches your normative belief, which is another mental state. This is why Enkrasia is about intending rather than acting.

Nevertheless, I still understand Enkrasia to be the principle that you should not be akratic. Defying tradition, I interpret akrasia as failing to intend to do something you believe you ought to do.

References

BROOME, J. (2013): *Rationality Through Reasoning.* Wiley-Blackwell.

FINK, J. (ms): A Constitutive Account of 'Rationality Requires'. Submitted to *Erkenntnis.*

GIBBARD, A. (2003): *Thinking How to Live.* Harvard University Press.

KOLODNY, N. (2005): Why Be Rational? *Mind* 114, 509-563.

SETIYA, K. (2007): Cognitivism about Practical Reasoning. *Ethics* 117, 649-673.

WEDGWOOD, R. (2002): Internalism Explained. *Philosophy and Phenomenological Research* 65, 349-369.

Is the *Enkratic* Principle a Requirement of Rationality?[1]

ANDREW REISNER

Department of Philosophy. McGill University. Leacock Building, Room 908
855 Sherbrooke St. W. Montreal, Quebec. H3A 2T7 Canada
andrew.reisner@mcgill.ca

ABSTRACT: In this paper I argue that the enkratic principle in its classic formulation may not be a requirement of rationality. The investigation of whether it is leads to some important methodological insights into the study of rationality. I also consider the possibility that we should consider rational requirements as a subset of a broader category of agential requirements.

KEYWORDS: John Broome – enkrasia – enkratic principle – rational requirements – rationality.

0. Introduction

There is a tradition in the philosophical literature that treats rationality, or at least part of rationality, as imposing requirements on the relations amongst our mental states. One part of this tradition is in particular con-

---

cerned with consistency amongst the contents of the relevant mental states. As a basis for rational requirements, consistency looks like an attractive feature, especially when thinking about theoretical rationality and practical rationality as comprising separate islands of requirements. But at least some rational requirements that purport to link theoretical and practical rationality are not as obviously apt for explanation in terms of consistency.

One such requirement is the *enkratic* principle (EP). EP tells us that we are rationally required to intend to φ whenever we believe that we ought to φ. Pre-theoretically, EP holds much appeal. While familiar, the failure of individuals to act (or intend to act) as they believe that they ought is a persistent and disturbing aspect of human agency. On the view that rational requirements are consistency constraints on relations amongst the contents of an agent's mental states, the failure to satisfy EP will only be a rational defect if it turns out to be a form of inconsistency. Not all defects in agency are rational defects, or so I shall contend. Thus the question remains as to whether the failure to satisfy EP is a rational defect, or one of some other kind.[2]

In attempting to determine how to classify the defect of failing to conform to EP – as a rational defect or a defect of some other kind – it is tempting to rely on our intuitive judgements about what is rational and what is not. In many areas of practical philosophy, we give a certain priority to our intuitions and build our theories to fit them. Applied to theorising about rational requirements, we can identify two extremes in the relationship between intuitions and theories.

At one extreme, there is what one might call 'intuitionism'. Intuitionism about rational requirements relies, as its name suggests, on our intuitions to tell us which putative rational requirements are in fact rational requirements. At the other extreme there is what one might call 'algorithmic systematicity'. We develop an admissibility algorithm to tell us whether a relation amongst mental states is a requirement of rationality, and we designate it appropriately based on whatever features the algorithm takes into account.

Recent influential work by John Broome on EP takes a broadly intuitionistic approach, although he sees a kind of family resemblance amongst

---

[2]   The view that this is a rational failure is widespread and relates to the tradition of viewing *akrasia* as a rational failure. On EP as a positive rational requirement, see Broome (2013), Coates (2013), and Wedgwood (2007).

the requirements of rationality.[3] At the far opposite extreme, there is decision theory, in which rational requirements are based on a strong notion of logical or mathematical consistency. Under intuitionism, EP looks plausible to many philosophers, and I shall not challenge its pre-theoretical appeal.

The project of this paper is instead to explore how plausible it is to classify EP as a rational requirement, when we move towards the algorithmic end of the methodological spectrum and when we understand rationality as being principally concerned with consistency amongst an agent's mental states. In this sense, the paper offers a conceptual experiment, examining how well a particular methodological approach and a particular criterion for being a rational requirement can be squared with the common intuition that EP is a rational requirement.

This experiment, I shall argue, has three interesting consequences. The first is that EP cannot be generated from considerations of strict consistency. This is not a surprising consequence, but it does provide some insight into the limitations of treating rationality strictly in terms of consistency. The second consequence is that a widened notion of consistency can accommodate an EP-like requirement, but whether EP itself can be accommodated depends on exogenous theoretical commitments in action theory and the theory of normative reasons. The third conclusion is that we must be careful to remember that there is more to agency than rationality. Not all requirements of agency are rational requirements, and not all agential defects are rational defects.

1. The *enkratic* principle

The *enkratic* principle gets its name from John Broome (see Broome 2013), but it appears under a variety of other guises and came to central prominence in the work of A.C. Ewing.[4] There are also a number of va-

---

[3]   See Broome (2013) for a discussion of how rational requirements relate to a broad notion of coherence.

[4]   Although there has of late been a revival of interest in the *enkratic* principle, there is an interesting literature on it dating back now over 50 years. See Dancy (1977), Ewing (1959), and Greenspan (1975) for some examples. In more recent literature, see Broome (2013), Skorupski (2010), and Wedgwood (2007).

riants of EP. We should distinguish between three versions of EP at the outset:[5]

E1. Synchronic narrow-scope EP: You are, if you believe that you ought to φ, rationally required to intend to φ.

EF1. $BO\varphi \rightarrow RR(I\varphi)$

and

E2. Synchronic wide-scope EP: You are under a rational requirement such that if you believe that you ought to φ, then you intend to φ.

EF2. $RR(BO\varphi \rightarrow I\varphi)$

and

E3. Diachronic narrow-scope EP: You are required to intend to φ at time t_2 if you believe that you ought to φ at time t_1.

EF3. $B_{t_1}O\varphi \rightarrow RR(I_{t_2}\varphi)$

E1 and E2 are synchronic rational requirements, and E3 is a diachronic rational requirement. In this paper, I shall be working with E2. Of course, philosophers who think that rational requirements in general are narrow scope and synchronic or who think that all requirements are diachronic may find this unsatisfactory. To them, I offer a brief apology.

This paper asks about EP's place as a rational requirement, where being a rational requirement is at least in part being a consistency requirement. Considerations of consistency are inadequate for producing interesting narrow-scope synchronic requirements. For very similar reasons, they are inadequate for generating interesting diachronic requirements.[6] So, I shall not evaluate EF1 or EF3 here.[7]

---

[5] Various qualifiers may be in order for each of these versions. I omit them to ease explication, as they do not affect the substance of the discussion in this paper. For an extensive discussion of further qualifiers, see Fink (2012).

[6] See Reisner (2009a) for a discussion why interesting diachronic requirements require more than consistency. Kolodny (2005) provides a good discussion of what beyond consistency is needed to generate interesting narrow-scope synchronic requirements. See Broome (2007) and Kolodny (2007) for more on the relationship between narrow scope synchronic requirements and diachronic rational requirements.

[7] See Broome (2007) for an in depth discussion of the difference between wide and narrow scope synchronic requirements.

It should also be noted that rational requirements in this paper are treated as sets of local judgements. It is tempting to think that a global judgement of an agent's rationality supervenes in a straightforward way on the satisfaction of all of her local rational requirements. The idea is roughly that *A* is fully rational if she has satisfied all of her local rational requirements. Alternatively, *R* is a local rational requirement for *A*, only if *A* is not fully rational unless she satisfies *R*. While such a relationship may hold between local and global rational requirements, none is assumed here.[8]

2. Consistency and unity

To understand why the *enkratic* principle does not fit easily with other putative synchronic rational requirements, it will be helpful to consider two ways in which those other requirements count as consistency requirements. The first way is what I shall call 'strict consistency'. A requirement that can be explained by appeal to the consistency of the contents of the mental states that it governs, or by appeal to closure on the contents of those mental states, is a requirement that is explained by appeal to strict consistency. The second way is what I shall call 'unity', which I shall discuss in §3. This includes requirements that are not explained by appeals to strict consistency, but which might nonetheless be a type of requirement we mean to include in developing a consistency based theory of rationality.[9]

At least some practical and certainly many theoretical rational requirements may be apt for explanation or justication in terms of ensuring strict consistency amongst the contents of the relevant subsets of an agent's mental states. We can begin by considering the theoretical rational requirement not to hold contradictory beliefs (given here in its conditional form):

---

[8]   Some of the complexities of working out the relationship between local and global rational requirements are discussed by Cherniak (1986) and Fink (2012).

[9]   I am taking it as an assumption that theories of rationality that might be said to pick up on a certain kind of psychological consistency are not presumed to be interested *per se* solely in strict logical or semantic consistency. If that were the presumption, then all rational requirements would have to be semantic consequences of the semantics of 'rationally required' and the attitudinal terms (and their contents). See Wedgwood (2007) for a discussion of this approach.

T1.   You are rationally required that (if you believe p, then you do not believe not p).[10]
TF1. RR(B$p \rightarrow \neg$B$\neg p$)

This requirement preserves the consistency of the contents of a subset of an agent's beliefs. Consistency requirements can plausibly be generated for larger complexes of an agent's beliefs:

T2.   You are rationally required that (if you believe x and you believe if x then y, then you do not believe not y).
TF2. RR{[Bx & B($x \rightarrow y$)] $\rightarrow \neg$B$\neg y$}

This requirement ensures that the contents of a particular subset of an agent's beliefs are not logically inconsistent. It can be strengthened so that the contents satisfy *local closure*:[11]

T2a. You are rationally required that (if you believe x and you believe if x then y, then you believe y).
TF2a. RR{[Bx & B($x \rightarrow y$)] $\rightarrow$ By}

Closure requirements entail consistency requirements, but not *vice versa*. In both cases, one may explain the requirements by appealing in a general way to rationality's being concerned with the conformity of the contents of one's mental states to logic. Such are strict consistency requirements, whether they are grounded in consistency in the strictest sense or in closure.

---

[10]   This may be an idealised form of the requirements. Concerns about implausibility due to demandingness on cognitive resources lead many authors to add additional relevancy or interest constraints. These constraints are orthogonal to the discussion at hand and have been omitted to avoid complicating the requirements unhelpfully. For discussion of these constraints, see Broome (2013), Fink (2012), and Kolodny (2005).

[11]   I am sceptical about closure requirements, at least for beliefs, because of concerns about demandingness. If there are closure requirements, I assume that they must be *local closure* requirements, which operate in a controlled way on the relevant subset of an agent's mental states. Spelling out how these requirements are localised is difficult to a degree that leads me to remain sceptical about their correctness. TR2a requires conformity only to *modus ponens* and is at least a candidate for being appropriately demanding.

At least some practical rational requirements are, or at least may be, explicable by an appeal to strict consistency. One example is the practical analogue of TF1:

P1.   You are rationally required that (if you intend to φ, then you do not intend not to φ).
PF1. $RR(I\varphi \rightarrow \neg I\neg\varphi)$

Another is the instrumental principle, which in the simplified form presented here is an analogue of the theoretical *modus ponens* requirement:

P2.   You are rationally required that (if you intend to φ and believe that φ only if $\downarrow$, then you intend to $\downarrow$).[12]
PF2. $RR\{[I\varphi \ \& \ B(\varphi \rightarrow \downarrow)] \rightarrow I\downarrow\}$

Although there are both beliefs and intentions in P2, the core requirement remains that their contents conform to *modus ponens*. It is important to note that recent work on the instrumental principle adds many constraints to PF2. For example, the believed necessary means must be an action that you take to be within your power to effect. It must also be an action, the results of which you do not believe will occur without your doing it. Indeed, other constraints may be appropriate, but once they are satisfied, it is relations amongst the contents of the states that ground the instrumental principle.[13]

There may also be probabilistic versions of both the practical and the theoretical rational requirements in the preceding examples, but I shall not attempt to set them out here. If there are genuine probabilistic requirements of rationality, then the underlying consistency norm will come from probability theory rather than classical logic.

Both the practical and the theoretical requirements set out above are justified or explained by an appeal to strict consistency (or some version of a closure principle). It is an implicit assumption of the schemata above that if any of the contents of the mental states in the requirements are norma-

---

[12]   Broome offers an elaborated form of this principle with many more constraints in place. See Broome (2013).

[13]   This is discussed in Broome's earlier work on normative requirements. See Broome (2002).

tive, the fact that they are normative does no special work and need not be reflected explicitly in the formulation.

It is worth noting that, as with alethic modal formalisms, we give an explicit formal representation of deontic modal terms, because we think they do some load-bearing work and should be tracked in a way distinctive from their non-modal counterparts. This observation is at once obvious and important, when we come to evaluate putative rational requirements.

To make the point about alethic modals vivid, consider two uncontroversial claims:

M1: $\Box P \rightarrow P$

and

M2: $P \rightarrow \Diamond P$

I shall, idiosyncratically, call M1 and M2 'modal bridge principles'. By this I only mean that these are valid inferences from or to alethic modal propositions from or to non-modal propositions. The non-identity of the modal claims with the non-modal claims can be seen easily in that we deny M3 and M4, whereas removing the modal operators would yield tautologies:

M3: $P \rightarrow \Box P$

and

M4: $\Diamond P \rightarrow P$

Out of such modal bridge principles, we can build modal bridge rational requirements. One example will suffice:

MR1: $RR(B\Box p \rightarrow \neg B\neg p)$

I believe that this is a plausible rational requirement. A competent grasp of 'necessarily' requires this inference. It is logically[14] inconsistent to believe that necessarily p while also believing not p. The alethic modal rational requirement is generated the same way non-modal consistency requirements

---

[14]   It may be preferable to say 'semantically inconsistent', if modal logic is understood as giving a semantics for 'necessity'.

are: by, in effect, grafting mental state operators onto the logic of a set of contents.

No such straightforward approach can be adopted in the case of deontic modals. Consider deontic versions of M1 and M2:

D1: $OP \rightarrow P$

and

D2: $P \rightarrow \neg O \neg P$

Neither D1 nor D2 are valid inferences. Consequently, we cannot explain any rational requirement merely by grafting appropriate mental states onto this set of contents and its logic. Importantly, D1 shares the contents of the mental states in EF2. If EF2 is to be counted a rational requirement, where rational requirements are broadly understood as consistency requirements, it will have to be on account of something other than what I have called 'strict consistency'.

There are no doubt many possible strategies for brining deontic modal inferences under the scope of rational requirements. One way, of course, is by intuition. Intuitively, many philosophers think, EP is a rational requirement. That may be the right way to do things. If, however, we want an algorithmic approach, and conditional on our having started with strict consistency as a criterion, we should look for the most conservative expansion of the admission criteria that we can. Excessive broadening of the admission criteria grows ever closer to intuitionism. I suggest something called 'unity' as a conservative, and not *prima facie* implausible, expansion on strict consistency.

3. Unity[15]

In this section, I shall look at a class of putative rational requirements that I shall call 'matching attitude requirements'. I shall argue that they are best understood as being grounded on the basis of *unity*, and that unity at

---

[15]   Much of the thought in this section originates from a discussion with Louis deRosset, who suggested to me looking at matching attitude requirements as a way of explicating EP.

least may be plausibly counted as part of what a consistency-based system of rational requirements is in fact trying to capture.

Matching attitude requirements (MARs) take the following form:

MAR: You are rationally required that: if you believe that you ought to have attitude A towards contents C, then you have attitude A towards contents C.

MARF: RR(BOAc → Ac)

For example, it is rationally required of you that if you believe that you ought to fear the hungry tiger, then you fear the hungry tiger. These are matching attitude requirements, because the attitude within the scope of the ought is the same as the attitude that appears without a normative operator in the consequent.[16]

In the context of this paper, the question to ask about matching attitude requirements is why they are (or might be) requirements of rationality. The answer, I propose, is what I call 'unity'. This is perhaps not the best name for the phenomenon, but I lack a better one for it.

As agents who have the capacity to be responsive to perceived norms, we are able to couch our thought in both explicitly normative and explicitly descriptive terms. Thus, I can attribute certain descriptive states to myself: that I am afraid, for example. I can also make an evaluation about whether I am in fact as I ought to be – in this case, whether I ought to be afraid. Such judgements are, naturally, fallible, but without privileged access to the book of *oughts*, agents must rely on their self-ascriptions of normative requirements to decide whether they are in a correct or a defective state.

Consistency between one's normative beliefs, which are a form of normative self-ascription, and one's related attitudes constitutes a rationally successful unification of an agent's psychology. To make this point somewhat less gestural, consider an agent who sincerely utters, 'I believe that I ought to fear the tiger, but I do not fear the tiger'. It would be natural to hear such an utterance as an admission of a defect. It implies that by the agent's own lights, she is not as she ought to be.

It is not difficult to understand what an agent is saying in cases like this, and there is no logical inconsistency. The worry instead is of a kind of

---

[16] Brunero (2013) discusses the impact of adopting something like MARs requirements. As he notes, MARs also follow from Kolodny's 'c+' and 'c-' requirements. See Kolodny (2005).

agential disunity. An agent's self-evaluation of correctness and her actual states do not match up. She has failed with respect to being guided by her normative beliefs, and thus she exhibits disunity in her theoretical reasoning about attitudes and in her actual state of being.

Disunity is a kind of inconsistency, although not the logical sort. I have no special argument that theorists who think that rational requirements stem from strict consistency must accept that some rational requirements also stem from considerations of unity. Indeed, I am not committed to the view that they must. Instead, I want again to emphasise that adding unity as a source of rational requirements is a conservative expansion on strict consistency. As far is it goes, it is a *prima facie* plausible and attractive way of generating bridge requirements between beliefs about how one ought to be and how one in fact is. Strict consistency alone secures for an agent the possibility of valid reasoning.[17] If one may employ, for example, two contradictory beliefs in one's theoretical reasoning, then one's reasoning could permissibly go anywhere. Two contradictory beliefs about necessary means to a given end would prevent instrumental practical reasoning from issuing the appropriate intention. Strict consistency based rational requirements, when conformed to, partially secure the possibility of a central agential activity: good reasoning.[18]

Unity serves something of the same kind of function. Agents need not merely act according to their inclinations, but instead are capable of reflective reasoning about what they ought to do. This kind of self-regarding normative reasoning involves a fallible self-ascription of a normative requirement, and failure to conform to it involves the implicit the self-ascription of a defect. Thus, the man who believes that he ought to fear the tiger, but does not, implicitly regards himself as defective, as do we. It is tempting to classify this kind of defect as a rational error, because while not strictly paradoxical, there is a sense of tension between theoretical

---

[17]   And it only does so in quite a limited way, by avoiding reasoning that leads to logically invalid sets of mental states.

[18]   It takes much more than consistency to fully secure the possibility of (good) reasoning. See Bratman (1999), Broome (2013), and Reisner (2009) for more discussion of reasoning.

commitment to a normative requirement and an attitudinal failure to make good on that normative commitment.[19]

Some philosophers worry about MARs. John Broome, for example, rejects them, as does Derek Parfit.[20] If we reject MARs, and unity, which generates them, as a consequence, then there is little hope for a non-*ad hoc* consistency based justification for EP in its current standard form, EF2. However, as unity is sufficient for generating a requirement that is similar to EF2,[21] it is worth considering – however briefly – why MARs may not be problematic as rational requirements. As a further point, although it can only be gestured at in passing here, it may prove difficult to argue against MARs without raising doubts about EF2 for reasons independent of one's account of consistency.

Worries about MARs typically arise because of the (debatably) non-pukka ways that one can form beliefs about which attitudes one ought to have. Suppose one accepts that there are state-given reasons for propositional attitudes, *i.e.* reasons for having a particular attitude that arise from incentives to have that attitude. State-given reasons stand in opposition to object-given reasons, which arise from the conceptual relation between an attitude type and its contents.[22] For example, suppose that Robert has been told by a billionaire that she will give him half of her fortune, if he fears that he will be eaten alive by a budgie. Robert consequently forms the belief that he ought to fear that he will be eaten alive by a budgie. It seems peculiar to think that on this basis, Robert has failed a requirement of ra-

---

[19]   Matching attitudinal requirements are wide-scope. The language in this section suggests some kind of priority is being placed on the belief that one ought to have a particular attitude towards C, but that is not in fact the case. One ceases to be disunified if one stops having the normative belief, and thus the wide-scope requirement may be satisfied either by negating the antecedent or affirming the consequent.

[20]   See Broome (2013) and Parfit (2011).

[21]   That is RR(BOIφ → Iφ). See section §5.

[22]   It has proven difficult to say just what an object-given reason is or just what relation obtains between an attitude type and its contents. A schematic way of understanding the relation is that it is the one of *fittingness*. If it is fitting for A to desire p, then A has an object-given reason to desire p. For a useful discussion, see Danielsson and Olson (2007) and also Piller (2006).

tionality by not fearing that he will be eaten alive by a small songbird, as that is an intuitively irrational seeming fear.[23]

Although the intuition in this example deserves respect, it does not tell against MARs *per se*. Like other rational requirements, MARs may be subject to certain restrictions. For example, consider any rational requirement, the consequent of which is an intention. For these requirements, one possible restriction is that one must at least not believe that one cannot carry out the intended action. And it is plausible that in the case of the instrumental principle, the requirement only applies to believed necessary means that an agent does not believe will obtain unless the agent takes action to cause them to obtain. For MARs, there may be a restriction on the basis upon which an agent comes to believe that she ought to have the relevant attitude. For example, a restriction could be added to the effect that if the agent believes that she ought to fear something solely because she has incentives to do so, then the requirement does not apply.

I am, nonetheless, sceptical that one need draw on such restrictions to defend the plausibility of MARs. When one has come to believe that one ought to fear something and does not or cannot fear it, there are several possible responses. I shall discuss two in particular.

One response is for the agent to reconsider whether she really ought to believe that she ought to fear *x*. There are requirements of theoretical rationality governing beliefs about which attitudes one ought to have and towards what. If state-given reasons are not reasons at all, or if they should not feature in grounding beliefs about which attitudes one ought to have, then what is rationally problematic about Robert is that he has come to believe at all that he ought to fear being eaten alive by a budgie. Nonetheless, given that he has that belief, the wide-scope MAR applies to him.

A parallel point may be made about EF2. One may have formed a belief about what one ought to do on insufficient or outright problematic grounds. An agent's background normative theory might be wrong, or even incoherent. Likewise, she may have systematically skewed empirical beliefs.

---

[23]   What ways of forming beliefs are pukka and which ones are not may depend on the background theory about reasons for having propositional attitudes. Morauta (2010) defends an entirely state-given reason account of reasons for intending and Booth (2012) does the same for belief. If views like Morauta's or Booth's are correct, then presumably there would be no rational defect evidenced by forming beliefs about what one ought to intend or believe in accordance with one's beliefs about one's state-given reasons for doing so.

The fact that one can arrive at beliefs about what one ought to do in strange ways is not a mark against EF2,[24] and it is not clear why it should be a mark against MARs.

Perhaps the difference between EF2 and MARs is supposed to be that, whereas one can voluntarily form an intention to do something that one believes that one ought to do, affective attitudes like fearing seem to arise spontaneously from one's beliefs and are not voluntary. Putting aside the question of whether one can will one's emotions or affective attitudes, it is certainly evident that there are indirect means of causing oneself to have certain attitudes. A hypnotist might be able to help Robert develop the fear that he will be eaten alive by budgies, or perhaps watching *The Birds* would be sufficient to have the same effect.

This brings out the second way of responding to worries about MARs. Wide-scope rational requirements govern material conditionals. A MAR would be satisfied if one believed that one ought to have attitude *A* towards contents *C* and also had attitude *A* towards content *C*. Unlike requirements of reasoning, which govern dynamic processes in which each sequential change in one's attitudes is grounded in one's prior attitudes, synchronic wide-scope requirements just specify acceptable combinations of mental states. Thus, while it might be difficult to fear being eaten alive by a budgie,[25] the fact that there is no immediate step in reasoning from believing that one ought to fear being eaten alive by a budgie to fearing being eaten alive by a budgie does not rule out the applicable MAR.[26]

And again, there is a parallel point to make for EF2. Sometimes one psychologically cannot form an intention to do what one believes that one ought to do. There may be many reasons for this. Jill may detest desiccated coconut, but she may also believe that her rare disease requires her to eat it in order for her to stay healthy. For her, too, it may take hypnosis, or a trip to the hospital to see what happens to those who do not eat their prescribed serving of desiccated coconut, to cause her to form the intention.

---

[24]   See Broome (2013) and Wedgwood (2007) for more discussion on this point.

[25]   Or it might not be difficult to fear it. Perhaps a childhood trauma has led Robert to a lifelong fear that he will be eaten alive by budgies, and this happily eases his compliance with the MAR in question.

[26]   Niko Kolodny disagrees. He argues that requirements of rationality and reasoning processes are conceptually linked. See Kolodny (2005; 2007). There is not space to take up his worries here. I argue against his view in Reisner (2009a).

Rather than the belief that she ought to eat desiccated coconut, it may be fear in the latter case that causes her to intend to eat it and hypnosis in the former, but she is still in compliance with EF2.[27]

4. Unity and the enkratic principle

On the surface, unity looks like a promising explanation for EF2. It matches an action-related attitude, intention, with a normative belief about actions. However, in no strict sense is EF2 a matching attitude requirement. What we would need to explain EF2 would be something like unity and a rule that allowed unity to apply to requirements like EF2. In this section, I shall argue that such a rule looks *ad hoc*, although I shall say something about why it may in fact not be. More straightforwardly, I shall argue there is a nearby alternative to EF2 that can be explained on the basis of unity.

MARs match a normative belief about an attitude with the having of that attitude. Here is the general schema again:

MAR: You are rationally required that: if you believe that you ought to have attitude A towards contents C, then you have attitude A towards contents C.

MARF: $RR(BOAc \rightarrow Ac)$

We can see that the same attitude appears both within the scope of the belief in the antecedent and on its own in the consequent. That EF2 is not a MAR is easy to see, because there is an intention in the consequent that does not appear within the scope of the belief in the antecedent:

EF2. $RR(BO\varphi \rightarrow I\varphi)$

Something closer to a MAR could come in two forms:

---

[27] An option that remains open to someone who wants to argue against MARs and for EF2 is to note that intentions may sometimes be formed voluntarily and directly, whereas other propositional attitudes may not. I am sceptical about arguments in rationality and normativity that rely on the difference between certain attitudes being voluntary and others not so. For more discussion in the specific case of belief, see Reisner (2009b) and Reisner (2013).

EF2a. $RR(BO\varphi \rightarrow \varphi)$

and

EF2b. $RR(BOI\varphi \rightarrow I\varphi)$

We may reject EF2a as a rational requirement on the view, assumed in much of the literature about this style of rational requirement, that rationality supervenes on the mental.[28] One may fail to perform an action for reasons having nothing to do with any mental failures. As an example, we may imagine someone who mistakenly, but consistently, believes that there are unicorns, and he is deceived into believing that a particular horse at the local stable with a narwhale horn glued to its forehead is in fact a unicorn. For whatever reasons, he forms the belief that he ought to ride a unicorn, and soon he canters off on the corned horse. The person in this example has, of course, failed to do what he believed that he ought to do, but this is not (or at least need not be) through any defect in his mental life. There simply are, unbeknownst to him, no unicorns, and thus he will certainly fail to do what he believes that he ought to. The failure here is not one of rationality, but of knowledge.

Since on the view considered here it is the price of entry for being a rational requirement that the *relata* of the requirement relation are mental states, only EF2 and EF2b are possible rational requirements. EF2 and EF2b are not logically inconsistent with each other. Whether EF2, EF2b, or both are rational requirements depends on whether we accept unity and certain restrictions thereon.

Let us consider a simple case for EF2's being the right version of the *enkratic* principle.[29] The story might go something like this. Intentions, like beliefs, are distinctive when we engage in reasoning about them in that they are transparent. They are transparent in the sense that first person reasoning about what to believe and what to intend has the character of being reasoning about the contents of the belief or the intention; ex-

---

[28]   As Christian Piller reminded me, the view that rational requirements also govern actions has a number of considerations in favour of it. See Anscombe (1957) for the modern *locus classicus* of this view. I am nonetheless persuaded that rationality does not govern actions.

[29]   A much more elaborate case is made in Broome (2013).

plicit mention of the attitude itself is often otiose in the context of reasoning.[30]

Consider an example of reasoning that issues a belief and an example of reasoning that issues an action. I wonder whether it is raining out. I look out my window and see that the ground is wet and that ripples are forming in the puddles on the pavement. I consider that wet ground and ripples in puddles normally occur as a consequence of its raining out, and I say to myself 'It's raining'. At no point in the reasoning process do I use 'belief' or make any oblique reference to it. I just consider what is the case. In the case of intending, I might wonder whether I ought to go to the shops today, or if I ought to wait until tomorrow. I consider that the traffic will be more manageable tomorrow, once the weekend has started, so I say to myself 'I shall go to the shops tomorrow' or perhaps 'I ought to go to the shops tomorrow'. Here again, I need not make mention of my intention during any part of the reasoning process.

By way of contrast, when I consider whether to fear the tiger, I must either mention or make an oblique reference to 'fearing'. I consider that the tiger is likely to eat me if it is hungry, and that I cannot outrun it. I thus say to myself 'I ought to fear the tiger'. Substitute other attitudinal terms, and the story will be the same, whether for desiring, admiring, wishing, etc. Yet, reasoning about what to do does not directly yield an action, it yields an intention. Similarly, reasoning about what is the case does not yield some event in the world such that the conclusion of my reasoning must take hold.[31] Rather it results in a belief.

This observation about reasoning might be used to support EF2. We could adopt a rule that says something to the effect that when a particular kind of reasoning has a transparent conclusion, then the most closely associated rational requirement, if there is one at all, should take a belief or an intention as the *relatum* that has the transparently presented content. This rule might be used to explain why EP, specifically in the form of EF2, takes a different form to that of a MAR, but is still a rational requirement.

---

[30]   See Hieronymi (2005), Shah and Velleman (2006), and Broome (2013) for more on the transparency of belief and intention.

[31]   To put this point another way, just because I represent the world to myself as being thus and so, it does not result in the world's being thus and so. Representing it that way just is having the belief.

We may again notice that EF2 is not a MAR because of its antecedent. The normative belief is about what to do (strictly speaking, it is about a first-person action proposition); it is not about what to intend. Thus, unity does not straightforwardly apply. An application of something like the unity principle for generating rational requirements would yield EF2a:

EF2a. $RR(BO\varphi \rightarrow \varphi)$

Of course, EF2a is not a rational requirement, insofar as we accept, as I do, that rational requirements exclusively take mental states as their *relata*. The suggested rule for dealing with cases like these attempts to resolve this problem. The implied line of argument is something to the effect that when we form normative beliefs about what to do, we would like there to be a rational requirement connecting the normative belief and our satisfying it, just as there is for attitudes that must be explicitly mentioned when we reason about them. Given that this is not possible due to the mentalistic restriction on rationality, we should adopt a kind of saving rule that picks up the nearest action related mental state, *i.e.* an intention.

This line of reasoning, however, is *ad hoc*, and it is essentially a concession to our having to decide what is and is not a rational requirement, in marginal cases, by our intuitions. Indeed, beyond being *ad hoc*, it is *post hoc*. Adding the transparency rule in this context is a way of arriving at a particular intuitive outcome: having EF2 as a rational requirement under a consistency-based theory of rationality. While it is a rule, its connection to the notion of consistency or unity is at best unclear. Adopting transparency suggests the method of slackening our understanding of consistency until we get all the requirements from our theory that we want. This method is a more appropriate response to an inquiry with different starting assumptions: one that assumes that EF2 is a consistency-based rational requirement and that consistency must be understood in such a way as to generate EF2.

To get something like EP from unity, we would need a MAR, and we have one in the form of EF2b. EF2b inserts an intention within the scope of the normative belief:

EF2b. $RR(BOI\varphi \rightarrow I\varphi)$

EF2b is a MAR, and someone who were to fail to conform to it would have the kind of self-ascription failure that unity describes. Because the *re-*

latum in the antecedent has an attitude within the scope of the ought, the requirement would be exactly parallel to the fear requirement. Nonetheless, we rarely see a requirement of this kind proposed.[32] In the next section, I shall consider whether EF2b is a rational requirement, while also considering the view that EF2, the classic form of the *enkratic* principle, is a requirement of a broader agential kind.

Before doing so, however, it is important to point out that under some restrictions, EF2b collapses into EF2.[33] Let us restrict the kind of reasons that feature in the belief that one ought to φ to object-given reasons, *i.e.* those reasons that depend on the goodness of the intended action. This ensures that all reasons to intend are also reasons to do. Whenever you believe you ought to φ, you have the same grounds for believing you ought to intend to φ, as there are exactly the same reasons for both, and necessarily so. It would therefore be a self-ascription error to believe you ought to φ and also fail to intend to φ. Scanlon's view requires a stronger claim, namely that what it is to believe that you ought to intend to φ just is to believe that you ought to φ. This would make EF2b strictly equivalent to EF2. The outstanding question is whether there is a good reason to restrict the kind of reasons that feature in the belief that one ought to φ to object-given reasons. I shall set this point aside briefly, before returning to it later in the next section. For now, it is interesting to note that there is one potentially non-*ad hoc* argumentative strategy for getting EF2 from unity: get EF2b as a MAR, and then conceptually restrict reasons for intending to φ just to being reasons to φ.

5. A rational requirement and an agential requirement?

In this section I want to consider two claims that may appear to be in tension with each other. The first is that there is a rational requirement in the neighbourhood of EF2, that is, in the neighbourhood of the classic *enkratic* principle, but that it is in fact EF2b. The second is that there is a good reason why we are interested in EF2, and that it is an important re-

---

[32]   See Marauta (2010) for a good discussion of the norms of intention that might support such a requirement. See Booth (2012) for a related discussion of belief.

[33]   I thank John Broome for directing me towards this point, originally made in Scanlon (forthcoming).

quirement, although not a requirement of rationality. I consider the effects of adopting Scanlon's view that EF2 and EF2b are equivalent, which solves one puzzle and raises a second. I then consider a puzzle that arises, if they are not equivalent. Jointly, they should perhaps make us think about positing a second category of requirements: agential requirements.

Since I am proposing that we at least consider a kind of requirement-type multiplication project, it is a good idea to defend the need to multiply. It is important to explain why it is theoretically desirable, or more likely to be true, that EF2b is part of the pantheon of rational requirements, while EF2 belongs, or might belong, in a new category of requirements. Let us consider first the case for including EF2b amongst the rational requirements.

Suppose that I have the following beliefs: BOIφ & B¬Oφ. These beliefs are not logically inconsistent, and they do not entail any very strong conclusions on their own. Consider a toxin puzzle case.[34] I shall receive a reward for having a certain intention, but performing the action will be bad for me to a limited extent. One reasonable enough reaction to such a case is the thought that I ought to have the intention, but that there is no particular normative requirement that I carry out the action. This is consistent with my believing that it is permissible for me to carry out the action, so there is no general difficulty with my ability to form an intention, when I believe that it is not the case that I ought to carry out the corresponding action.

In such a case, I would be guilty of a self-ascription error, of violating unity, if I were not to form the intention to φ. EF2 does not account for this case successfully,[35] but barring strong views about there being no state-given reasons for propositional attitudes,[36] it is difficult to see why having beliefs about what one ought to intend should not be just like having beliefs about what one ought to fear. Namely, by an agent's own lights, there is something wrong with her, if she fails to have the intention that she believes she ought to have. The only difference between EF2b and the fear

---

[34] See Kavka (1983). In the toxin puzzle, you receive a prize for intending at a particular time, say midnight tonight, to drink a noxious, but otherwise safe, liquid the next day. You can keep the prize if you have the intention, even if you do not carry out the action.

[35] More specifically, EF2 is not violated because its antecedent is false.

[36] Indeed, this seems like the wrong place to be fighting such battles. I have defended the view elsewhere that there are state-given reasons for propositional attitudes. See Reisner (2009b). More directly on this topic, Morauta (2010) provides a compelling argument that all reasons for action are at bottom state-given reasons to intend.

requirement is that the *A*-place in the generic MAR schema is filled by an intention rather than a fear.

It is not at all clear that we are being parsimonious, *vis-à-vis* the total number of requirements, by leaving EF2 as a rational requirement. This is because we would still need a further rational requirement, namely EF2b, that accounts for this reaction to the toxin puzzle. We would then have two very similar rational requirements: EF2 and EF2b.

Now consider a stronger reaction to the toxin puzzle, one which at least some philosophers would endorse. One ought to intend to drink the toxin, but one ought not to drink it. We can attribute to someone with this view, who finds herself with a toxin offer, the following beliefs: BOIφ & BO¬φ. If we accept EF2b and EF2, we get two competing rational requirements:

R1.   RR(BOI$\varphi \rightarrow$ Iφ)

and

R2.   RR(BO¬$\varphi \rightarrow$ I¬φ)

And let us assume a third rational requirement:

R3.   RR¬(Iφ & I¬φ)

These three rational requirements cannot be jointly satisfied unless one gives up the belief in the antecedent of R1 or the belief in the antecedent of R2. But, it is unclear what is wrong with holding both of those beliefs. While certain substantive views about how to address the toxin puzzle or about there being no state-given reasons for propositional attitudes might tell against holding both beliefs jointly, those are substantive theoretical commitments that one would need to accept to do the job at hand. As I, for one, do not accept them, it seems unreasonable to me to use them as a basis for ruling out the rational permissibility of holding both the belief that one ought to intend to φ and that one ought not to φ (see Reisner 2009a).

As we are investigating a notion of rationality that is underpinned by consistency, we ought not to have rational requirements that cannot be jointly consistently satisfied without having to give up two very plausible beliefs that are logically and conceptually consistent with each other:[37] be-

---

[37]   This point could be fortified by claiming that someone who takes this stronger line on the toxin puzzle is rationally required to believe that he ought to intend to drink the

liefs about what one ought to intend, and do, in reaction to being pre-
sented with a toxin puzzle-like case.[38]

Thus, we will be forced to remove one of R1-R3 from the list of ra-
tional requirements. R3 is explained by strict consistency, and it should not
be abandoned. R1, which is an instance of EF2b, is explained by unity.
R2/EF2 is not *prima facie* well explained. In this circumstance, it is R2/EF2
that has the least firm grip on its position as a requirement of rationality. If
it is a requirement at all, it looks to be a requirement of a different kind.

This is a straightforward thought that derives from the methodology
being considered in this paper. We develop a set of criteria for admitting
requirements to the pantheon of rational requirements, and EF2b meets it,
while EF2 does not. An initial thought is that insofar as we adopt this me-
thodology, we should be confident in saying that R2/EF2 is not a rational
requirement. Instead, we should move it to a new category: agential re-
quirements.[39] Agential requirements include all the rational requirements,
but not all agential requirements are rational requirements.

However, on reflection, the toxin puzzle should lead us to proceed cau-
tiously with recategorisation, at least based on the argument just given.
This caution is due to concerns arising from action theory. One important
aspect of the toxin puzzle is the question of whether or not one can win
the prize on offer at all. The subject of the bet, call her 'Nina', knows that
one need not drink the toxin to win the prize. Because of the time gap be-
tween when Nina is checked for having the relevant intention and when
Nina would need to drink the toxin to carry out the intention, Nina can
predict that she will not drink the toxin, there being no point to doing so.
Many action theorists believe that not believing you will not do something
is necessary for intending to do it,[40] and assuming that Nina is attentive to

---

toxin, and that he is rationally required to believe that he ought not to drink the toxin.
If this picture is correct, and I am unsure of whether it is, then the requirements of ra-
tionality would be formally inconsistent. For more discussion of the joint consistency of
rational requirements, see Reisner (2009a).

[38]   Note the strong similarity to Brunero's case of rational *akrasia*. See Brunero (2013).

[39]   John Broome has pointed out to me difficulties with this approach that arise from
considering the action-theoretic aspect of the toxin puzzle.

[40]   See Bratman (1999) for an in depth discussion of the role of beliefs about what one
can, and will, do and how they restrict intention.

her situation, it appears that she will not be able to win the prize. She will not believe that she will drink the toxin, and thus will not intend to.[41]

The severity of this problem hinges on there being a conceptual link between intending to φ and believing that it is not the case that one will not φ. Clearly, if this is the right view about intending and believing, it is a substantial concern for the puzzle set up by combining R1-R3, because the setup of the puzzle appears to make the MAR, EF2b, problematic, rather than EF2. The action-theoretic aspect of the toxin puzzle looks like the best argument for Scanlon's restriction on the contents of the normative belief in R1 and R2. We can get the prize in toxin cases only if we take our reasons to intend to be the same as our reasons to do.

If all these considerations add up, then we simply cannot makes sense of R1 being a different requirement to R2, and there is no inconsistency.[42] However, it is difficult to tell, at least if we assume that there are state-given reasons, whether all the considerations do add up. Take the individual belief that there are incentives for me to intend to φ. Suppose I have no view about whether I ought to φ, because I have not thought about it. I could be subject to R1 without the toxin problem arising.

I shall not try to address this problem further here. If Scanlon's restriction holds, perhaps due to toxin style worries, then R1 and R2 can never come into conflict. If we do not accept Scanlon's restriction, then they can conflict. What should be noted in either case is that the truth in action theory and the theory of reasons matters for how we resolve the conflict amongst R1, R2, and R3.

If R1 and R2 do conflict, then it may make sense to move R2 into a new sphere of requirements, those of agency. Of course, agency is a fuzzy notion. At minimum, it is a way of understanding human (and some non-human) beings as having the capacity to act and think on the basis of reflective normative thought. In philosophy of mind, this is commonly asso-

---

[41]   In fact, it is likely not required that Nina be attentive to her situation on many views about intentions. On such views it is constitutive of being an intention that one not believe that one will not carry out the intended action. I am sceptical about this as a constitutive requirement on intending. I may believe that I will not do as I intend to do, because something usually gets in the way of me carrying out my intentions. Nonetheless, I may still intend it, or so it seems to me.

[42]   At least if we take the equivalence of reasons/oughts to intend and reasons/oughts to do as conceptually identical, as Scanlon seems to. If they are conceptually distinct, but metaphysically identical, then EF2 does not collapse into EF2b.

ciated with there being mental-to-mental and mental-to-physical causation (see Kim 1996). Thinking can cause action, and thinking of a special sort, reaction to normative considerations, can do so in particular.

We need not make commitments of such a strong kind here, but rather make the thinner observation that agency importantly connects our thoughts about our proper place in the world with our interactions with the world. Requirements of rationality in particular reflect a distinctive pressure on rational agents to be internally consistent and responsive to such reflections as far as mental-to-mental relations are concerned. Our mental life leads to worldly action in a mediated way, through the state of intending; our judgements about what we ought to do only lead to our doing so under ordinary circumstances through the formation of intentions. To put it metaphorically, intentions are the doorstep to action.

Thus, it is important that our judgements about what we ought to do, so long as we believe that doing so is within our power, bring us to the doorstep of action – to the formation of an intention. Reasoning and thought have no other agential reach into the world but through that route.[43] Individuals who fail to connect their beliefs about what they ought to do with forming the intentions to do so are agentially deficient. Reasoning about what one ought to do stays as a theoretical, rather than a practical exercise for such individuals. Practical agency requires the connection. The classic version of the *enkratic* principle may or may not be a rational requirement under the method adopted in this paper. If it is not, it is because it cannot be explained by the distinctive appeal to consistency, understood to include unity. And, it can conflict problematically with a requirement, EF2b, that can be explained or justified in that way. It is, however, a plausible thing to require of an agent; our understanding of agency is partially rooted in the idea that normative thought can lead us to the doorstep of action. The pre-theoretical importance that we commonly assign to EF2 may be better, or at least equally well, explained by its role in setting the norms of well-functioning agency, rather than as a requirement of rationality *per se*.

---

[43] Except in deviant cases where thoughts, or their neural correlates, cause changes in ways unmediated by action. For example, consider an FMRI. One might intend that the FMRI show as active that part of the brain that lights up when one has the intention for the FMRI to show that part of the brain as lit up.

6. Conclusion

In this paper I have argued that EP, classically understood as EF2, may not be a plausible requirement of rationality under a particular, conservative method of generating rational requirements. I have offered an as yet sketchy case that it may belong to a broader class of agential requirements. There are questions that I have not answered. Part of the argument for moving EF2 to the broader class of agential requirements was based on its ability to come into conflict with EF2b in toxin puzzle cases. Whether it does so depends on contested views in the philosophy of action and on the nature of normative reasons for intending and acting. And, as I have tentatively suggested that rational requirements are also agential requirements, there is now a conflict at the agential level, rather than the rational one. How that conflict would be resolved remains to be seen.

Lest one worry that agential requirements come out of thin air, it is important to note that they ought not be very contentious, at least not to philosophers who want a consistency based notion of rationality. It is clear enough so as not to require argument that there is more to agency than consistency. At bare minimum, reasoning is certainly a part of agency, and it cannot be based on consistency alone, strict or otherwise (see Reisner 2009a). Affective responses to aesthetic considerations, artistic creativity, and our distinctive modes of interacting with others may also fall under the rubric of agency; if they do, then they, too, go far beyond consistency. To the extent that some ways of reasoning, certain responses to aesthetic considerations, and certain ways of interacting with others are agential and others are not, the features of our agency restrict how we live and how we are.[44] These restrictions are descriptive features of how agents can interact with the world. Restrictions that distinguish well-functioning from poorly functioning agency are requirements, agential requirements. We can expect as diverse a set of requirements as we might expect to find under the heading of agential requirements to conflict formally. A simple example was offered in §5 of how inconsistency might arise between EF2 and EF2b.

Agency as a source of requirements may allow for unresolved conflicts, or it may have the resources for ordering individual conflicting require-

---

[44]   Sufficient irrationality, arbitrary responses to one's environment, and radical discontinuities in one's mental life can diminish the degree to which a person is an agent and can even put her outside the scope of agency altogether.

ments, so that it is possible to be as agency requires overall. Spelling out in greater detail what agential requirements are, and whether they can conflict all-things-considered, is a project for another paper. At the same time, I think there is some reason to think that agential requirements admit of conflict. It is distinctive of rational requirements that they are derived from consistency. If they are distinctively the agential requirements that are so derived, this suggests that consistency may not be a necessary condition on the overall set of ordered agential requirements. Whether this is so remains to be seen.

Absent from this paper, too, is a discussion of the right method for determining the correct set of rational requirements. The right method may fall towards the intuitionistic end of the spectrum, or towards the algorithmic end. The right criteria may be different from consistency and unity, as far as they go, or they may be closely related to them – or even be them. While a discussion of the right method is absent, it has been an aim of this paper to show the importance of being explicit about method and criteria when developing a theory of rationality, insofar as rationality comprises a set of rational requirements. It has also been an aim of this paper to show that whether putative rational requirements are genuine rational requirements hinges not only on method and criteria, but also on exogenous considerations arising at least from action theory and the theory of reasons. In the specific case of EP, its status as a rational requirement is surprisingly contingent on what turn out to be the right method, right criteria, and right exogenous theories of action and reasons.

References

ANSCOMBE, G. E. (1957): *Intention*. Cambridge: Harvard University Press.

BOOTH, A. R. (2012): Two Reasons Why Epistemic Reasons Are Not Object-Given Reasons. *Philosophy and Phenomenological Research*, DOI: 10.1111/j.1933-1592.2012.00631.x.

BRATMAN, M. E. (1999): *Intention, Plans, and Practical Reason*. Cambridge: Cambridge University Press.

BROOME, J. (2007): Wide or Narrow Scope? *Mind* 116, No. 462, 359-370.

BROOME, J. (2013): *Rationality Through Reasoning*. Wiley-Blackwell.

BRUNERO, J. (2013): Rational Akrasia. *Organon F* 20, No. 4, 546-566.

CHERNIAK, C. (1986): *Minimal Rationality*. Cambridge: A Bradford Book, The MIT Press.

Coates, A. (2013): The Enkratic Requirement. *European Journal of Philosophy* 21, No. 2, 320-333.

Dancy, J. (1977): The Logical Conscience. *Analysis* 37, No. 2, 81-84.

Danielsson, S. – Olson, J. (2007): Brentano and the Buck-Passers. *Mind* 116, No. 463, 511-522.

Ewing, A.C. (1959): *Second Thoughts in Moral Philosophy*. London: Macmillan.

Fink, J. (2012): The Function of Normative Process-requirements. *Dialectica* 66, No. 1, 115-136.

Greenspan, P. S. (1975): Conditional Oughts and Hypothetical Imperatives. *The Journal of Philosophy* 72, No. 10, 259-276.

Hieronymi, P. (2005): The Wrong Kind of Reason. *The Journal of Philosophy* 102, No. 9, 437-457.

Kavka, G. S. (1983): The Toxin Puzzle. *Analysis* 43, No. 1, 33-36.

Kim, J. (1996): *Philosophy of Mind*. Boulder: Westview Press.

Kolodny, N. (2005): Why Be Rational? *Mind* 114, No. 455, 509-563.

Kolodny, N. (2007): How Does Coherence Matter? *Proceedings of the Aristotelian Society* 107, No. 3, 229-263.

Morauta, J. (2010): In Defence of State-based Reasons to Intend. *Pacific Philosophical Quarterly* 91, No. 2, 208-228.

Parfit, D. (2011): *On What Matters: Volume 1*. Oxford: Oxford University Press.

Reisner, A. (2009a): Unifying the Requirements of Rationality. *Philosophical Explorations* 12, No. 3, 243-260.

Reisner, A. (2009b): The Possibility of Pragmatic Reasons for Belief and the Wrong Kind of Reasons Problem. *Philosophical Studies* 145, No. 2, 257-272.

Reisner, A. (2013): Leaps of Knowledge. In: Chan, T. (ed.): *The Aim of Belief*. Oxford: Oxford University Press.

Scanlon, T. M. (forthcoming): Structural Irrationality. In: Brennan, G. – Goodin, R. – Smith, M. (eds.): *Common Minds: Essays in Honor of Philip Pettit*.

Shah, N. – Velleman, J. D. (2006): Doxastic Deliberation. *The Philosophical Review* 114, No. 4, 497-534.

Skorupski, J. (2010): *The Domain of Reasons*. Oxford: Oxford University Press.

Wedgwood, R. (2007): *The Nature of Normativity*. New York: Oxford University Press USA.

Understanding Enkratic Reasoning

OLAV GJELSVIK

Centre for the Study of Mind in Nature. IFIKK. University of Oslo
Georg Morgenstiernes hus. Blindernveinen 31. 0313 Oslo. Norway
olav.gjelsvik@ifikk.uio.no

ABSTRACT: Beginning from John Broome's approach to Enkrasia, the paper quickly moves to giving a condensed presentation of an approach to practical reasoning motivated by a Fregean approach to inference (in theoretical reasoning). The suggested account of practical reasoning avoids using rationality requirements to do explanatory work when accounting for correct reasoning, and thus avoids lots of problems. It is strictly conservative in its approach, and no new inference rules are required for moving from the theoretical to the practical case. It is suggested that we can stick to deductive reasoning when accounting for practical reasoning proper; the crucial premiss from theoretical reasoning about practical matters cannot normally be established this way. The paper moves on to tackle counterarguments to the effect that there will simply be too little correct practical reasoning on the present (deductive) approach. The simple account of correct reasoning has too high a cost, it is argued. The paper meets this objection when it argues that much reasoning is enthymematic or incomplete reasoning. By making specific claims about how there may be practical premises to which we do not attend even when they are, in some sense, before the mind, the approach is defended.

KEYWORDS: Akrasia – Broome – enkrasia – enthymematic reasoning – incomplete reasoning – inference – practical inference – weak will.

1. Introduction

The question of how the belief that you ought to do something relates or should relate to your intending to do that thing is of great interest. John

Broome's general strategy is to argue that *rationality* requires you to intend to do what you believe you ought to do, and that you can bring yourself to satisfy this requirement of rationality through *reasoning*. An important requirement of rationality in Broome's account is therefore the following, now commonly called *Enkrasia*:

> Rationality requires of you that, if you believe you ought that you F, you intend that you F.

This formulation will do for now. It exhibits a 'rational' connection between what you believe and what you intend, such that if you fail to exhibit this connection, you are not (fully) rational. What I shall now call Broome's *associated claim* is a claim to the effect that reasoning can make you satisfy this requirement. It is surely possible, I suggest, to accept the first claim, which we might call the *rationality claim*, without accepting this *associated claim*.

Broome takes a stand on various questions and endorses specific views on reasoning in making this *associated claim*. His account of reasoning has a rule of inference that identifies correct enkratic reasoning. Having this kind of rule might be seen as unmotivated and problematic, and, consequently, as unable to substantiate the *associated claim*.[1]

This paper will argue that we can accept a version of the *associated claim* without having to introduce a specific rule of inference for the enkratic case. In order to show this, I shall outline a view on correct enkratic reasoning which introduces no special rule. I shall make comparisons between my approach and Broome's, and argue in support of my view.

Actual cases of correct enkratic reasoning might be thought to be somewhat less frequent on my view than on Broome's. That is, correct enkratic reasoning might be thought to be too rare on my approach. I shall aim to show that this need not be so as long as a more substantive account of rationality is correct, and we also allow for enthymematic reasoning, as indeed Broome also does. The more general methodology of enthymematic cases will in fact favour the present approach, I shall argue. In the theoretical choice between this approach and Broome's, the former approach to reasoning and also to enthymematic reasoning should be favoured. First, let me outline the present approach to reasoning.

---

[1]   Correct reasoning can be enthymematic on Broome's account, as it also can on those of most others.

2. Background for the present approach

The approach I suggest starts from a general approach to inference. I think of an actual inference as an act. This is important in at least two ways: to see what inferences are and what they are not. What they are: Mental acts are things we engage in as persons, and should not be characterized merely by the relationship between propositional contents, which is how logicians often characterize inference. Inferences are acts by us. What they are not: I shall distinguish between acts and actions in contending that basic sorts of mental activity, like judging and inferring, are acts, not actions. Actions are processes in time, subject to agential control; judgings and inferrings are not subject to the same type of control, and are not temporal processes.

My theoretical starting point lies in current neo-Fregean approaches to inference. This sort of view can be best appreciated by considering the case of normal theoretical inference, where inferring is an act in which you move from the premises to the conclusion. Each premiss is characterized by the semantic content of a proposition, and also by the way in which we relate to that content when we judge it as true. The latter is represented by the Fregean judgment stroke, normally taken to indicate judgment or assertion. This approach, and the crucial use of the judgment stroke, has languished in some disrepute until recently when it comes to accounting for inference, but is gaining currency again. Here is Dag Prawitz's account of inference:

> An inference in the course of an argument or proof is not an assertion or judgment to the effect that a certain conclusion *B* 'follows' from a number of premises *A1, A2, ..., An,* but is first of all a transition from some assertions (or judgments) to another one. In other words, it contains the *n*+1 assertions *A1, A2, ..., An,* and *B,* and in addition, the claim that the assertion *B* is supported by the assertions *A1, A2, ..., An,* a claim commonly indicated by words like 'then', 'hence', or 'therefore'.
>
> This is how Frege saw an inference, as a transition between assertions or judgments. To make an assertion is to use a declarative sentence *A* with assertive force, which we may indicate by writing ⊢*A,* using the Fregean assertion sign. We may also say with Frege that a sentence *A* expresses a thought or proposition *p,* while ⊢*A,* the assertion of *A,* is an act in which *p* is judged to be true. (Prawitz forthcoming)

It is thus very important that we do not think of inference as ascertaining whether a relation of logical consequence holds between two expressions; we are not thinking of inference as judging that there is an entailment relation between two propositional contents, but as an act of transition from the premisses to the conclusion. And, moreover, the simplest case is the deductive case. It is therefore worth working with as long as we can.

Now the interesting thing when we think about Enkrasia is the possibilities that present themselves if we extend and generalize the Fregean picture and apply it to all inference, theoretical and practical. (Frege's concern was always demonstrative science.) Consider, then, the possibility of introducing a *practical stroke* parallel to the judgment stroke. This stroke then represents a practical way of being related to a propositional content, as opposed to the theoretical way represented by the judgment stroke. This practical way of being related to a propositional content might, in the primary case, simply be an intentional action, i.e. a case of doing something intentionally. The case of intending to do something I take as a derived case, to be explained and illuminated from the philosophically prior case of doing something intentionally.

In practical inference, on this picture, we need a relevant practical way of taking a premiss for a practical conclusion, and the conclusion might simply be an action B that follows from engaging in another action A together with a premiss of the sort "In order to do A I need to do B". If this is the picture, the pattern of reasoning might be instrumental deductive reasoning. Furthermore, it is basic to the account of reasoning and inference I am defending here that if it is true that I am practically related to p, I am theoretically/judgmentally related to p. One can, on this view, always go from premiss practical stroke p to premiss theoretical/judgment stroke p, but not the other way. (One is aware of what one is doing intentionally.) The relationship is modeled along the lines of the relationship between knowledge and belief, as that relationship has to be modeled on the knowledge first approach; it is not an inferential relationship, and believing is not an isolatable part of knowing.

We think of the way we are related to a propositional content that is signified by the judgment stroke, as characterized by the rule for assertion, as a rule we might think of as the constitutive rule for assertion. This has consequences, the first being that even if we typically move downstream when inferring q from p, the conclusion might be a proposition we subsequently might realize we cannot assert. When that happens, and the con-

clusion cannot be asserted, we have to withdraw the assertion of at least one premiss. This stems simply from the nature of assertion plus the entailment relations of the case. Now, if practical stroke p commits us to theoretical (judgment) stroke p, we get the same result in the practical case as in the theoretical case. This feature of the practical commitment to the propositional content at hand explains all there is to explain about what in the literature is discussed as the scope of rationality requirements. All the truths about the scope of those requirements, which John Broome has done so much to establish, are simply direct consequences of a Fregean approach to inference, a reasonable approach to assertion along the lines of Williamson, and a recognition of Anscombe's view that in doing something intentionally we know (are aware of) what we are doing. (This latter I see as legitimizing the move from practical stroke p to theoretical stroke p.) All these points are controversial, but enjoy nevertheless fairly widespread acceptance today.[2]

3. The relationship to Broome's current account of enkratic reasoning

Let us now take this basic picture and contrast it with John Broome's approach. Broome works out the rule for correct enkratic reasoning by starting with what he calls Enkratic Permission. Here is Broome:

Enkratic Permission
Rationality permits N that
N believes at some time that she herself ought that p, and N believes at some time that it is up to her herself whether or not p
N intends at some time that p, and N's intention that p is based on N's belief that she herself ought that p and belief that it is up to her herself whether or not p.

---

[2]   For a list of Broome's works, see the reference section below. The discussion about rationality requirements is highlighted in Broome's interaction with Kolodny (2005). I cannot here go into that interchange. My approach leaves no explanatory room for rationality requirements beyond what stem from the ways we are related to propositional contents. There is some similarity to Kolodny, but there are also differences; the ways we are related to contents in judging might change when we realize the consequences of what we judge to be true. For my references to Anscombe and Williamson, see the reference section below.

This is derived from Enkrasia. The corresponding rule is:

> From
> <I ought that *p*; belief> and
> <It is up to me whether or not *p*; belief>
> to derive
> <*p*; intention>.

Your reasoning, given below, about taking a break is correct only if it follows this rule. If it is enthymematic it might still follow this rule if the second premiss is before the mind in some sense. Fully spelt out and made explicit, correct enkratic reasoning would be:

> I ought to take a break.
> It is up to me whether or not I take a break.
> So I shall take a break.

Let me start by stating a point of agreement. I agree with Broome that we need to be able to distinguish between a practical and a theoretical way of being related to propositions in the account of correct reasoning. I accomplish this by means of the technical apparatus of strokes, the judgment stroke and the practical stroke. Broome does it with 'belief' and 'intention'.

We also agree that we need to be able to account for the correctness and incorrectness of reasoning. We differ on how we do that. Broome's view is given above by a specific rule for Enkrasia. I hold, and this stems from a generalized approach to reasoning, that we recognize correctness also by the semantic/logical relations between the contents in question, which on my view require entailment relations. We agree to a large extent about the roles of the practical/theoretical attitudes we have to the semantic contents, but I require a practical premiss to get a practical conclusion. Broome manages without that with his rule for Enkrasia, which bridges the theoretical and the practical. On this point, my view is the more conservative, as it only recognizes inference rules of the sort we find in theoretical reasoning (semantic entailment relations), and has no inference bridging the theoretical and the practical. I get back to this later.

Broome and I agree that we have uncontroversial cases of practical reasoning in instrumental practical reasoning. Think of cases where

I intend to F,
Realize (believe) that to do F I must do G,
and (via the route of practical inference)
intend to G.

One reason why this is uncontroversial between us is that we here have a parallel theoretical inference: it would also be correct to infer from

judging that I am F-ing, and also
judging that if I am F-ing I have to be G-ing, to
judging that I am G-ing.

Contrary to Broome, I think we can make use of this structure of instrumental practical reasoning for practical reasoning proper in its entirety, including enkratic reasoning. Broome takes a different approach, and formulates the rule for the correctness of the reasoning in question.

Broome considers objections to his account of reasoning. In the very last section of his 2009 manuscript "How to be Rational", an early version of his new book, but also in its present version, he considers a view which objects to his full account of enkratic reasoning by invoking the "Motivation out, Motivation in" principle, named so by Jay Wallace (see Wallace 2001). The principle boils down to saying that a practical premiss is needed for a practical conclusion, as is also the case in instrumental practical reasoning.

Here is something of what Broome says against Wallace:

If there is such a thing as enkratic reasoning, the premiss-belief does incorporate a motivation of sorts. Suppose your premiss-belief is that you ought to take a break, and suppose you are rational. Being rational, you are supposed to do enkratic reasoning. Therefore you are disposed to reason your way from your belief to an intention to take a break. This intention is itself a sort of disposition to take a break. So your belief that you ought to take a break constitutes a sort of disposition to take a break. And a disposition to take a break is a motivation to take a break.[3]

I agree with much of what Broome is saying here, but not all. I resist the view that a belief, or perhaps we should rather say my judging something to be correct, is at all a disposition to act in a *particular* way. I see the

---

[3]   Quoted from Broome's 2009 manuscript. More or less the same statement is found on page 305 in the present manuscript, first long paragraph.

essential nature of belief as given by its aim in truth (or knowledge). Of course, a belief, for instance the belief that 'snow is white', will interact with existing motivations and thus be a factor behind new motivational states and actions when such interaction in reasoning brings about new motivations. Think of a case where I intend to see something white and believe there is snow behind my house. I also believe that snow is white, and that the easiest way to get to see something white to go behind my house. I then intend to go behind my house and proceed to do so. The fact that a belief can interact with motivational states and in this way make up a factor in bringing about new motivational states does not make the belief that snow is white, or my believing that snow is white, itself into anything like a motivational state or disposition. I think this point generalizes to all belief.

There are thus two points here. There is in general no source in the nature of belief to see belief as a motivational disposition for doing particular things. Furthermore, the fact that belief in conjunction with motivational states lead to new motivational states does not show that belief is a motivational state. For that to follow, we would need to hold that there are cases of the sort we are interested in where all crucial factors behind a particular motivational state are beliefs. That is, I take it, the issue we are discussing, and that is a contested issue when it comes to beliefs about what we ought to do.

There is a caveat. The argument Broome gives is about reasoning. Of course, my believing that snow is white might simply cause me to intend to see some snow. Similarly, my belief that I ought to do F might cause me to intend to do F. Beliefs can cause intentions without this causal relationship being one of reasoning. When a belief causes an intention, and it also is rationally required that we have that intention if we have that belief, we have both a rational connection, and a causal connection. But the existence of such a causal connection does not automatically make of it a piece of reasoning. One causal route will be via reasoning, but there will also be other causal routes. There are constraints on *how* the enkratic requirement can be implemented by reasoning also on Broome's view. A causal connection is not enough in and of itself.

Note also that accepting Wallace's 'Motivation out Motivation in' principle does not in itself lead to rejecting the claim that we implement the enkratic requirement by reasoning. What it does, though, is rather to modify or constrain the bigger picture both of rationality and of reasoning, of causal connections and reasoning connections. One complication is this:

someone who accepts that rationality requires us to be enkratic might at the same time be subscribing to a more substantive and quite different conception of rationality than Broome endorses. According to the present approach, it is true that more premises 'will be required than on Broome's view to derive an intention by reasoning in the enkratic case where we believe we ought to do something. As I shall argue in a separate step below, we should think of these further premises as practical states (or intentions).

Still, it might be thought, cases of enkratic reasoning may nevertheless be very rare on this approach, though on a more substantive conception of rationality, certain standing intentions or standing actions of engaging in doing what you ought to do might in fact be required (in some sense) for you be rational. There might be many intentions to which we do not really attend. Clearly, if such 'higher-level' intentions were not in place, one would not be able to reach the conclusion required by Enkrasia by the present conception of reasoning. Broome at this point has a less substantive conception of what it is to be a rational person, and on his view enkratic reasoning, by the present standards, would be rare. Still, a different account of reasoning might also come with a more substantive account of what it is to be a rational person, and distinguish between intentions we have and intentions we attend to.

While there are clear similarities between Broome's 'intending' and my practical stroke, let me add something more about our differences. From my perspective Broome might be making a mistake by not digging more deeply into the possibly explanatory connections between doing something intentionally and intending something. If the explanatory connections go from doing something intentionally to intending something, and doing something intentionally connects with exercising specific types of rational capacities when acting in response to reasons, then that brings a conception of agency into the picture. On such a view, it is the relationship of intending to doing something intentionally that gives intending the practical character it has, and also restricts what we can intend. When seen this way, we might also be able to acknowledge the further point that the conclusion of practical reasoning might indeed be the act of doing something intentionally.[4]

---

[4]   At this point there are of course very clear differences between Broome's and my views. There are important questions about the relationship between prior intention and the general state of intending something, and how they both relate to doing something intentionally. I do not address them here. I take the explanatory prior state to be

My more substantive conception of rationality does, it seems, raise some very tricky issues about what it takes to have something like an intention, or to have 'in some sense' an intention in some case of enthymematic reasoning. Such issues need, however, to be faced by all of us. It is easy to overlook the possibility that many intentions and things we do intentionally do not have to be clearly attended to in consciousness even if we engage in them and are in some sense aware of them. It is therefore easy to overlook the possibility that the reasoning that we think is fine might be enthymematic. I shall defend the view that much practical reasoning is indeed incomplete, and that if it is widespread, then it raises many issues about self-knowledge I will not be able to address those issues in the present context, but I shall nevertheless defend the possible criticism that there will be too few satisfactory practical inferences on the present approach by arguing that it is not true; the inferences are, however, incomplete or enthymematic. We can extend this line to an answer to the view that that type of practical reasoning does not require validity or correctness; it simply provides an aspect under which the action is seen as good. This is naturally seen as enthymematic reasoning as well.

4. Enkratic reasoning on the present view

Let us then, at least for the sake of argument, consider the possibility that the actual cases Broome calls enkratic reasoning are incomplete or enthymematic pieces of reasoning, in this specific way, that there is a "hidden" further premiss. There are typically good grounds for ascribing this "hidden" (enthymematic) premiss to a person seriously engaged in judging that "I ought to do F". I shall for now also assume that this 'hidden' premiss is held in a motivational way. I will later argue for this assumption.

The case is Broome's in which you move from the recognition that you ought to take a break to an intention to take a break. 'B' stands for belief when we render Broome's approach, and stands for the judgement stroke on my approach. The 'I' stands for intention on Broome's approach, and for the practical stroke on my approach. This is done for comparative purposes; I don't think it causes any problems.

---

the latter, and the primary case of reasoning to be the case where both premisses and conclusion are factive states.

This is the practical reasoning (all simplified in various ways):

2. B (I ought to take a break)

3. I (I shall take a break)

On Broome's view this reasoning is enthymematic: I also need to believe that it is up to me whether I take a break or not. The correct reasoning is something like this:

2. B (I ought to take a break)

2*. B (It is up to me whether or not I take a break.)

3. I (I shall take a break)

On the view I am suggesting this is not correct reasoning. There is a move from the modality of 'ought' to 'shall', which is not correct reasoning in as much as the one modal predicate does not entail the other.

This is the practical reasoning on my view:

1. I (If I ought to take a break then I shall take a break)
2. B (I ought to take a break)
3. I (I shall take a break)

Let us now compare these views of what is enthymematic in these cases. Some points are obvious. In the type of enkratic reasoning I endorse, the premiss-belief does not have to incorporate any motivation of any sort. I therefore need not subscribe to the following statement by Broome, quoted above: "So your belief that you ought to take a break constitutes a sort of disposition to take a break. And a disposition to take a break is a motivation to take a break." The belief in question, the belief Broome mentions, would still be the same belief even if you had no disposition whatsoever to do what you ought to do in relation to taking a break. In that case there would not be a disposition to take a break, and no corresponding intention.

The very possibility that the first (possible) premiss has no direct phenomenological presence seems to support Broome's approach, which tries to do without such an intention among the premisses. But if the reasoning is incomplete or enthymematic, then this point does not necessarily support his approach any more than it supports my approach. That being the case the situation is left quite open, and the favoured option must be the

most theoretically satisfactory approach that fits in with a number of other philosophical commitments.

Enthymematic reasoning is tricky business. I follow Grice here, who has done interesting work, in taking enthymematic reasoning to imply that in addition to the actual argument there is also a non-actual or ideal argument that is formally valid (see Grice 2005, especially Chap. 2). The premises of this ideal argument are in some sense to be specified present to the mind of the reasoning person. The 'some sense' is the hard bit; a brute disposition to act is not enough for what we are after. That presence before the mind is furthermore, in some sense, required in order for the reasoning person is to be considered rational. I note that Broome's approach here differs from Grice's. Of course, there is great need to say more about the notion of being present to the mind in some sense, but that is another a problem we both face, indeed, we all face.

In the rest of the paper I shall develop and support my view further. In doing this, I shall invoke Grice's somewhat neglected work on incomplete reasoning. From it we learn how hard it is to specify what reasoning is and what it isn't. We can also learn this: Methodologically speaking we ought to start with paradigmatic and undisputed examples of reasoning. In these cases the activity of the reasoning person is characterized both by the formal/semantic relations between the contents entertained and also by the reasoning person's relations to such contents. And as we also will shortly see, we need a broader conception of reasoning than the actual cases of reasoning that fit the paradigmatic cases. This is particularly so in the case of incomplete reasoning.

5. Incomplete reasoning

In trying to pin down what reasoning is, we face some deep methodological issues. The prime examples involve correctly engaging in impeccable deductive or inductive structures. We need, for all cases of reasoning, an independent grasp of the correctness conditions. We normally get this from logic, from soundness proofs, etc.

Reasoning involves more than the prime examples, and we want to extend our account of reasoning, possibly to bad reasoning and definitely to incomplete reasoning. When we do so, we extend our conception of reasoning to new cases, to be understood against the background of the primary cases.

As I have explained, my worries about Broome's approach to Enkrasia have their root in worries that the account of reasoning invoked does not satisfy correctness criteria that can be seen as independent from the case at hand. It moves from and 'ought' premiss to a 'shall' conclusion, which is not a valid inference in the normal sense of valid. On the other hand, a conception of reasoning as conscious explicit reasoning is too limited. We need to include incomplete reasoning.

While incomplete reasoning is in fact very common, it is also hard to pin down. I shall give a very brief exposition of Paul Grice's account of incomplete reasoning (cf. Grice 2005).

Here is one case discussed by Grice:

Jack sustains a head injury, Jill says
He is an Englishman, so he will be brave.

Grice identifies three ways of dealing with this as a piece of reasoning, and he endorses the third. The first holds that there is a suppressed premiss (namely that 'All Englishmen are brave'). There is a particular problem with this suppressed premiss thesis: you cannot object to Jill "That does not follow". This is because it does follow, given the suppressed premiss.

The second approach sees "Englishmen are always brave" not as a premiss but as an inference rule. This second approach accepts contingent inference rules. They carry with them lots of problems. One might, after all, believe that Englishmen seldom are brave etc.

The third approach holds that there are two arguments at play, the actual and the non-actual or ideal. The ideal argument incorporates the premiss which in some sense is before Jill's mind, namely that "Englishmen are always brave", and is formally valid. Jill's actual argument is (informally) valid just in case there is a formally valid argument with premisses Jill in some sense has in mind.

Without going further into Grice's reasons for preferring the third approach, let us note there will be a range of cases where the "suppressed" premiss is not explicit because it is generally, in some sense, taken for granted by all parties. Grice has a lot to say about the pragmatics around such cases – maybe the premisses can be taken for granted because they are part of a general concept of rationality that we must see as applicable to the reasoning person.

6. Enkratic reasoning as incomplete/enthymematic

Focus then on the enkratic case, illustrated above:

2. <u>B (I ought to take a break)</u>

3. I (I shall take a break)

There are two competing enthymematic approaches. The crucial bit of reasoning on Broome's account that is not formally valid on my view is the move from one modality to another, from "ought" to "shall" or "will".[5] We seem from the perspective of this view to be missing a premiss that connects the two modal verbs, ought and shall/will in the required way. This is not difficult to supply, however. I have supplied it above, in the extra premiss. Can we think of this premiss as being present in some sense? What does it take to be present? Are there tests that can illuminate whether the premiss is present in some sense or not? To provide something like a test I will introduce predictions. But I stress that the test gives only a strong indication, not a proof, and is not a reductive account of what it is to have the premiss before the mind.

6.1. Predictions as tests

Imagine you are doing predictions about yourself or another. At the moment you are not doing practical reasoning, you are simply trying to predict whether A will do F. Stipulate that A realizes that he ought to do F. We can on that basis try to predict that A will F, but this is neither a solid nor well-founded prediction as long as it is unknown whether A is likely to do as he ought to do. The prediction would be on a substantially more secure footing if we also knew that A was set to do what he ought to do in relation to doing F or not. In that case we could make a solid prediction.

Notice that we might also be able to make this same prediction if we took this "suppressed" premiss to be an inference rule. Doing so, however, would open its own can of worms. Think of a very different situation where Akratic A knew or believed that he would not do what he ought to

---

[5]   This modal verb could possibly be understood differently from the way I understand it. I understand 'shall' to be proximate in meaning to 'will', and that we use 'shall' in the first person, 'will' in the third. 'Ought', on my view, is straightforwardly normative, the other two are not.

do in relation to F. In that case the picture would be that he did not reason according to the inference rule in question and also that the fact that he did not reason in accordance with the inference rule was known. This is a bad picture of what seems to be happening in akrasia: it is extremely hard to make sense of an agent who knowingly goes against inference rules. (Remember the view of inference as act, not action.) Weak will/akrasia is hard enough to understand without this complication.

Let us look at what is in effect a generic statement, to the effect that A normally does what he thinks he ought to do. This statement is not much help for a really solid prediction, but there is some help in it. What it is good for is what we might think of as an explanation post factum. Many explanation theorists who insist that prediction and explanation come apart, hold that much explanation is post factum.

Let us move closer in on the enkratic case. Imagine first a self-prediction along the same lines as the prediction above. The self-prediction would normally take the same form as the above, the only difference being that it is in the first person. All the same points seem to apply.

It is natural to think, as I do, that such a self-prediction and a piece of practical reasoning can involve the exact same propositional contents. There are general grounds for believing this to be the case, it is the case in instrumental practical reasoning, and it would connect the various types of reasoning in a simple picture.

To be sure, the difference between the predictive case and the practical reasoning case is that practical reasoning has an intention as a conclusion, i.e. it ends in relating to a propositional content the way we do when we relate practically, in intending or in action. That is, at least for the moment, agreed territory. What I take to have pointed to is the following.

If we are to see the predictive case and the practical reasoning case as involving the same propositional content, as it is natural to do on general grounds, then it is also natural to see the practical case given by Broome as incomplete. These issues about prediction seem to provide substance as to whether the premiss that is definitely not explicit is present in some sense or not. A claim to the effect that there is a missing premiss has real substance if its presence or absence really matters to the solidity of the predictions we can make.

Against this background we can consider various approaches to incomplete reasoning. Broome's version of the enkratic reasoning person can be seen as following Grice's second way of dealing with incomplete reasoning,

with an exception for the 'up to us-ness'. (I shall ignore that last part of it for now.) This way is not to be recommended on general grounds. It treats the enkratic requirement as if it were an inference rule, and makes what in this case is weak will into a case where we knowingly do not follow an inference rule. More generally, the approach faces the problem of finding some 'independent' way of "grounding" the presence or absence of the suppressed premiss. This makes it difficult for us to distinguish the case where we have a causal transition from the premiss(es) to the conclusion from the case where this transition is via reasoning, albeit enthymematic. We need the distinction, even if it is hard to classify actual cases by it.

If we go for Grice's third option, as I do, we hold that there are two arguments, one actual but incomplete, and one ideal but complete. The premisses of the ideal one are in some sense before the mind of the reasoning person, but not directly present in the sense of in so far as they as they are not being attended to. We provide some substance to presence or absence of the ideal argument by looking at a parallel piece of predictive reasoning about a formally identical situation. Here we find resources that can ground the ascription of the missing premiss, and provide some substance to the presence or absence of the premiss. The substance resides in how solid a prediction we can make. If the predictions are solid, we have some ground for saying that the suppressed premiss is present in some sense. We have, possibly, further ground when the premiss is something approaching an analytical truth, to the effect that you ought to do what you ought to do.

Note that we have reached this conclusion about the enkratic case without taking a stand on the motivation in, motivation out principle. It was reached on general grounds by thinking of the case as a case of incomplete reasoning, and just on the assumption that the reasoning is a formally correct piece of reasoning. A further step is required to make the case for the claim that the attitude to the premiss that is present in some sense and occurs in the ideal argument is a motivationally held premiss. I shall finally argue in further support of this step.

7. The missing premiss

By focusing on the change from one modal verb in the premiss to another modal verb in the conclusion, I have argued that in order to see the enkratic reasoning as reasoning we should see it as incomplete reason-

ing. There are various ways of approaching the issue; I suggest opting for Grice's preferred way of dealing with incomplete reasoning.

On Grice's approach there are various general explanations why certain premisses are not attended to; typically they are entirely obvious and not worth mentioning. Bearing that in mind, let us return to our starting points. Here is Broome again (as quoted above):

> Being rational, you are supposed to do enkratic reasoning. Therefore you are disposed to reason your way from your belief to an intention to take a break.

What I have been insisting on is that it is indeed possible also on my view to accept the first of these sentences, and also the second as long as we conceive of the actual reasoning in question as incomplete or as enthymematic.

The premiss, I think, is (in some sense) present when enkratic reasoning being performed is the premiss that I, in relation to doing F, shall do what I ought. Grice-type mechanisms for explaining why many things go unmentioned can be seen as carrying over to this case as well if it is a characteristic trait of rational persons that they normally do what they recognize they ought. This latter point about a generic trait might be understood in many ways, but understood rightly it might be approaching a true generic statement about the sort of rational beings humans are.

I admit to a certain type of haziness around the presence of such premisses to which this approach appeals. If we were to delve further into these presences or absences, we would meet up with some tricky issues in self-knowledge, the extent to which our own mental states are transparent to us, and the relationship between the contents of our minds and what we attend to.

At this point it might be useful to remind ourselves of typical cases of weak will. They seem to be exactly the kind of case where the true generic statement underlying the ascription of rationality has an exception. It is not true that you can be described as engaging in the full ideal argument, even if you accept the premiss that you ought to do F, in a case of weak will where you weak-willedly do something else. (You do not intend to do what you ought to do when you are Akratic.) What is the case is precisely that you do not form the intention of doing F as a rational person would. And in so far as you do not, the likely explanation is not that there is an error of reasoning, but that the premiss needed in the ideal argument to make a

conclusion out of the intention to do F, is not true of you. There is then, in your case, no valid argument to the conclusion that you do F. Weak will and akrasia reside in the weakening of a premiss a rational person ought to hold on to, not in erratic reasoning.

Broome is possibly in the neighbourhood of having us see weak will as an error of reasoning. My objection is that weak will (Akrasia) does not have the recognizable features of slips and errors of reasoning, which again are connected with the inferences being acts. Weak will is a different type of rationality failure. The error of Akrasia is that the true generic statement about rational beings as humans is matched by a momentary failure; it is not true at the moment of weak willed action that the agent forms the intention to do what he ought.

7.1. Support for Wallace's principle

Let us then again address the question of whether the attitude to the missing premiss should be motivational or not. I say it should be motivational. Here is an argument in support of that. If the premisses leading up to a conclusion are the same in two arguments, both in propositional content and attitude towards that content, then the conclusions should be the same. The cases of prediction and practical reasoning proper can have the same content but different attitudes towards the conclusions: belief in the predictive case, intending in the practical case. This difference in the conclusions must be explained by some difference in the premisses. Since the contents are the same, there has to be a difference in attitude to one premiss if the difference in conclusion is to be explained. Since there clearly are only belief attitudes towards the premisses in the predictive case, there has to be a motivational attitude to a premiss in the practical case to reach a practical conclusion. This follows as long as we accept the need to provide full accounts of validity in semantic relations, and note the difference in conclusion in the predictive case and the practical case. So here we have an argument for the 'motivation out motivation in'.

8. General conclusion

We now have a more complete picture of enkratic reasoning as normal practical reasoning, and practical reasoning simply as generalized instrumental practical reasoning. The enkratic reasoning may typically be incom-

plete reasoning. The missing premiss then has to be present to the mind in some sense in an ideal argument. It is quite natural that such a premiss is not present to consciousness or attended to if the premiss would express something general about the nature of a rational person; that is also to be expected from Gricean principles of economy etc. We can also, when we see this, provide a precise argument for the view that the 'extra' premiss has to be held motivationally. Only in that way can we account for the difference between the practical case and the predictive case.

The picture of enkratic reasoning is, then, in a certain sense, very simple, but the picture of the mind, its rationality and the rational person is not. When accounting for enkratic reasoning, we touch on deep features of the mind and its activity in the lives of our sort of rational beings. The fact that we do enkratic reasoning shows that we typically engage in doing what we ought to do. A substantive account of what it is to be a rational being is needed to accommodate this, which, it seems to me, is both right and very interesting.

If intentional actions are to be understood as ways of being related to propositions, then rationality obviously extends to them as well. If all intentional actions are ways of moving our bodies, they are also minded movements. Facts about the bigger picture, and the limitations in our access to the contents of our own minds beyond what we attend to, all come to light by thinking through the issues around Enkrasia. Enkrasia seems simple, but the problems are deep, and to get things right we need to get the big picture right.

The view here defended shows, I believe, that Enkrasia is a direct reflection of the deeper fact that rationality consists in responsivity to the reasons there are, and that intentionally doing things might be among the responses required by the reasons there are. That, however, I cannot argue here (see Gjelsvik 2007).[6]

References

ANSCOMBE, G. E. M. (1957): *Intention*. Oxford: Blackwell.

ANSCOMBE, G. E. M. (1995): Practical Inference. In: Hursthouse, R. – Lawrence, G. – Quinn, W. (eds.): *Virtues and Reasons*. Oxford: Oxford University Press, 1-34.

BROOME, J. (1999): Normative Requirements. *Ratio* 12, 398-419.

BROOME, J. (2001): Normative Practical Reasoning. *Proceedings of the Aristotelian Society* 74, No. 1, 175-193.

BROOME, J. (2003): Practical Reasoning. In: Bermudez, J. L. – Millar, A. (eds.): *Reason and Nature. Essays in the Theory of Rationality*. Oxford: Oxford University Press, 85-112.

BROOME, J. (2009): The Unity of Reasoning. In: Robertson, S. (ed.): *Spheres of Reason*. Oxford: Oxford University Press.

BROOME, J. (2009): *How to be Rational*. Circulated book manuscript, the latest version edited as Broome, J. (2013): *Rationality Through Reasoning*. Wiley-Blackwell.

BROOME, J. (2010): Rationality. In: O'Connor, T. – Sandis, C. (eds.): *A Companion to the Philosophy of Action*. Blackwell, 285-292.

GJELSVIK, O. (2007): Are there Reasons to be Rational? In: Rönnow-Rasmussen, T. et al. (eds.): *Hommage á Wlodek. A Festschrift for W. Rabinowicz*. Lund.

GRICE, P. (2005): *Aspects of Reason*. Oxford: Oxford University Press.

KOLODNY, N. (2005): Why Be Rational? *Mind* 114, 509-563.

PRAWITZ, D. (forthcoming): Explaining Deductive Inference'. In: Wansing, H. (ed.): *Dag Prawitz on Proofs and Meaning*.

WALLACE, R. J. (2001): Normativity, Commitment and Instrumental Reason. *Philosophers' Imprint* 1, No. 3, 1-26.

WILLIAMSON, T. (2000): *Knowledge and its Limits*. Oxford: Oxford University Press.

Akrasia and Uncertainty[*]

RALPH WEDGWOOD

School of Philosophy. University of Southern California
Los Angeles. CA 90089-0451. USA
wedgwood@usc.edu

ABSTRACT: According to John Broome, *akrasia* consists in a failure to *intend* to do something that one believes one *ought* to do, and such *akrasia* is necessarily irrational. In fact, however, failing to intend something that one believes one ought to do is only guaranteed to be irrational if one is *certain* of a *maximally detailed* proposition about what one ought to do; if one is uncertain about any part of the full story about what one ought to do, it could be perfectly rational not to intend to do something that one believes one ought to do. This paper seeks to remedy this problem, by proposing an anti-*akrasia* principle that covers cases of uncertainty (as well as cases of such complete certainty). It is argued that this principle is in effect the fundamental principle of practical rationality.

KEYWORDS: Act individuation – *akrasia* – John Broome – decision theory – practical rationality – probability – uncertainty.

1. The irrationality of *akrasia*

According to Socrates and Aristotle – at least as I shall interpret them here – *akrasia* (if it is possible at all) would involve *voluntarily* doing something that one *knows* one *ought*, *all-things-considered*, *not* to do.[1] It is widely

---

[1]   For Socrates' investigations of *akrasia*, see especially Plato's *Protagoras* (351a–358d); for Aristotle's discussion, see *Nicomachean Ethics* VII.1–10.

agreed that if such *akrasia* is possible, it is irrational – indeed, it is a paradigmatic form of irrationality. To be *akratic* in this way at a given time *t*, by voluntarily doing something at *t* that one simultaneously knows one ought not to do, is incompatible with being fully rational at that time *t*. As I shall use the term here, to say that you are at a given time *t* "rationally required" to φ is just to say that it is *necessary* for you to φ if you are to be fully rational at *t*.[2] So it seems that every agent is rationally required not to be *akratic* at any time.

In this essay, I shall not investigate the question that Socrates and Aristotle puzzled over, of how such *akrasia* is possible. Instead, I shall focus on a more basic question: assuming that we are rationally required not to be *akratic*, what is the most precise account of the general principle that underlies this rational requirement?

I shall start my exploration of this question by considering the formulation of this "anti-*akrasia*" principle that is given by John Broome. We shall see that Broome's principle is unsound unless it is restricted to a narrow range of cases – specifically, to cases where there is an extremely *fine-grained* way of carving up the available options or courses of action, such the agent in question is for all practical purposes *certain* about which of these fine-grained options she ought to do.

Nonetheless, if it is restricted in this way, Broome's principle seems to be sound. This raises the question of how we can generalize this restricted version of this principle so that it covers a wider range of cases, where the agent is uncertain about which of these fine-grained options she ought to do. The last three sections of the paper are devoted to this question.

2. Broome's principle

When Socrates undertook his investigations of *akrasia*, he assumed that if it existed at all, it would consist in voluntarily doing what one *knows* one ought not to do. But in fact, it seems plausible that even if you merely *believed*, and did not know, that φ-ing was something that you all-things-

---

2   Some philosophers – most notably, John Broome (2013) – understand the phrase 'rational requirement' rather differently. In my view, however, this is simply a terminological issue: the phrase 'rational requirement' is a semi-technical term, which is not in regular use by ordinary citizens; so it is quite legitimate for me simply to stipulate how I shall use the term here.

considered ought not to do, you would still be being *akratic* if in spite of this belief, you were voluntarily to φ. In such cases, there is a kind of conflict or incoherence between your beliefs and your will: your beliefs in some sense tell you not to φ, while your will voluntarily embraces φ-ing. Such incoherence or conflict between your beliefs and your will seems to be irrational. Rationality requires us to avoid such conflicts.

John Broome accepts this point that rationality requires a kind of coherence between one's beliefs and one's will. He has attempted to give a more precise account of exactly what this coherence consists in (see Broome 2013, Section 6.5). In order to articulate this account as precisely as possible, he makes use of several technical or semi-technical terms (which he first defines in careful detail). Since many aspects of his account are not central to the issues that I shall explore in this essay, I shall take the liberty of rephrasing his account in slightly more ordinary terminology. This rephrasing will not change the meaning of Broome's principle in any respects that are relevant to the purposes that I am pursuing here – but it should *not* be taken as a completely accurate presentation of Broome's view for all purposes whatsoever.

For our purposes, then, we may interpret Broome's principle as equivalent to this:

> Rationality requires of you that:
> If
>
> (1) You believe at *t* that you yourself ought to φ, and
> (2) You believe at *t* that, if at that time you intended to φ, then because of that, you would indeed φ, and
> (3) You believe at *t* that, if at that time you did not intend to φ, then because of that, you would not φ,
>
> Then
>
> (4) You must intend at *t* to φ.

The purpose of conditions (2) and (3) of this principle is effectively just to narrow down the scope of this requirement to cases in which, if the beliefs mentioned in these conditions are correct, you will avoid voluntarily failing to φ if and only if you *intend* at *t* to φ. I shall not worry about this aspect of Broome's principle here. I shall simply focus on cases where these two conditions (2) and (3) are met. The main focus of my discussion here will be on the relation between condition (1), believing at *t* that you ought to φ, and condition (4), intending at *t* to φ.

Broome's claims about this principle have already been widely debated (see, for example, Kolodny 2005). But it seems plausible, at least to me and to many others, that there is at least *one* interpretation of the principle such that whenever you violate Broome's principle, on this interpretation of what it means, you are being *akratic* – and so irrational. The main task of this essay is to work out what exactly this interpretation of the principle is.

One of the main issues that arise about the principle concerns the interpretation of 'ought'. In fact, the word 'ought' in English and its equivalents in other languages seem to be systematically polysemous, and capable of expressing a range of different concepts in different contexts. In effect, 'ought' has many different senses in different contexts. So, one of the central questions that we have to address is this: Which senses of 'ought' will make all instances of Broome's principle true?

Suppose that Broome's principle is true in all instances for *more than one* sense of 'ought'. Now, unless one of these senses of 'ought' analytically implies the other, it would surely be possible, at least in principle, for you to be in a case in which you are rationally required to believe that in one of these senses you "ought" to φ, and also required to believe that in another of these senses you "ought not" to φ. (For example, suppose that an oracle whose pronouncements have the most extraordinary track record for reliability announces that you are in a case where you "ought" in the first sense to φ, but "ought" in the second sense not to φ. Then it seems that you could be rationally required to have both beliefs.)

In that case, however, if Broome's principle were true in all instances for both of these senses of 'ought', you would be simultaneously rationally required to intend to φ and rationally required to intend *not* to φ. But having contradictory intentions of this sort also seems paradigmatically irrational – not something that can result from one's complying with rational requirements.

For this reason, I shall assume from now on that there is just *one* sense of 'ought' for which Broome's principle is true in all instances. I shall return later on to the question of what exactly that sense of 'ought' is. For the time being, I shall just try to read the principle sympathetically and charitably – that is, in effect, to read the principle as involving that sense of 'ought', whatever it is, that makes it most plausible that the principle is true in all instances.

3. Two issues in the philosophy of belief and intention

There are two extremely well-known issues in the philosophy of belief and intention, which seem obviously relevant to the evaluation of Broome's principle.

i. Beliefs come in *degrees*; we believe some propositions *more strongly* or *more confidently* than others.
ii. The *acts* or *options* that an agent can intend to perform can be *individuated* in different ways – sometimes more *finely* and sometimes more *coarsely*.

To illustrate the first issue, we may note that it seems that I believe the propositions that I exist, and that 1 + 1 = 2, with more confidence than the proposition that Dushanbe is the capital of Tajikistan. So, when Broome's principle refers to what "you believe" (as it does in each of its first three clauses (1), (2), and (3)), we need to know: What degree of confidence must you have in the relevant proposition for it to be true in this context to say that you "believe" the proposition?

To illustrate the second issue, we need to remember that some kinds of acts are very general and unspecific, like *moving one's hands*, while others are much more specific, like *signing a cheque to pay November's phone bill*. In this example, the first kind of act is less specific than the second kind, because it is necessary that whenever one performs an act of the second kind, one also performs an act of the first kind, but not *vice versa*. So, again, we need to know, when Broome's principle uses the schematic letter 'φ', can this letter 'φ' take the place of any act-description, or can it only take the place of an act-description that is at a certain level of generality or specificity?

Broome's principle seems most compelling when the following two conditions are met:

i. The relevant *beliefs* (referred to in clauses (1), (2), and (3) of the principle) are beliefs held with *maximum confidence*.
ii. The act of *your φ-ing* (referred to in every clause of the principle) is individuated extremely *finely*, so that it is a highly *specific* act, capturing *everything* of importance in the relevant situation.

When these conditions are met, you are *totally convinced* of a proposition that – if true – gives the *whole truth* about what you ought to do in

your situation. If you voluntarily act contrary to a conviction of this sort, you are surely being irrational in some way.

However, this is a severely restricted version of Broome's principle. This restricted version of the principle says nothing about the cases where you cannot be *rationally certain* of any such highly detailed proposition about what you ought to do. It seems clear that such uncertainty can undoubtedly arise, since facts about what you ought to do seem not to be, as we might put it, *rationally luminous*. That is, it is not in general the case that *whenever* such a fact obtains, it is possible for you rationally to have an attitude of *maximum confidence* in the proposition that that fact obtains. Indeed, a version of the famous "margins for error" argument of Timothy Williamson (2000, 93-106) seems to show that normative facts cannot be rationally luminous in this way. There could be a continuous series of cases, such that in the case at the *beginning* of the series, the only attitudes that it is possible for you rationally to have towards the proposition *that you ought to φ* all involve a high level of confidence in that proposition, while in the case at the *end* of the series, the only attitudes that it is possible for you rationally to have towards that proposition all involve a high level of *disbelief* in that proposition (and so also a high level of confidence in the negation of the proposition). We may also make two further stipulations about this series of cases: first, there are no cases in this series in which it is *both* possible for you rationally to have a high level of confidence in the proposition that you ought to φ and *also* possible for you rationally to have a high level of disbelief in this proposition; and secondly, for every case in the series after the very first case, the range of attitudes that it is possible for you rationally to have in that case differs at most only *very slightly* from the range of attitudes that it is possible for you rationally to have in the immediately preceding case.

Given these stipulations, it follows that there must be some cases, somewhere in the middle of this series, where the only attitudes that it is possible for you rationally to have in the proposition that you ought to φ (if indeed it is possible for you rationally to have *any* attitudes towards that proposition at all) are all *intermediate* levels of confidence, rather than high levels of confidence or high levels of disbelief. In those cases, given classical logic, either the normative proposition that you ought to φ is true, or its negation is true – where this negation is equivalent to the normative proposition that it is *permissible* for you *not* to φ. Either way, then, there is a true normative proposition in which it is not possible for you rationally to

have a high level of confidence. Thus, true normative proposition are not (as I put it) rationally luminous. Cases can arise in which it is impossible for you to be rationally *certain* about what you "ought" (in the relevant sense) to do.

Now, it may be that in at least some cases of this kind, there is some *other* sense of 'ought' such that you *are* rationally certain about what you "ought", in this other sense, to do. Even if that is true, however, it is irrelevant. We are assuming that there is exactly one sense of 'ought' that features in Broome's principle. The cases that we are considering are cases in which one is not certain about what one "ought" to do *in this crucial sense*. It is irrelevant if in some of these cases, there is some other sense of 'ought' such that you are certain of what you "ought" in that other sense to do. That other sense of 'ought', whatever it may be, is not the sense that appears in Broome's principle, and so need not concern us here.

It is crucial to see that this Williamson-inspired argument shows that *no* non-trivial sense of 'ought' is rationally luminous: for every non-trivial sense of 'ought', cases can arise in which a rational agent cannot be certain about what she ought to do. For this reason, any version of Broome's principle that is restricted to cases in which the agent is certain about what she ought to do is, as I have said, a severely limited principle. The significance of this point will emerge in the sequel.

4. Against the unrestricted form of Broome's principle

In this section, I shall consider an *unrestricted* form of Broome's principle. In this unrestricted form of the principle:

i.   The "beliefs" referred to in the principle can be held with *any* degree of confidence that is at least as great as some threshold t, where $t < 1$.

ii. The schematic letter 'φ' can stand for *any* act, regardless of whether it is a finely individuated, highly specific act or a coarsely individuated, highly general act instead.

As I shall argue here, this unrestricted form of Broome's principle is open to fatal counterexamples. I shall start by presenting a counterexample in which *none* of the agent's beliefs are held with the maximum level of confidence; this counterexample does not depend on the issue of how finely or coarsely the relevant acts are individuated. Then I shall present a

second counterexample, in which the agent is *certain* about which *coarse-grained* acts she ought to do, and is uncertain only about which *fine-grained* acts she ought to do.

4.1. Counterexample (i) to the unrestricted form of Broome's principle: Uncertainty about all options

The first counterexample to the unrestricted form of Broome's principle is a case in which there are two options available to you, *A* and *B*, such that these two options form a *partition* – i.e., you are certain that you will do *one*, and *no more than one*, of these two options.

In this case, although you are *not certain* whether you ought to do *A*, or ought to do *B*, you have a *very high* degree of confidence that you ought to do *A* – a degree of confidence x such that $x \geq t$. In other words, your degree of confidence that you ought to do *A* is at least as great as the crucial threshold t. Still, in your view, you cannot absolutely rule out the rival hypothesis that you ought to do, not *A*, but *B* instead; and so your degree of confidence x in the proposition that you ought to do *A* is less than certainty – that is, $x < 1$.

Now, let us also assume that in this case, you are *conditionally* certain, given the assumption that it is *not* the case that you ought to do *A*, that *B* is not just slightly better than *A*, but *astronomically* better than *A*. (For example, perhaps, if it is not the case that you ought to do *A*, doing *A* will result in the destruction of the whole world or the like.) On the other hand, you are also conditionally certain, given the assumption that it is the case that you ought to do *A*, that *A* is only very *slightly* better than *B*.

In this case, it seems possible for you to be rational, to have beliefs of this sort, and simultaneously to intend to do not *A*, but *B* instead. If that is right, then this case is a clear counterexample to the unrestricted version of Broome's principle. You are perfectly rational, you have a degree of belief above the threshold t that you ought to do *A*, and yet you do not intend to do *A* – you intend not to do *A*, but to do *B* instead.

Some readers might suspect that we can get round counterexamples of this sort simply by amending the unrestricted form of Broome's principle so that the kind of 'belief' referred to in the principle must consist of beliefs of which the agent is, for all practical purposes, completely *certain*. As we shall see in the next subsection, this suspicion is incorrect: this amendment does not make the principle immune to counterexamples of this kind.

4.2. Counterexample (ii) to the unrestricted form of Broome's principle: Uncertainty about the fine-grained options

Our second counterexample concerns a case – like Frank Jackson's "three drugs" case[3] – where there are *three* fine-grained options available to you: *A*, *B*, and *C*. Again, suppose that these three options form a partition (that is, you are certain that you will do exactly one of these three options). There are also some coarse-grained options, like *doing A or B*, or *doing B or C*, or *not doing A*, and so on. Since *A*, *B*, and *C* form a partition, the coarse-grained option of *not doing C* and the coarse-grained option of *doing A or B* are effectively equivalent.

In this case, suppose that you are certain that *either* you ought to do *A or* you ought to do *B*; and you are also certain that you ought *not* to do *C*. However, you are radically *uncertain* about whether the option that you ought to take in this situation is *A* or *B*.

In addition, in this case, you are *conditionally* certain, given the assumption that you ought to do *A*, that doing *B* will be utterly disastrous (it will result in the destruction of the world or the like), and you are also conditionally certain, given the assumption that you ought to do *B*, that doing *A* will be equally disastrous. However, you are *also* certain that doing *C*, though it falls short of being what you strictly ought to do, is not *too* bad: it is far less bad than doing *A* would be if you ought instead to have done *B,* and equally far less bad than doing *B* would be if you ought instead to have done *A*.

In this case, you might be rational, and be certain that you ought not to do *C* (or, equivalently, that you ought to do either *A* or *B*), without intending not to do *C* (or, equivalently, without intending to do *A* or *B*). If that is right, then this case is also a counterexample to the unrestricted form of Broome's principle: you are rational, you are certain that you ought not to do *C,* and yet you do not intend not to do *C*.

In this way, then, the unrestricted form of Broome's principle seems to be faced with fatal counterexamples. This unrestricted form of the principle is unacceptable.

---

[3]   See Jackson (1991); another famous case of this sort is the "three mineshafts" case of Parfit (2011, 159), which was inspired by an example of Regan's (1980, 264-265, n. 1).

5. A better way of generalizing the restricted form
of Broome's principle

Still, as I commented above, the restricted form of Broome's principle – restricted to cases where the agent is *certain* about which *fine-grained* option she ought to do – seems compelling. It does seem necessary that if an agent is perfectly rational, and is certain of the truth of such a fully specific proposition about what she ought to do, then the agent will intend to do the fine-grained act that she is certain she ought to do.

However, as I argued in Section 3, the restricted form of Broome's principle is severely limited to a narrow range of cases. It seems plausible that if this restricted principle holds in this narrow range of cases, that will be because of some more general truth that explains why it holds in these cases. But what is this more general truth? How can we generalize this restricted form of the principle, in order to cover cases of uncertainty about which fine-grained option the agent ought to do, while avoiding these troublesome counterexamples?

5.1. Generalizing (i): Uncertainty

First, let us just focus on the issue of *uncertainty*. Let us leave aside the issue of option-individuation, for the time being – let us simply assume that we are considering only super-finely individuated options.

It seems clear that in each of the troublesome cases that we have just considered, the rational intention is an intention that *maximizes* some kind of *expectation* of some kind of *value*. We might try to treat the principle that rational intentions maximize the relevant sort of expected value as if it were a completely separate principle from the restricted form of Broome's principle. But it seems as if it would be more promising to unify these principles somehow. This is what I shall try to do here. Specifically, I shall try to find a version of the idea that rationality requires us to have intentions that maximize expected value which implies the restricted form of Broome's principle as a special case.

In defining any notion of expected value, we need to appeal to two real-valued functions: first, a *probability* function, and secondly, a *value*-function of some kind. To identify a version of the idea that rationality requires one's intentions to maximize expected value which implies Broome's principle as a special case, we must interpret the probability function that is involved in determining the relevant expectation as modelling the rational

agent's *degrees of belief*; and we must interpret the relevant value-function as a measure of how *closely* the fine-grained options approximate to being *what the agent ought to do*.

The idea of modelling a rational agent's degrees of belief by means of probability functions is familiar. The idea of using value-functions to measure how closely such fine-grained options approximate to being what the agent ought to do is less familiar. But it is not too hard to get the hang of this idea. Intuitively it seems clear that we can *compare* the available fine-grained options to each other in terms of how closely they approximate to being what you ought to do. Out of the options that fall short of being what you ought to do, some of these options fall only *slightly* short of being what you ought to do, while others fall *atrociously* far short of being what you ought to do. In other words, of the options that it would be wrong or inappropriate for you to choose, some are more *badly* or *seriously* wrong than others. As I shall say, some are *less choiceworthy* than others.

It seems plausible that there is a way of talking about "reasons for action" on which – at least wherever φ-ing is a fine-grained option – an agent has "most reason" to φ if and only if the agent ought all things considered to φ. So these comparisons of fine-grained options in terms of their degrees of choiceworthiness are effectively equivalent to comparisons of options in terms of *how much reason* there is in their favour.

Moreover, there are reasons for thinking that this notion of choiceworthiness gives us more than just a *ranking* of these options. Specifically, there are reasons for thinking that this notion allows us to make sense of the *cardinal measurement* of choiceworthiness. It seems that we can not only compare *options* in terms of their degrees of choiceworthiness; we can also compare the *differences* in choiceworthiness between options – e.g., we can say that the difference in choiceworthiness between options A_1 and A_2 is a *small* difference, compared to the much *larger* difference between options B_1 and B_2. This supports the view that we can make sense of the cardinal measurement of choiceworthiness.[4] Then we could say that rationali-

---

[4]   A system of four-place relations, comparing the differences between pairs of items with respect to some quantity or value, is known as a *difference structure*. If a difference structure gives a *complete* ranking of all differences with respect to a certain value between *infinitely many* pairs of items, then there is in fact a *unique* interval scale (given an arbitrary choice of a unit and zero point) on which the value in question can be measured. For this point, see Krantz et al. (1971).

ty requires one to have an intention that (out of the relevant set of available alternative intentions) *maximizes expected choiceworthiness*.[5]

To make sense of this notion of expected choiceworthiness, we must suppose that it is rational for the agent to have various *degrees of belief* in various propositions about the *degrees of choiceworthiness* of the relevant options. If these propositions form a partition (that is, it is rational for the agent to be certain that one and no more than one of these propositions is true), the expected choiceworthiness of an option is the weighted sum of its degree of choiceworthiness according to each of these propositions, weighting each degree of choiceworthiness by the degree of belief that it is rational for the agent to have in the relevant proposition.

If the agent has degrees of belief of this sort, then one special case of such degrees of belief is the case in which the agent has the *highest possible* degree of belief in the proposition that option A has a *greater* degree of choiceworthiness than *all alternative options* – that is, in effect, the agent is *certain* that A is what she *ought* to do. This is the case to which the restricted version of Broome's principle applies.

We may give a more formal representation of this notion of expected choiceworthiness, in the following way. First, let us suppose that there is a set of probability functions, including all and only those probability functions P that faithfully represent the degrees of belief that it is rational for the agent to have; let these probability functions be defined over a set of epistemically possible worlds – where these possible worlds are, intuitively, the most specific and detailed propositions that it is rational for the agent to regard as potentially relevant to the decision in question.

Secondly, let us represent these degrees of choiceworthiness by means of a *set* of *real-valued value-functions*. Suppose that every value-function V in this set assigns a real number to every relevant world W – where V assigns a number to each world W based purely on the degree of choiceworthiness of the *fine-grained* option that the agent does in the relevant situation in W.

Let 'Intend: A' stand for the first-personal present-tensed proposition that the agent could express by saying something of the form 'I intend to do A'. Then, if A and B are both fine-grained options, we may say that the intention to do A has greater expected choiceworthiness than the intention

---

[5]   This proposal is not totally unprecedented. For example, in the context of a discussion of "moral obligation", Peter A. Graham (2010) has suggested (in effect) that a morally conscientious agent will seek to *minimize* her conduct's *expected* degree of moral *wrongness*.

to do B if and only if, for every pair $\langle V, P \rangle$ consisting of one of these value functions and one of these probability functions:

$$\Sigma_W \, V(W) \, P(W \,|\, \text{Intend: } A) > \Sigma_W \, V(W) \, P(W \,|\, \text{Intend: } B)$$

This notion of expected choiceworthiness can be used to make the following proposal about the intentions that it is rational for an agent to have: When each of the relevant alternative intentions is an intention to do a fine-grained option, rationality requires the agent to have an intention that (out of these alternatives) maximizes expected choiceworthiness in this sense.[6]

This proposal may *look* similar to classical decision theory (according to which rational choices must maximize "expected utility"). In fact, however, if the restricted version of Broome's principle is to be a special case of this proposal, there have to be some crucial differences.

In particular, the restricted version of Broome's principle concerns cases in which the agent is rationally certain that a certain fine-grained course of action is what she ought to do. So, if this version of Broome's principle is to be a special case the proposal that I am making here, then having certain degrees of belief in certain possible worlds must be *equivalent* to having a certain degree of belief in a (normative or evaluative) proposition about which of the relevant options the agent ought to do.

In effect, then, we should think of each world W as an extremely detailed conjunctive proposition, some conjuncts of which are evaluative propositions about the degree of choiceworthiness of the fine-grained act that the agent does in the relevant situation. To get a rough picture of what this amounts to, we might imagine that the relevant evaluative conjunct of this world W is the proposition that the agent might express by saying 'The fine-grained act that I do in this situation is choiceworthy to degree n'. In this case, we could imagine simply that the real number $V(W)$ that the value-function V assigns to W is precisely n.

---

[6]   In this definition of an intention's expected choiceworthiness, I have appealed to the *conditional* probability of each world W given the assumption that the agent has the intention in question. For my reasons for defining the notion in this way, see Wedgwood (2011a). For the purposes of this paper, however, this point is not important. These "evidential" conditional probabilities could easily be replaced with a more "causal" notion of probability without affecting my arguments. For my reasons for appealing to the conditional probability of the world given that you *have the intention* (rather than given that you actually *carry out* the intention), see Wedgwood (2011b).

In fact, however, we need to recognize that the precise number that this value-function V assigns to each world is really just an arbitrary device for representing the structure of these degrees of choiceworthiness. The relevant possible worlds themselves do not need to assign any real numbers to these degrees of choiceworthiness; they just need to imply enough about these degrees of choiceworthiness so that what the worlds imply about the choiceworthiness of the agent's actions in those worlds can be represented by means of a value-function like V.

At all events, we must not think of these possible worlds as encoding only *empirical* uncertainty about non-normative non-evaluative matters of fact; we must think of them as encoding the agent's uncertainty about *normative* and *evaluative* matters as well. The value-function is simply a way of representing a feature of the content of these worlds – specifically, it represents what each world implies about the degree of choiceworthiness of the act that the agent performs in the relevant situation.

In this way, this value function differs crucially from a "utility" function, since the number that a utility function assigns to a world is not determined purely by the content of the world; it is also determined by the agent's subjective preferences, of which this particular utility function is a measure.[7] These preferences can vary independently of the worlds (for example, different agents' utility functions might rank two different worlds in very different ways); so a utility function is clearly not just a way of representing any feature of the content of the worlds.

Admittedly, we have said nothing so far about the precise meaning of the relevant kind of 'ought', or about the nature of this value of "choiceworthiness". So we have not ruled out the suggestion that (like "utility") this value is determined by purely the agent's subjective attitudes. However, even if this value is determined by the agent's subjective attitudes, this proposed principle does not imply that the rational agent's intentions must cohere or harmonize in any way with these subjective attitudes themselves: it requires only that the agent's intentions must cohere with the agent's *de-*

---

[7]   This conception of "expected value" also differs crucially from the conception that appears in Jackson's (1991) "decision-theoretic consequentialism", since the value-function that Jackson appeals to is the value-function that corresponds to the *truth* about morality, whereas in my approach this value-function is simply a way of formulating the *content* of the agent's *degrees of belief* about the relevant options' *degrees of choiceworthiness*.

grees of belief in propositions about the degrees to which the relevant options exemplify this value.

Here is another way of bringing out the distinction between this proposal and classical expected utility theory. Classical expected utility theory is compatible with a strictly *expressivist* and *non-cognitivist* treatment of evaluative and normative language, according to which (at the most fundamental level of analysis) evaluative and normative statements do not express beliefs in ordinary propositions, of the sort that are expressed by ordinary factual statements, but instead express mental states of some fundamentally different "non-cognitive" kind. This sort of expressivist non-cognitivism supports the conclusion that what it is for the mental states that are expressed by these normative and evaluative statements to be rational or justified will ultimately be crucially different from what it is for ordinary factual beliefs to be rational or justified. If this sort of expressivist non-cognitivism is correct, then it is natural to think that we should not model our normative and evaluative attitudes in the same way as our ordinary factual beliefs, as an assignment of degrees of belief across a space of epistemically possible worlds. Instead, we should model these normative and evaluative attitudes as a system of subjective preferences or the like.

By contrast, the proposal that I am making here coheres most straightforwardly with a *cognitivist* and *truth-conditional* interpretation of normative and evaluative statements. According to this sort of cognitivism, even at the most fundamental level of analysis, the meaning of these statements involves a *proposition* – the proposition that gives the truth-conditions of those statements – and in making these statements, speakers express an ordinary attitude of *belief* towards these propositions – an attitude that is of fundamentally the same kind as the attitude of belief that we have towards ordinary factual propositions. This interpretation naturally encourages the view that we have degrees of belief in evaluative and normative propositions, in just the same way as in ordinary factual propositions; and in consequence the relevant epistemically possible worlds, over which our degrees of belief are defined, must be thought of as big conjunctions of both normative and non-normative propositions.

The core of this approach, then, is the idea that the rational agent is guided by her degrees of belief in normative or evaluative propositions about the relevant available options' degrees of choiceworthiness. To that extent, this proposal is in line, not with the Humean tradition, according to which reason is necessarily "the slave of the passions", but rather with

the broadly Aristotelian or Thomistic tradition, according to which the rational will is fundamentally "moved by the intellect".[8]

5.2. Generalizing (ii): Option-individuation

How can the account given above about which *fine-grained* options it is rational to intend be extended into an account of which *coarse-grained* options it is rational to intend?

To capture any requirements that apply to your intentions to do coarse-grained options, we will need a *holistic* constraint on the total set of intentions that you have at the relevant time. Suppose that we can define a notion of the expected choiceworthiness of a whole set of intentions. Then we can say that if you are rational, you will have a whole set of intentions that (out of all relevant alternative sets of intentions) maximizes expected choiceworthiness (in this sense).

Let 'Conj-Intend: *P*' stand for the proposition that the *conjunction* of all the contents of your intentions is the proposition *P*. If the conjunction of the contents of one set of intentions S_1 is *P* and the conjunction of the contents of a second set of intentions S_2 is *Q*, then S_1 has greater expected choiceworthiness than S_2 if and only if:

$$\Sigma_W V(W) \, P(W \mid \text{Conj-Intend: } P) > \Sigma_W V(W) \, P(W \mid \text{Conj-Intend: } Q)$$

Using this notion of the expected choiceworthiness of a set of intentions, we may now make our most general proposal about what it is rational for an agent to intend: To be rational, an agent must have a *set* of intentions that (out of the relevant alternative sets of intentions) maximizes expected choiceworthiness in this sense.

In other words, the basic idea is this: the agent's intentions must make it rational for the agent to have an expectation for the degree of choiceworthiness that her conduct will exemplify in the relevant situation that is *at least as great* as the expectation that every relevant alternative set of intentions would make it rational for her to have.

It is intuitively clear, it seems to me, that this new principle entails the restricted form of Broome's principle as a special case. That is, this new

---

[8]   As Aquinas puts it (cf. *Summa Theologica* IaIIae, 9.1), "*intellectus movet voluntatem.*" In the terms that were suggested by Cullity – Gaut (1996), this is fundamentally a "recognitional" rather than a "constructivist" conception of practical reason.

principle guarantees that if you violate the restricted form of Broome's principle, you are irrational.

Suppose that φ-ing is a *super-finely* individuated option, and you are *certain* that in situation S you ought to φ (and Broome's clauses (2) and (3) are met), but at t you do not intend to φ. Then there are two possible cases. In the first case, it is not rational for you to be certain that in this situation you ought to φ. In this case, it is clear that you are being in at least one way irrational.

In the second case, it is rational for you to be certain that in this situation you ought to φ. Given that φ-ing is a super-finely individuated option, the proposition that you ought to φ is by definition equivalent to the proposition that φ-ing is more choiceworthy than all the relevant alternatives. So you are in effect rationally certain that φ-ing is more choiceworthy than all alternatives. In that case, if you are rational, *all* the propositions about the available options' degrees of choiceworthiness in which you have any non-zero degree of belief assign a higher degree of choiceworthiness to φ-ing than to every alternative. At least assuming that Broome's conditions (2) and (3) are met with respect to each of the relevant alternative options, it follows that the intention to φ is the only intention that maximizes expected choiceworthiness. So, in this case, you are rationally required to intend to φ, and if you do not intend to φ you are irrational.

At the same time, this new principle entails the intuitively correct answer to the cases that I put forward in Section 4 as counterexamples to the unrestricted form of Broome's principle. For example, in the case considered in subsection 4.1, no set of intentions that includes the intention to do A will maximize expected choiceworthiness, since on the assumption that you intend to do A, there is too high a risk that your conduct in the relevant situation will involve destroying the world. On the other hand, on the assumption that you intend to do B, there are no such risks of your destroying the world; and so it seems that a set of intentions containing an intention to do B could well maximize expected choiceworthiness, and so could count as a rational set of intentions.

6. Expectations *vs.* beliefs

As we have just seen, in the most general formulation, the principle proposed here says, not that rational agents' intentions are in line with

their *beliefs* about what they *ought* to do, but that their intentions are in line with their *expectations* of *choiceworthiness.*

Beliefs and expectations are crucially different mental phenomena:

a. Each of your beliefs is an attitude towards a *single* proposition.
b. Each of your expectations is determined by your degree of belief in each member of a *partition* of propositions.

There is admittedly one special case in which a belief coincides with an expectation – namely, in the special case in which you are absolutely *certain* of the relevant proposition (in which case the relevant partition of propositions in effect has just one member). Except in this special case, however, beliefs and expectations are importantly different. As David Lewis (1988) taught us, beliefs and expectations behave quite differently in response to new evidence. So long as your initial degree of belief in a proposition $p < 1$, new evidence can lower your degree of belief in p below any threshold t; that is, in effect, new evidence can *deprive* you of having any belief in p. By contrast, if you are rational, new evidence will never deprive you of having any expectation of choiceworthiness for the intention to do A.

So, according to the proposal that I am making here, the fundamental account of rationality is that the rational agent's intentions are in line with her expectations of choiceworthiness, not with her beliefs about what she ought to do.[9]

Some philosophers will be inclined to object that I am underestimating the importance of beliefs about what one ought to do. In particular, some of these philosophers will object along the following lines. According to the principle of Section 5, whenever there is a unique set of intentions that maximizes expected choiceworthiness, you are rationally required to have those intentions. Arguably, the notion of a "rational requirement" is a kind of 'ought' – specifically, it is a "subjective 'ought'", in the sense that what the agent "subjectively *ought*" to do is determined by facts about the evidence or information that is available to that agent (not by facts about the external world of which the agent is ignorant). If that is right, then in all of these cases, there is a sense of 'ought' such that the agent *ought* in that sense to have the intentions that she is rationally required to have. So, in

---

[9] One philosopher who appreciated this point more than twenty years ago, ironically enough, was John Broome (1991), in his commentary on Lewis (1988).

every one of these cases, there is a kind of 'ought' – the 'ought' of rational requirement – that the rational agent's intentions will conform to.

Moreover, some philosophers might think that because this kind of 'ought' – the 'ought' of rational requirement – is fixed by the degrees of beliefs that it is rational for the agent to have, there is no room for any real difference here between the claim that the agent "ought" in this sense to intend to φ (that is, the agent is rationally required to intend to φ) and the claim that the agent *rationally believes* that she "ought" in this sense to intend to φ. So these philosophers would think that a principle very similar to the unrestricted form of Broome's principle is true – namely, the principle that to be rational, one must intend to φ whenever one rationally believes that one ought in this sense to intend to φ.

As tempting as this line of thought may seem to some philosophers, it cannot be reconciled with the Williamson-inspired point that no non-trivial 'ought' is rationally luminous. Since the notion of a "rational requirement" is a kind of 'ought', rational requirements are not luminous either: cases can always arise where it is *true* that you are rationally required to intend to φ, but it is impossible for you rationally to believe that you are rationally required to intend to φ. So, it seems that for *every* sense of 'ought', there is a gap between the proposition that you ought, in this sense, to intend to φ, and the proposition that you *rationally believe* that you ought, in this sense, to intend to φ. Cases can arise where the first proposition is true and the second proposition is false.

As a result, the principle that says that for you to be rational, your intentions must be in line with your rational beliefs about what intentions you are rationally required to have is at best significantly *narrower* – that is, covers a significantly smaller range of cases – than the principle that I am advocating here, according to which for you to be rational, your intentions must actually maximize expected choiceworthiness. If the more general principle is correct, as I am advocating here, then it seems that the narrower principle has no interest except as a special case of that more general principle.

Moreover, many philosophers would accept that it is possible for you to have a rational but *false* belief in the proposition that you are rationally required to intend to φ. (For example, if a suitable oracle pronounces that you are rationally required to intend to φ, perhaps you could be rationally required to have a high degree of confidence in the proposition that you are rationally required to intend to φ, even if on this one occasion, the oracle's

pronouncement is actually false.) If this is indeed possible, then these two principles are not just different but *inconsistent* with each other. In that case, since I am advocating the principle that to be rational, your intentions must maximize expected choiceworthiness, I would be committed to denying the principle that to be rational, your intentions must be in line with your rational beliefs about what intentions you are rationally required to have.

In general, then, it is not rational agents' *beliefs* about what is rationally required of them that fundamentally guide their deliberations. It is these *rational requirements themselves* – and the facts about the agents' degrees of belief in propositions about the available options' degrees of choiceworthiness, on which these rational requirements supervene – that will guide the agents' deliberations.

This point helps us to understand the precise sense of 'ought' that appears in the true instances of Broome's principle. This sense of 'ought' is closely connected to the notion of "choiceworthiness": as I put it above, choiceworthiness is precisely a measure of how closely each of the available options approximates to being what the agent ought, all things considered, to do in the relevant situation. Thus, the very concept of choiceworthiness guarantees that for any fine-grained option *A*, the agent *ought* in the relevant sense to do *A* if and only if *A* is the *most choiceworthy* option available to that agent in the relevant situation.

According to the principle that was proposed in Section 5, uncertainty is handled by means of the *degrees of belief* (and the sets of probability functions that represent those degrees of belief) that are involved in defining the notion of *expected choiceworthiness*. For this reason, it would be *double-counting* if the concept of 'choiceworthiness' that appears in this proposed principle also took account of uncertainty. So, this concept of "choiceworthiness" must be a *maximally objective* normative notion: the truth about the relevant available options' degrees of choiceworthiness depends on the objective facts of the agent's situation, which may include facts that the agent is not even in a position to know at the relevant time.

I argued at the end of the previous section that the restricted version of Broome's principle is just a special case of the principle that was proposed in that section. Given the fundamental connection between the concepts expressed by these uses of 'ought' and 'choiceworthiness', it follows that the sort of 'ought' that appears in this restricted version of Broome's principle must also be a maximally objective 'ought' – the kind of 'ought' for which

what one ought to do may depend on facts that one is not in a position to know at the relevant time.[10]

7. The fundamental principle of rationality

Someone might object to the principle that was proposed in Section 5, along the following lines. It seems possible for there to be an agent who is capable of rational choices and intentions, but has no degrees of belief in *any* propositions about the relevant options' degrees of choiceworthiness. How could such an agent have a set of intentions that maximizes expected choiceworthiness? And if it is not possible for an agent to have a set of intentions that maximizes expected choiceworthiness, how could the agent be rationally required to have such intentions?

My definition of expected choiceworthiness appeals to a set of probability functions – specifically, the set of probability functions that models the degrees of belief that it is rational for the agent to have. This objection to the principle that I proposed in Section 5 fails, because there can *be* degrees of belief that it is rational for the agent to have, even if the agent does not *actually* have these degrees of belief.

According to most theories of rationality, the degrees of belief that it is rational for an agent to have at a given time are determined by such things as the evidence that the agent has at that time, or by the mental states and mental events that are present in the agent's mind at that time. Even if the agent does not actually have these degrees of belief, it is presumably still possible for the agent to be guided, in forming and revising her intentions, by this body of evidence, or by these mental states and mental events. In this way, it could be that it is no accident that the agent has a set of intentions that maximizes expected choiceworthiness – if the agent is guided, in forming and revising these intentions, by whatever evidence or mental states determine the degrees of belief that it is rational for the agent to

---

[10]   At the same time, it is easy to explain why the *assertions* that we make about what agents ought to do will often involve a less objective kind of 'ought'. When we make assertions, we are generally highly confident of the truth of what we say. We are often not very confident of the propositions involving this objective 'ought'. Still, even if this objective 'ought' is often not used in our assertions, it may nonetheless appear in propositions towards which we have partial degrees of belief.

have in the relevant propositions about the options' degrees of choiceworthiness.

Once the principle proposed in Section 5 is clarified in this way, it becomes clear that if this principle is true, it is not just one principle of practical rationality among many. It is in a sense the *fundamental* principle of rationality. This is because this principle applies quite *generally*, to all cases of rational intention; so this principle conflicts with all other proposed principles of rational intention, except for those that are implied by it.

For example, at least on most interpretations of what "utility" is, the principle proposed in Section 5 conflicts with the principle that rational intentions must maximize expected utility. The reason for this is that on many common interpretations of "preference", it is possible for a rational agent to have a "preference" for A over B, even if the expected choiceworthiness of B is greater than that of A. The dominant interpretation of "utility" – ever since John von Neumann and Oskar Morgenstern (1944) – is simply as a measure of the relevant agent's "preferences". So (on those common interpretations of "preference") it is possible for A to have greater expected choiceworthiness than B even if B has greater expected utility than A. It follows that the principle proposed in Section 5 is inconsistent with the view that rational intentions must maximize expected utility.

There are still some ways in which this principle needs to be clarified. In particular, we need to figure out exactly how to understand the "relevantly available alternative sets of intentions" that are mentioned in the principle. What is it for a set of intentions to be one of these "relevantly available alternative sets"? This point clearly needs to be clarified if we are to understand the exact implications of this principle.[11]

Subject to these clarifications, however, our search for an anti-*akrasia* principle that can handle cases of uncertainty and intentions involving coarsely individuated options seems to have led us to the fundamental principle of practical rationality.

---

[11] I have tried to explore this question in a little more detail elsewhere; see Wedgwood (2011b).

References

BRATMAN, M. (1988): *Intentions, Plans, and Practical Reasoning*. Cambridge, MA: Harvard University Press.

BROOME, J. (1991): Desire, Belief and Expectation. *Mind* 100, 265-267.

BROOME, J. (2013): *Rationality through Reasoning*. Oxford: Wiley-Blackwell.

CULLITY, G. – GAUT, B. (eds.) (1996): *Ethics and Practical Reason*. Oxford: Clarendon Press.

GRAHAM, P. (2010): In Defense of Objectivism about Moral Obligation. *Ethics* 121, 88-115.

JACKSON, F. (1991): Decision-Theoretic Consequentialism and the Nearest and Dearest Objection. *Ethics* 101, 461-482.

KOLODNY, N. (2005): Why Be Rational? *Mind* 114, 509-563.

KRANTZ, D. H. – LUCE, R. D. – SUPPES, P. – TVERSKY, A. (1971): *Foundations of Measurement I: Additive and Polynomial Representations*. London: Academic Press.

LEWIS, D. K. (1988): Desire as Belief. *Mind* 97, 323-332.

MORGENSTERN, O. – VON NEUMANN, J. (1944): *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.

PARFIT, D. (2011): *On What Matters*. Vol. 1. Oxford: Oxford University Press.

REGAN, D. (1980): *Utilitarianism and Cooperation*. Oxford: Clarendon Press.

WEDGWOOD, R. (2011a): Gandalf's Solution to the Newcomb Problem. *Synthese*, Online First: 15 March, 1–33. DOI: 10.1007/s11229-011-9900-1.

WEDGWOOD, R. (2011b): Instrumental Rationality. *Oxford Studies in Metaethics* 6, 280-309.

WILLIAMSON, T. (2000): *Knowledge and its Limits*. Oxford: Clarendon Press.

The Distinctive Rationality of Intentions[1]

DAVID BOTTING

Instituto de Filosofia da Linguagem. Faculdade de Ciências Sociais e Humanas
Universidade Nova de Lisboa. Av. de Berna, 26 - 4º piso. 1069-061 Lisboa. Portugal
davidbotting33@yahoo.co.uk

ABSTRACT: In this paper I intend to defend Broome's cognitivist view that reduces practical normativity to theoretical normativity, but argue that this leaves unaccounted for distinctively practical norms that I seek to capture as a system of local obligations to have particular intentions. The krasia requirement dictates what obligations we have relative to the normative beliefs that we have but does not tell us what intentions it is rational to have all-things-considered.

KEYWORDS: Broome – cognitivism – krasia – practical rationality.

1. Cognitivism and normativity

This paper attempts to account for a distinctively practical normativity. What normativity is this, exactly? Is there a normativity involved in practical reasoning that is distinct from that involved in theoretical reasoning? It is not enough to say simply that these are different kinds of reasoning with different subject-matters and have intentions or actions as their conclusion. In other words, before I can even get started, I need to narrow down what is distinctive, and to do this I need to discuss cognitivism.

Cognitivism generally is the thesis that norms of practical rationality are to be accounted for in terms of norms of theoretical rationality, i.e., norms concerning doxastic states like belief. If cognitivism is true, then there is no distinctively practical normativity; hence, we must find some normativity that cognitivism does not explain. I will take a norm of practical reasoning and show what normativity remains unexplained by a given cognitivist account. This residuum will then be whatever is distinctive and will be the explanandum for a non-cognitivist account. The basis of this non-cognitivist account will be the *krasia requirement*.

For my norm of practical reasoning I will give the *instrumental requirement*:

(IR$_W$) Rationality requires that (if I intend to x and believe that y is a means to x-ing, then I should intend to y).[2]

This seems to be basic to any reasoning about what to do. For my cognitivist account I will give Broome's (2009) attempt to give a cognitivist account of IR$_W$. Broome attempts to show that the normativity involved in IR$_W$ can be accounted for by the normativity involved in norms that say, for example, that you should not hold inconsistent beliefs. In particular, he attempts to show that it is by satisfying requirements on theoretical reasoning that we bring ourselves to satisfy IR$_W$.

Broome's method of accounting for IR$_W$ is to describe a process of reasoning that satisfies it by satisfying a requirement of theoretical rationality. Kolodny (2005) argues that "for any rational requirement on you, there must be a process of reasoning through which you can bring yourself to satisfy that requirement," but Broome (2006, 2) quite explicitly rejects Kolodny's arguments for this view, expressing agnosticism towards its conclusion, and in Broome (2009) we see why, for he says:

The attitudes that can figure in your conscious reasoning must be ones you are conscious of, which is to say ones you believe you have. Conversely, if you believe you have an attitude that actually you do not have,

---

[2]   There is ongoing debate I do not mean to survey here whether wide or narrow-scope requirements are to be preferred and over what issues should determine our preference. I will assume a wide-scope requirement. If there is a cognitivist account of a wide-scope requirement then there ought to be a similar account of the corresponding narrow-scope requirement, so the distinction should not matter for my purposes.

the fact that you do not have it cannot impinge on your conscious reasoning.

The consequences are disturbing from a theoretical point of view. They leave us with an unpleasant dilemma... [R]easoning cannot [always] bring you to satisfy *Instrumental requirement* [sic]. We shall have to conclude either that *Instrumental requirement* [sic] is not a genuine requirement of rationality, or alternatively that it is a genuine requirement but not one that reasoning can always bring you to satisfy. (Broome 2009, 18)

By describing such a process Broome does not take himself to be proving IR_W to be a genuine requirement,[3] and, there being some cases (where the reasoner believes that they have an intention they do not actually have) where IR_W cannot be so satisfied, Broome explicitly leaves this issue open. By appealing to requirements that can be satisfied by reasoning Broome is limiting what a cognitivist account has to explain, viz., the normative force of IR_W and not its substance, where this means that it is not required that practical norms be equivalent in extension to theoretical norms but only that the normative force of the former be derivable from or reducible to the normative force of the latter.

I will give the useful summarization of Broome's account given in Bratman (2009, 31-32), who refers to a requirement of theoretical reasoning he calls BC (Bratman 2009, 14) where this can be defined as:

(BC_W) Rationality requires that (if you believe that E and you believe that E will only occur if M, then you should believe M).

Armed with IR and BC (which I have called IR_W and BC_W and substituted into what follows) he shows how, by satisfying BC, Broome hopes to show that a reasoner can satisfy IR:

(1)   Intention is not belief.
(2)   It is not in general true that if you intend E you believe E.

This last is because (and here Broome and I agree)

(3)   You can sometimes intend E but fail to believe you intend E; and in such cases you may well not believe E.
(4)   But, if you do believe you intend E then you will believe E.

---

3   Høj (2009, 2) mistakenly attributes Kolodny's view to Broome too.

(5)   And for your intention to E to enter into practical reasoning about means to E, you need to believe you intend E.

(6)   So when your intention to E enters into your practical reasoning about means to E, you believe E.

(7)   And it is this belief that E that provides the premise for your reasoning, namely: E.

(8)   If you also believe that E only if M, and if these beliefs do not change, BC_W requires that you believe M; and that is where your reasoning can lead.

(9)   But if in the "background" you believe that M will obtain only if you intend M, then if you do arrive at the belief that M this will normally be by way of intending M. In satisfying BC_W in this way you will satisfy IR_W.

This, then, is Broome's broadly cognitivist picture of reasoning from intended ends to intended means. Broome wants to acknowledge, though, that it remains possible to intend E, believe that this requires both M and that you intend M, but falsely believe that you intend M. Broome grants that in such a case you fail to satisfy IR_W, though you may well satisfy BC_W. But, says Broome,

(10)   In such a case "your false belief blocks any reasoning that can bring you to satisfy IR_W." So,

(11)   Insofar as IR_W is a rational demand that can be satisfied by reasoning it is demand [sic] that derives from BC_W. Insofar as IR_W seems to impose demands that go beyond what is imposed by BC_W, these are not demands that can be satisfied by reasoning.

In this sense, it is BC_W that is fundamental for a theory of practical reasoning from ends to means.

This requires some commentary.

There are basically three parts to this summary. The first part consisting of steps (1) to (7) are meant to justify "I believe E" as a premise that can be got from the intention that E and that is then used in the conscious reasoning that is modelled in the second part consisting of steps (8) and (9). This reasoning goes:

a) I believe that E.
b) I believe that E only if M.
c) I should believe M.

Step (8b) is the means-end belief and (8c) is derived from (8a) and (8b) by BC_W. Step (9) is meant to go from the belief that M to the intention that M. According to Bratman this is because you believe that M will obtain only if you intend M and so "if you do arrive at the belief that M this will normally be by way of intending M" (Bratman 2009, 32). Høj (2009, 6) breaks this into two stages: normally you would only form the belief that M by forming the belief that you intend that M, and normally you would only form the belief that you intend that M by forming the intention that M.

What Broome actually says is slightly different. He begins:

> Straightforward theoretical reasoning would bring you to believe you will buy a ticket... However, you are limited by a special constraint, which prevents you from doing straightforward theoretical reasoning. You believe you will not buy a ticket unless you intend to do so. Therefore, if you are rational, you cannot come to believe you will buy a ticket... unless you also come to believe you intend to buy a ticket. (Broome 2009, 14)

Since we are prevented from doing "straightforward theoretical reasoning" we cannot go immediately from (8a) and (8b) to (8c); we are prevented from forming the belief that M because we do not now believe we have the intention that M. If, however, the beliefs in (8a) and (8b) remain unchanged [as stipulated in (8)] then the only way to satisfy the rational requirement is to form this belief too. So the belief that M and the belief that I intend that M are formed as part of the same step of reasoning, so to speak, or to put it slightly differently, the conscious reasoning has two conclusion-states which an unconscious process ensures match each other so as not to violate any requirements or introduce any incoherence. This seems similar to the first stage described by Høj although it is not clear whether Høj has realised that the belief that M in (8c) is not caused to occur before the belief that I intend that M is caused. Up to the formation of these conclusion-states it is the theoretical requirement BC_W that captures the normativity of the reasoning.

Broome continues that in normal circumstances "you cannot acquire the belief that you intend to buy a ticket except by actually acquiring the intention of buying one" (Broome 2009, 14), which is to say that the process by which you come to form the belief that you intend that M normally operates by causing you first to intend that M. This seems plausible and similar to the second stage described by Høj. However, there is a dis-

analogy with what Høj says because Høj argues as if both stages are likewise ways of avoiding violation of a requirement of theoretical rationality; this is true of the first stage, but not the second. The connection between the belief that I intend that M and the intention that M is merely a causal dependency and not a normative constraint at all, whether theoretical or practical; it is not that an intention that M is rationally required for forming the belief that I intend that M.

The third part consisting of steps (10) and (11) is Broome's defence of the idea that IR_W can in some sense be derived from BC_W or more generally from requirements on theoretical reasoning despite the fact that a reasoner may be completely theoretically rational and yet be instrumentally irrational because failing to have the appropriate intention. Broome (2009, 17-18) says that when this happens there is no way of getting out of this situation by reasoning because the false belief blocks any such reasoning. You cannot, as Bratman would like to do, reason to an actual intention.[4]

Let us look at these three parts in order.

What is the relation between intending that E and the belief that you will E? We need the latter for use as a premise in (7) and it is questionable whether we are entitled to it. Broome's view expressed in (1) to (3) contrasts with the view that intending something entails that you believe you will do it. Bratman (2009, 21) calls this the Strong Belief Thesis and rejects it. There are two reasons why an agent may be reluctant to believe that E: because they believe that their attempt to carry out the intention will not succeed or because they believe that they will not carry out the intention at all, possibly because at the time the intention should have been carried out they will have ceased to believe (having forgotten) that they have the intention. The first might be described by saying "I will try that E," but the second is best described simply by saying "I intend that E (but may not)". It is the second that seems to be reported in (3) which says that you may well not believe that E because you fail to believe you intend that E although you do in fact intend it, and it is this that Broome wants to allow for by replacing the Strong Belief Thesis with what he calls the *belief-intention link* (Broome 2009, 12):

---

[4]   Bratman objects that we are not blocked because we can go through the same reasoning again and on this occasion the intention may be caused to occur. This is true, but a more charitable interpretation of Broome would ascribe to him the view that no process of reasoning can be guaranteed to result in an intention.

Belief–intention link:
If N believes that she herself intends that p, then, because of that, N believes that p.

If N has forgotten that she has the intention that E then she will not believe that E, but if she has not forgotten then this link implies that she must believe that E. However, this does not seem to cover the case where you may not believe that E because you think your attempt to E may fail. Broome seems to say that in such cases what you have is not really an intention but something weaker that is not subject to the same norms, but it will be shown that this kind of response is evasive. An equally evasive alternative is to say that it is true if the content of the intention does not involve a success-verb, such as would be the case if the intention were for a basic action (a bodily movement). Taking all the necessary steps for buying a ticket might then count as buying a ticket and we can be confident that we will do this even if we are not confident that, having taken these steps, we will be in the state of having bought a ticket, e.g., because there are no tickets left. But this is doubtful as well, for we use success-verbs in our reasoning and do not seem to be conscious of using them in anything other than their ordinary senses.

The reasoning process as Broome presents it consists of saying things to yourself. Saying to yourself (or another) "I will E" as an intention and not as a predictive belief expresses both the intention that E and the belief that you intend that E, since you are not in a position to say sincerely that you intend that E unless this intention is one you believe you have. In saying it sincerely you are also expressing the corresponding predictive belief [i.e., the belief in (7)], argues Broome:

> You cannot sincerely express an intention to do something without believing you have that intention. Consequently, when you sincerely express an intention, you must believe you will do what you intend. That is why you are in a position to express this belief simultaneously.
>
> The evidence I offer for the existence of this belief-intention link is just that expressing an intention is also expressing a belief; both take the form of saying the same indicative sentence. So you cannot express an intention without expressing a belief that you will do what you intend. And you cannot do that sincerely without having the belief. (Broome 2009, 12-13)

This is Broome's argument for the *belief-intention link*. The first sentence seems reasonable: if I do not believe that I have the intention that E, I cannot sincerely say "I will E" or even "I intend that E," and I cannot engage conscious reasoning on the basis of attitudes that I do not believe I have.

How "consequently" I must believe that E does not seem at all obvious. The appeal to sincerity suggests that perhaps Broome believes that it is a sincerity condition of the speech act of expressing the intention that E that one believes that E, that a speaker or reasoner would not be entitled to say this unless they believe they will do what they intend to do. This would be an attractive solution if there were such a sincerity condition. Unfortunately there is not: the sincerity condition for expressions of intention and for commissives in general is simply having the intention and not any belief that the intention will be fulfilled (cf. Searle 1979, 14). Nor is it obvious that we do always express an intention and a predictive belief by the same indicative sentence (perhaps Broome believes this by generalizing only from examples that do not involve success-verbs); certainly we do sometimes but it does not seem well-motivated to limit the account to those cases where we would, i.e., intentions for which the *belief-intention link* is true. The norms we are trying to account for seem to apply equally to those cases we would not express by saying "I will E"; whether we call such tentatively held attitudes 'intentions' or not is less important than whether the norms at issue apply to them. Bratman and Høj say that they do and I am inclined to agree that this is correct. However, I think this is less of a problem for Broome's model than may be thought. Broome does not need to limit his account to intentions that satisfy the *belief-intention link*; reasoning with weak intentions is simply to reason *ex hypothesi*, just as we may construct and follow arguments whose premises we believe to be false. We do not need to actually be sincere in order to reason as if we are.

To say that an intention is weak is only to say that one is more easily disposed to cease having it, in which case, under normal circumstances, one will cease to believe that one has it, and, if the *belief-intention link* is true, this will mean that I will not believe that it will be fulfilled. Let us look at the wording of (8): "If you also believe that E only if M, and *if these beliefs do not change*, BC$_W$ requires that you believe M." The grammatical form is deceptive, for it makes BC$_W$ look like a narrow-scope requirement. In fact, it is a wide-scope requirement, viz. "BC$_W$ requires that (if you also believe that E only if M, and *if these beliefs do not change*, BC$_W$ requires that you be-

lieve M)." So this conditional is satisfied if the beliefs do change due to weakness of the intention. In such a case we would not have (8c) as the conclusion-state, but the model never actually said that you would. If they do not change, then the weakness of the intention is immaterial – a weak intention that does not change is on a par with a strong intention that does not change. If this is correct then the problem of weak intentions is a pseudo-problem.

However, this presupposes that the *belief-intention link* is true, and we have seen no convincing argument for this yet. The linguistic fact that we often express both intentions and predictive beliefs as "I will E" is not convincing, yet I think there is something right about what Broome says. Certainly there is something at least a little paradoxical in saying "I intend to return the library books, but they will not be returned," and this seems to be because if I say that I will do something, I usually do so with the intention to make it known to the listener that they are entitled to expect its being done – I am not merely describing my own state of mind but making a kind of normative commitment. It is not entailed that it will be done or that I believe that it will be done – Broome is wrong to present the link as an entailment – but there is an implicature. Like all implicatures it can be cancelled, as we have already seen, by saying something like "I intend that E but I may not succeed".

The interesting thing is, if we accept Broome's idea that reasoning amounts to saying things to oneself, an implicature is enough to get to step (7), and once we have got that far, it is conceded that the reasoning in (8) and (9) is valid. The relation between intending that E and the belief that you will E is the implicature from "I intend that E" to "E" or "E will be done." In turn, this means that when we say to ourselves "I intend that E" we are entitled to say "E" unless we cancel this implicature. In saying this we can still be counted as reasoning if reasoning is a process of saying things to oneself, despite the lack of any deductive entailment.

Has Broome then succeeded in giving a cognitivist account of the instrumental requirement? There are still serious problems. There are timing issues, for surely I am already instrumentally irrational if I do not have coherent intentions, even before I have the belief that I intend the end and even if I never form the belief that I have this intention. Furthermore, the pressure to make myself instrumentally rational should begin as soon as I have this intention and the belief that I have the means to achieve it and not, as Høj (2009, 13-16) says of Broome's model, only after the intention

and the means-end belief have been expressed or even later at the last moment that the means is still effective as a means. Although I am not necessarily irrational when I leave taking the means "to the last moment" it seems that I am under some rational pressure to take the means before then. Lastly, suppose that I do have the intention to take the means yet I do not have coherent beliefs, which is to say that I satisfy IR_W but not BC_W. If I then lose that intention, then surely this is where I begin to be instrumentally irrational and under pressure to remedy that problem, and (says Brunero 2009, 316-317) this is an additional problem and an additional pressure to that due to the fact that I am *already* violating a requirement of theoretical rationality.

Broome has already conceded that there are cases where the theoretical requirements are satisfied and instrumental requirements are not (because the belief that one intends the means is false) and, presumably, also vice versa. Now it seems he must also concede that even in cases where both types of requirement are satisfied, they can be satisfied at different times. Now, we could say that we are not aware of being in violation of the instrumental requirement unless we are conscious of being in an incoherent doxastic state, where this is brought about by being in violation of a theoretical requirement, and it is only then that we can purposely get out of this state by reasoning. Here we appeal to the idea that the normative force of the instrumental requirement can be captured by the normative force of theoretical requirements even if the instrumental requirement itself is not captured by theoretical requirements. Yet it might still be objected that we should not be in a state where we intend an end yet do not intend what we believe to be our preferred means to that end; by being in this state we are violating a rational (state-)requirement. This is a special case of the wider assumption that it is irrational to have incoherent intentions just as it is irrational to have incoherent beliefs. Is this assumption true?

I do not believe that intentions have to be coherent in the same way as beliefs do. This is because I believe that any change to the strength or content of any practical, goal-oriented attitude must be explained by the fact that success at achieving the goal is more likely after that change than before. Let me put it this way: suppose that I form the belief that my goal (moving a log, for instance) is going to be harder to achieve than originally thought. I will be less likely to say, if asked what I am doing, that I am moving the log. I will likely say that I am *trying* to move the log. But what am I actually doing differently from when I would have said "I am moving

the log"? More to the point, what am I actually thinking differently? Has my intention changed to a mere trying, something that is less than a full-blown intention? No: the intention or goal-oriented attitude has the same strength as before. Has the content changed so that I no longer intend to move the log but only intend to try to move the log? No: I would not achieve my goal by reaching the state of having tried. No change of these kinds would make it more likely that I succeed in moving the log. Of course, I may decide that moving the log is so difficult that I abandon it as a goal; this kind of change obviously would not be explained as making achievement of the goal more likely, but this is a different matter. Call this the Evolutionary Advantage Constraint. Making changes to your intentions to make them more coherent will sometimes go against that constraint; it will put you at a disadvantage.

One of the clearest cases is the natural thought that you cannot have inconsistent intentions. I believe that you can have inconsistent intentions. By this I do not mean that one can have directly contradictory intentions: it is more questionable whether one can have the intention to x and the intention not to x at the same time without being irrational (though I will give a scenario where I think this is in fact possible). What I claim is that I may have the intention to x and the intention to y even when I believe that I cannot achieve both x and y together. I believe this is the case in Bratman's video game scenario. In this scenario a person is playing two games simultaneously in each of which one is attempting to hit a target, but the two games are linked in such a way that if he succeeds in one game then the other game shuts down. In short, because of the way it is set up it is impossible for the person playing the game to hit both targets thereby fulfilling both intentions, and the person knows it. The person might describe what he is doing when he is playing the games as *trying* or *intending to try* to hit each target, but Bratman seems to think it would be wrong for them to say that they *intend* to hit each target because this will leave them in the irrational state of having inconsistent intentions.

On the face of it there are two ways in which we might unpack this "trying," two ways in which we might weaken the claim that in each case the person intends to hit the target, referred to above as changing the strength or the content of the attitude. One is to put "try" into the content of the intention, i.e. instead of I(hit the target) having I(try to hit the target). This seems a very unnatural way of doing it – it is not as if the end being aimed at is having tried (although having tried may sometimes be

aimed at, e.g., if what you intend to do next depends on whether you have tried. If you have tried and failed to x by y-ing then you should try to x by z-ing, but you should not intend to z before trying to y).

A more natural way to unpack "trying" is to take it as modifying the intention rather than its content, in other words instead of I(hit the target) having I*(hit the target). Bratman argues that the person does not really *intend* to hit the target (because this results in inconsistent intentions) in either case but has a weaker attitude towards hitting the target that is not subject to the same kind of consistency constraints (although still subject to norms of means-end reasoning). For the sake of convenience I will call this intending-to-try. However, one can ask what this really accomplishes. Should we introduce a distinct kind of attitude simply in order to avoid inconsistency when a full-blown intention is just as likely or perhaps even more likely to succeed? Suppose that one begins playing both games and only learns that it is impossible to fulfil both of one's intentions whilst playing. What advantage to achieving one's intended ends does one gain by modifying one's intentions into intentions-to-try? None at all. Perhaps one might, on learning this, think oneself better off concentrating on just one of the games, but surely one is not rationally required to do so.

What, then, does the person playing the game mean when they say that they will try or intend to try to x? Only that they intend to x but are not confident of success. "Trying" is simply the pragmatic means by which the implicature that we will x is cancelled; it means "I intend to x (but I may not x)".[5]

If this is right then either there is no rational requirement for intentions to be consistent, or there is no evolutionary advantage in being rational in this way and we are better off being irrational. I prefer the first horn of this dilemma; we are not always irrational when we are in the state of having inconsistent intentions.[6] But are we irrational when we are in the

---

[5]   Kolodny suggests something similar in (2008, 395 ff. 34).

[6]   Kolodny (2008, 377-378) comes to the same conclusion for the video game case for much the same reasons. He expresses the point in terms of maximizing expected utility. Interestingly, he says that beliefs must always be coherent because avoiding believing a falsehood always has a higher utility than believing a truth. Therefore, if you have contradictory beliefs you believe one falsehood and one truth, yielding a net loss in utility. But this is not always the case with intentions. Having one intention that can be fulfilled and one intention that cannot does not necessarily have lower expected utility than having neither of these intentions. However, it seems to be implied that if fulfil-

state of having an intention for an end and not having the intention to take the means to that end?

Suppose that, because it is the one and only means to some other end that we have, we have the intention not to y. That is to say $I(x_1)$, $I(x_2)$, $I(\text{not-}y)$, $\text{not-}I(y)$, $B(y$ is the one and only means to $x_1)$, and $B(\text{not-}y$ is the one and only means to $x_2)$ where x_1 and x_2 are the intended ends. Obviously x_1 and x_2 can never both be achieved since the means for one denies the means for the other – they are incompatible. But I don't think this means that we must drop $I(x_1)$ or $I(x_2)$; after all, we might think of another means, or if we fail in intentionally taking the means to one of our ends we may wish to switch our focus to the other intention. If I try to carry out my intention not to y but fail then I can form the intention to y or, since it is implied that I have already y-ed (albeit unintentionally) I will find myself in the state of having achieved x_1 unintentionally. Note that this state is not necessarily the state of having x_1-ed; one of the x_1's logical preconditions is false, namely that you are not already in a state that logically precludes the action: you cannot close a window that is already closed, you cannot climb up a mountain when you are already at its summit, you cannot raise an arm that is already raised as high as it will go.

Despite their incompatibility I do not think that it is irrational to keep both $I(x_1)$ and $I(x_2)$. Obviously, we would not make achievement of x_1 and x_2 more likely by dropping them as intentions or weakening them in some way. By extension, the same goes when we add in $I(\text{not-}y)$. We have yet to find normativity still requiring explanation.

However, supposing for a moment that intending the means to our end is the rational state to be in, then we are rationally required to have $I(y)$ as well as $I(\text{not-}y)$, and, as I have already said, although having incoherent intentions is tolerable having directly contradictory intentions is somewhat

---

ling one of the intentions has higher utility than the other then the other should be dropped on the grounds that by intending the other you make it more likely to fulfil that other and, ipso facto, less likely to fulfil the intention that you have more reason to have. I am not so sure of this. The Evolutionary Advantage Constraint does not, then, give precisely the same results as Kolodny's utilities-based account, although it is basically the same idea.

Incidentally, this may tell against Broome's (2007) use of deontic dilemmas to argue against narrow-scope formulations of requirements of practical rationality – we need not assume that the intentions detached from such narrow-scope principles are irrational because incoherent.

more questionable. If we are rational, then we are irrational. In this situation we are not always irrational when we do not intend the means to our end. Perhaps it might be objected that what this means is that we are irrational in keeping both $I(x_1)$ and $I(x_2)$ after all, since it seems that either x_1-ing or x_2-ing must be logically impossible due to a logical precondition being false. This is true, yet as long as we are unaware of which is logically impossible it does not seem to be irrational to be in this state. This is the normativity needing explanation, and it seems to be a distinctively practical normativity, for the cognitivist account is not nuanced enough to allow for cases where the instrumental requirement can rationally be violated, cases where it would be irrational to have the intention even though it is rational to believe that you have the intention.

I will claim now that we *ought* to have the intention for the means and because of this it is generally rational to have this intention, but sometimes it is rational not to have this intention, although this does not mean that the rational requirement to have this intention has not been violated.

Before the theoretical treatment, let me give an example. Suppose, on a cold day, my pet cat stands on the window-ledge, wanting me to open the window and let her in from the cold. I intend to let her in, and, as the means for letting her in, I intend to open the window. At the same time, I have an intention not to let cold air in, and this is what will happen if I open the window. I am in a bit of a dilemma. I decide to do the decent thing and let the cat in. Do I thereby cease to have the intention not to let cold air in, and, once the cat is safely inside, reform that intention and then intend to take the means by closing the now open window? This all seems rather unnecessary. It seems far better to say that I had the intention not to let the cold air in all along. What if the window has got stuck in the cold, that is to say, I intend to open the window and make a genuine attempt to open the window, but I fail? In failing, I have already succeeded in achieving my intention to not let cold air in, not by actually closing the window, it is true, but because the window is already closed. Does this mean that I would be irrational to have the intention of closing the window? There are two reasons why it might be thought so: a) because it contradicts my intention to open the window, and; b) because it is pointless to form intentions to bring about states that already obtain (viz., the window's being closed). On the other hand, it is nevertheless the means to my end of not letting cold air in, and generally it is rational to intend the means to ends that you have and irrational not to. Also, should the window become open,

should it be necessary to form the intention to close it? Wouldn't it be better to say that this intention was already there but dormant in some sense? We do not need to imagine intentions vanishing and being reformed depending on whether the window is opened or closed; this would be a cognitively uneconomical way of doing things. All we have to do is recognize there is more than one way in which an intention can be satisfied – it can be successfully carried out or it can be, I shall say, *escaped*.

Oughts are things that we can satisfy either by doing what we ought or by escaping it. Schroeder (2004, 252-253) gives an example of Al who does not wish to keep his promise to meet Rose for lunch and so asks her to release him from his promise. If she does then it is no longer true that he ought to keep his promise since there is no such promise anymore; it is not that he has kept his promise by being released from it. Rather, he has *escaped* it. Because it does not predict this asymmetry between satisfying an obligation and escaping it, a wide-scope requirement to do what you promise to do is inadequate, Schroeder argues.

The idea I want to examine is whether we can reduce the instrumental requirement to a special case of the requirement that you ought to do (or intend to do) whatever you believe you ought to do. This is often called the *krasia requirement* or following one's conscience. Like a promise, the normative content of the belief generates an obligation to have certain intentions, but the *krasia requirement* does not go so far as to say that you are irrational all-things-considered if you do not have these intentions. There may be good reasons, reasons that can be offered in excuse or mitigation, for not having the intention one ought to have, which reasons may include an opposing obligation generated by another normative belief, another end. One can have incoherent oughts, but it is not the case (simply because it is impossible) that you ought to satisfy both oughts; sometimes it is more rational to violate the *krasia requirement* than to satisfy it. Put another way, violations of local rational requirements are tolerable and often necessary to satisfy global requirements that determine ultimately what is rational all-things-considered. It is a mistake, however, to suppose that because it is rational to violate a rational requirement the requirement is not really a requirement or not really exceptionless. The reason we give for not having the intention one ought to have functions by opposing its normative force to the normative force of the reason for having that intention – it does not function descriptively as an explanation of why the *krasia requirement* does not apply in this case. The *krasia requirement* does apply, and is violated;

the 'ought' involved is local but not defeasible, and these two things should not be mixed up.

Schroeder's version of the narrow-scope *krasia requirement* is:

Narrow ConscO: If you believe that O(you do A), then O(you do A).[7]

---

[7]   Schroeder rejects this principle, appealing to the bootstrapping objection: it cannot be the case that you ought to do something just because you believe you ought to – believing something does not make it so unless the belief is infallible, which it is not. If your belief that you ought to do A is false then by definition A is not something that you ought to do and so a principle that says that you ought to do A [i.e., that concludes O(you do A)] must be false. He prefers another narrow-scope principle formulated in terms of subjective reasons, viz.,

Narrow ConscSR: If you believe that O(you do A), then you have a subjective reason to do A.

If your belief that you ought to do A is false it is not the case that you cease to have a subjective reason to do A. For this reason Narrow ConscSR is to be preferred to Narrow ConscO (cf. Schroeder 2004, 349-351).

Against this, I would argue that when the belief is true you do not satisfy the requirement just by having the subjective reason. Suppose that this reason to do what you ought to do is defeated by another reason subjectively stronger. I am not saying that in such a situation it would be wrong to act on this other reason. I am saying that this other reason must be offered as mitigation for violating the requirement, but this can only mean that the requirement is violated despite having the subjective reason Narrow ConscSR says that you should have.

It seems that not everybody shares this intuition. Kolodny's intuition (2008, 379) is that no requirement is broken at all in having incoherent intentions, rather than that the requirement is broken but overridden. I am not sure this is true for the krasia requirement; the necessity to give a reason for not doing what you intend suggests to me that there is a requirement in place. I am tempted to say the same of Brunero's (2010, 34-38) interesting example of the girl who believes she has conclusive reason to go to the lecture but does not believe that she has sufficient evidence for this belief. Brunero says that although it is not irrational to form the intention to go to the lecture, she is not required to form this intention; it is not that there is a requirement she has better reason to comply with. My intuition about this situation clashes with Brunero's. I think that she does violate the krasia requirement when she fails to have the intention to attend the lecture but that she has reason to do this, and this is shown by my intuition that she is obliged in some way to explain why she has not formed the intention, and the reason performs this function because it has a normative force that conflicts with the normative force of another reason and not because it simply describes a situation where the krasia requirement does not apply or provides a counter-example to this particular formulation of the requirement.

Now, consider the situation where Al finds himself in the right time and place for the promised lunch with Rose but did not appear at this time and place in order to keep his promise, and further that he knows this. It seems wrong to say that Al has satisfied his promise or that he has done what he ought to do, yet he seems to have put himself in a situation where it is actually impossible for him to keep his promise. It makes no sense to ask Rose to release him from his promise. He has escaped his promise, but not through Rose's agency. The obligation to keep his promise is void unless he gets up to go somewhere else before Rose arrives, in which case it reappears. Similarly, if my left arm is already raised it cannot be the case that I ought to raise my left arm, and if I believe that my left arm is already raised then I can no longer believe that I ought to raise my left arm. I have escaped the *krasia requirement*.

What I would like to claim is that the oughtness involved in the *krasia requirement* is transmitted from ends to means. If I ought to x_1 (let my cat in) and to have the intention to x_1 then I ought to y (open the window) and to have the intention to y, and if I escape from the requirement (e.g., my cat is already in and it is my neighbour's cat on the window-ledge) so that it is no longer the case that I ought to x_1 and to have the intention to x_1 then equally it is no longer the case that I ought to y and to have the intention to y. The reason why I may be in an irrational state when I do not have the intention to take the means to an intended end is because the means inherits from the end the feature that it is something that ought to be done.

This does not rule out the fact that taking the means may have consequences that are things you believe you ought not to do. It is quite possible that you ought to y and ought not to y. Consider again x_1 and x_2. These are both things we believe we ought to do (and are assumed to be, and the states corollary to their fulfilment are assumed to be, equal in desirability). According to the *krasia requirement*, they are things we ought to do, and this implies that the intentions to take the means are intentions that we ought to have and that we are rational when we have them. However, suppose that the logical precondition for x_1 is not satisfied because we are already in the state that would result from x_1-ing had we been in a different state (my cat is already in). In other words, we *escape* the *krasia requirement* for x_1; there ceases to be any obligation to have $I(x_1)$ and, thence, any obligation to have $I(y)$ (which does not necessarily mean that we cease to have those intentions or that we are rationally required to cease having those in-

tentions). Yet at the same time that I(y) ceases to be obligatory it becomes a reason not to have I(not-y), on the grounds that we are already in one of equally desirable states and not-y-ing will not change that situation. Although having I(x₂) and I(not-y) satisfies the *krasia requirement* for x₂ and is rational in so far as it does, this is one of those cases where it is more rational to violate the *krasia requirement*. Intentions that one ought to have (in the sense of 'ought' being used here) are not necessarily the ones it is most rational to have. This implies that, in this situation, when one becomes aware of what state one is in it is more rational not to have the intention for a means that will take you out of that state than it is to have that intention, irrespective of the fact that one violates the *krasia requirement* in not having that intention. In other words, if it were not for the fact that letting the cat in was considered by you to be the decent thing to do and you were indifferent between that and not letting cold air in, it would be more rational to not let cold air in, on the grounds that this is what you are already doing and there is no motivation to change. Having contradictory intentions is rationally preferable when you do not know which of the states you are in (the normative forces of the reasons being in dynamic equilibrium), but ceases to be when you do know what state you are in.

2. Conclusion

The Strong Belief Thesis is false. The *belief-intention link* is false. Yet something like the *belief-intention link*, with a pragmatic implicature in the place of a material implication, is true, and, perhaps surprisingly, this is enough to provide the reasoner with the predictive belief that he will carry out his intention successfully as a premise, even, also perhaps surprisingly, when he does not believe that he will carry out his intention successfully. Our reasoning is not limited to what we sincerely believe or intend but can be carried out *ex hypothesi*, as it were. This gives Broome's model more scope and versatility than Bratman and Høj, and probably Broome himself, thinks.

Practical reasoning as such ends in the belief that I will fulfil my means-end intention and the belief that I intend to carry out the means. Providing (as the conditional says) that my premise-beliefs remain unchanged this is the only way to make my beliefs coherent. This does not, however, necessarily make my intentions coherent – satisfying the requirements on the

beliefs does not guarantee satisfying the requirements on the intentions, which is to say that the *instrumental requirement* does not have the same extension as the theoretical requirements. Nevertheless, it is true that it is only by satisfying the requirements on the beliefs that we can by reasoning bring ourselves to satisfy the *instrumental requirement* and only, I argued further, by violating the requirements on the beliefs that we can be led to believe (sometimes falsely) that the *instrumental requirement* is likewise violated. Because of this I tentatively suggested that the normative force of the *instrumental requirement* could be given in terms of the normative force of theoretical requirements even though they differed in extension, and this was sufficient for a broadly cognitivist account of the *instrumental requirement*.

This left a residuum of normativity to account for – a distinctively practical normativity – that at least sometimes one is in an irrational state if one has an intention for an end but not the intention for what is believed to be the means to the end, but also that sometimes this is not an irrational state, even if rationality requires you to believe you have this intention. I held as a constraint on this normativity that it must comply with what I called the Evolutionary Advantage Constraint. Paying close attention to this assumption I came to the conclusion that one is not necessarily irrational when one has incoherent and even directly contradictory intentions, so norms against such combinations of intentions could not explain the irrationality of at least some violations of the instrumental requirement, for there were no such norms.

The explanation, I urged, is a set of obligations to have certain intentions. When intended ends are not incoherent or incompatible there is an obligation to intend their means also, and so one is not rational if one intends to E, believes that M is the preferred means to E, and not intend M. But if an intended end is incoherent or incompatible with another end then it may be rational to intend the means to both even if this means having intentions that are directly contradictory. In the kind of incompatibility I considered there is always one *krasia requirement* that could be escaped, and which intentions it is rational to have depends on which end has higher priority or which is reflected in the current state, it being irrational to have the intention to take any means to leave a state unless this is the means to a state *more* desirable than the current one. This is to say that it can be rational to violate a requirement of practical rationality, viz., the *krasia requirement*, when one has reason to do so. Not all incompatibilities are

like this, but ones that lead to directly contradictory intentions are. Ones that lead to only inconsistent intentions are not irrational.

The *krasia requirement* has here been put forward as a genuine requirement of practical rationality and can account for our intuitions in the one case of the *instrumental requirement* that Broome's account did not seem to account for, the one place where there seemed to be a distinctively practical normativity.

References

BRATMAN, M. (2009): Intention, Belief, and Instrumental Rationality. In: Sobel, D. – Wall, S. (eds.): *Reasons for Action*. Cambridge: Cambridge University Press.

BROOME, J. (2006): Reasoning with Preferences. In: Olsaretti, S. (ed.): *Preferences and Well-Being*. Cambridge: Cambridge University Press.

BROOME, J. (2007): Requirements. In: Rønnow-Rasmussen, T. – Petersson, B. – Josefsson, J. – Egonsson, D. (eds.): *Homage à Wlodek: Philosophical Papers Dedicated to Wlodek Rabinowicz*. Lund: Lund University.

BROOME, J. (2009): The Unity of Reasoning? In: Robertson, S. (ed.): *Spheres of Reason: New Essays in the Philosophy of Normativity*. Oxford: Oxford University Press.

BRUNERO, J. (2009): Against Cognitivism about Practical Rationality. *Philosophical Studies* 146, 311-325.

BRUNERO, J. (2010): The Scope of Rational Requirements. *The Philosophical Quarterly* 60, No. 238, 28-49.

HØJ, J. B. (2009): Problems for Broome's Cognitivist Account of Instrumental Reasoning. *Acta Analytica* 25, No. 3, 299-316.

KOLODNY, N. (2005): Why Be Rational? *Mind* 114, 509-563.

KOLODNY, N. (2008): The Myth of Practical Consistency. *European Journal of Philosophy* 16, No. 3, 266-402.

SCHROEDER, M. (2004): The Scope of Instrumental Reason. *Philosophical Perspectives* 18, 337-364.

SEARLE, J. (1979): *Expression and Meaning*. Cambridge: Cambridge University Press.

The Practical Authority of Normative Beliefs: Toward an Integrated Theory of Practical Rationality

ROBERT AUDI

Department of Philosophy. University of Notre Dame
100 Malloy Hall. Notre Dame, IN 46556. USA
Robert.Audi.1@nd.edu

ABSTRACT: People who do not act on their judgments regarding what they ought to do—their practical judgments—are often considered weak-willed, especially when the judgment is made at a time when the act it favors is plainly possible. Is this a kind of practical irrationality, perhaps due to an incoherence between practical reason, which should guide intention and action, and behavior that fails to conform to a guiding directive? More generally, do normative beliefs with the same kind of self-directive content as practical judgments possess the same sort of rational authority, if indeed they must have any such authority? At least since Aristotle, weak-willed action has often been considered irrational. This paper indicates why that view is plausible, but also why it is too strong. The paper shows how the practical authority of normative judgments can be overridden and why, on the theory of rational action suggested by analysis of such cases, certain initially plausible action-guiding principles are too strong. The concluding part of the paper briefly indicates how that theory can do justice to the analogy between practical and theoretical reason and to the essential connection between the two.

KEYWORDS: Coherence – *Enkrasia* – grounds – intention – instrumentalism – integration – judgment – practical reasoning – rationality – reasons – weakness of will – well-groundedness.

Our knowledge of other people's mental states is at least largely based on our perceptions or knowledge of their behavior, and in normative matters it is especially natural to think that "actions speak louder than words."

It certainly seems that, usually, when people say they ought to do something and, given an opportunity, do not, we tend to wonder whether they really believe they ought to do it. It is also natural, moreover, to think that people tend to feel motivated to act on their practical judgments. Taken together, these widely assumed ideas incline many moral philosophers and action theorists to think that, at least normally, if one judges that one ought to do something, then one intends to do it. That rationality requires conformity to this principle is an *enkrasia requirement*. The main action-theoretic question here is the relation between such self-addressed judgments – and I assume the focus is *overall* judgments – and motivation, above all intention.[1] The question may be broadened to include normative beliefs as distinct from normative judgments, and I shall address both. My major concern, however, is with a related question in the theory of practical reason: Does rationality require conformity to an "*enkrasia* requirement" such as this: if one judges (or even simply believes) one ought to do something, then one intends to do so (Broome 2010, 290)? This question is important for the overall theory of rationality, which concerns not only rational action but also what I call *global rationality,* the overall kind that encompasses both practical and theoretical rationality. Roughly and somewhat metaphorically, global rationality is a matter of the rationality of both intellect and will.

1. Weakness of Will, Practical Judgment, and Rational Action

My point of departure is the nature and status of that perennial challenge to the theory of action, weakness of will. There is at best limited agreement on what weakness of will is, but a common element in at least most of the plausible accounts is the notion of action against one's better judgment. To act against one's better judgment – "incontinently," for short – is roughly to do something intentionally, such as take another drink, while in some sense aware of one's judging that doing something else would be best (or, on some accounts, better). More explicitly, we might plausibly adopt the following rough criterion of incontinence:

---

[1]   Overall judgments are sometimes called "all things considered" judgments. But here I avoid such terminology, in part because not all such judgments emerge from or are even preceded by *consideration*.

An agent, S, acts against S's better judgment at time t, in A-ing, if and only if, at t (1) S A's intentionally (or at least knowingly),[2] (2) there is at least one other action (type) B which S takes to be an alternative and with respect to which S has judged, or makes or holds a judgment, that it would be best (or that it would be better) to B, (3) S has not abandoned this judgment, and (4) S is aware (under some appropriate description) of (2) and (3).

The relevant kind of awareness need not involve entertaining the judgment; but where S is plausibly said to act *against* S's better judgment in the way required for incontinence, there must be a sense in which S is aware of holding the judgment.[3]

If action can be incontinent, surely intention can be. My will can be weak when I form the intention to have another drink even if the bartender wisely refuses me. Thus, incontinence extends beyond action. Intention is of course not action, and even forming an intention is at least not typically action. That intention and its formation can be incontinent should not be surprising. After all, if action should be responsive to practical judgment, so should intention, which is commonly a route to action and is rational for a person under conditions at least approximately equivalent to those applying to the act intended. Incontinence may also be instantiated by omissions;

---

[2]   For purposes of this formulation, the difference between knowing and believing is not crucial; moreover incontinence may be manifested by a "side effect," as where S judges S must not offend X, believes that nominating Y will do it, and intentionally nominates Y, truly believing but not knowing that S is thereby offending Y, and doing the latter *non*-intentionally (though not *un*intentionally). Both actions, of course (or one under both descriptions), may be incontinent.

[3]   The account as stated does not preclude compelled actions, as Audi (1990) does. Perhaps it should be revised to do this. Suppose S is tortured for information about fellow soldiers and, after two days of excruciating pain, gives it against S's better judgment and with a deep sense of guilt. Here S satisfies (1) – (4) yet acts under compulsion and is not plausibly said to have exhibited weakness of will. Granted, there is some degree of weakness *in* S's will as compared with the will of the very bravest. But is the highest humanly possible level of strength of will the standard for determining whether an act is *weak*-willed? Very strong people – and wills – need not be the strongest there are. Given normal standards for weakness *of* the will, as opposed to lacking the strongest of wills, S has held out longer than most normal agents would and even then acts quite *unwillingly*, though, it would seem, not incontinently. This paper does not turn on the difficult question whether compulsion rules out incontinence, and I leave it open.

these may be actions as opposed to mere non-performances. If one judges that (despite risking retaliation) one must speak in support of a friend but fails to form the intention to speak so, is this not incontinence? Here, as with "positive" action, the agent fails to respond to a self-addressed directive that apparently has, as its proper function, guiding the will.

The conception of weakness of will just articulated makes it easy to see why, on a kind of consistency view of practical rationality, incontinent actions are prima facie irrational: they contravene one's judgment of what one should do and thus exhibit a kind of inconsistency between one's action and one's assessment – which is often backed by good reasoning – of what one's action should be.[4] The same applies, of course, to other forms of weakness of will, such as incontinent intention.

Perhaps, however, even practical judgments of the self-addressed, overall kind in question are not something a rational agent must obey. Why should our intentions and actions never deviate from what our practical judgments call for? This question of course concerns why there should be an *enkrasia* requirement – one calling for intention to conform to such practical judgments – at all.

We should first indicate whether the point of view of our inquiry is cognitivist. If so – and cognitivism is the position I take here – then there is (at least normally) a truth of the matter concerning what we ought to do, what is good, what is obligatory, and so forth for other normative notions indicating goals of action (we can leave open whether, as some philosophers deny, cognitivism entails realism). Second, we must indicate whether the rationality of practical attitudes is considered subordinate to normative standards, prominently including the good and the obligatory – two notions central in (respectively) axiology and deontology. I assume this, with the qualification that actions and, correspondingly, intentions are rational when, given the agent's rational beliefs or evidential basis for rational beliefs, the acts or intentions are reasonably aimed at the good or some other appropriate normative goal, such as achieving what is obligatory. This cannot be shown here, though it will be supported by some of our examples. In any case, few would deny that it is at least plausible to hold that if I do

---

[4]   Cf. Scanlon's claims that "A rational creature is, first of all, a reasoning creature – one that has the capacity to recognize, assess, and be moved by reasons" (Scanlon 1998, 23) and "Irrationality in the clearest sense occurs when a person's attitudes fail to conform to his or her own judgments" (Scanlon 1998, 25).

what I *rationally* believe will realize more overall goodness than any alternative I have, I do something that is rational for me. The same holds if we put overall obligation in place of overall goodness.

There is one quite general point that, without a great deal of elaboration, can be seen to support not only some version of the "*enkrasia* requirement" but also the quite general view that practical rationality depends partly on theoretical rationality – or at least on cognitive attitudes. This is the point that (intentional) action is aimed at bringing something about and that our aim in acting is either broadly intrinsic or broadly instrumental: we normally act to bring about something, such as enjoying a tennis game, either for its own sake or for a further purpose, such as getting exercise. But this strongly suggests that actions must be guided by beliefs: either instrumental or to the effect that what we do will have the property for which we want to do it, say being enjoyable.[5] In any case, if it is irrational to believe that playing tennis will be enjoyable, it will not be rational to play it for enjoyment. One might still rationally play it for exercise, but even that will not be rational if we cannot rationally believe that playing will constitute or lead to exercise. In this way, theoretical reason, as a source of knowledge and rational belief, has authority over practical reason in certain cases. Above all, if we rationally believe, or even have grounds sufficient for rationally believing, that an act will not realize the relevant end(s), then (on the view I am taking here) it is not rational for us to perform it. Positively, beliefs ("outputs" of theoretical reason) are normally needed to guide action to its end(s); this implies a kind of authority to tell us how we can (and often how best) to realize our aims. (Practical reason does not have a comparable authority over theoretical reason, but that is not the point here and will not be argued.)

2. Four Models of Rational Action

There are many ways to conceive the rationality of actions and intentions, even in the special case that concerns us, in which S holds a judg-

---

[5]   There is a possible exception: if I perform a simple basic action, such as raising my arm, for its own sake, must I believe, e.g., that *this* is how one does it (where my belief is indexical), or can my desire yield the action "directly"? If not, note that there is still a *potential* negative role for belief: if I believe, even falsely, that my arm is paralyzed, I might not raise it or even try to raise it, even when I have a very strong desire calling for that act.

ment (or at least a belief) favoring a specific act. One model, perhaps found in places in Plato's Socratic dialogues, *the executive model*, takes that cognition to be a normatively governing directive of practical reason. A second model, *the instrumentalist model*, is derived from a kind of Humean instrumentalism: on a generic version of this model, actions are rational on the basis of how well they serve the agent's non-instrumental desires. A third model, *the fidelity to premises model* – is based on the idea that what we have best reason to do, and hence are rationally required to do (and to intend), is what is favored by our *practical reasoning* (or at least the relevant premises) – at least in the very common cases in which we engage in practical reasoning on a matter before acting in that matter. The fourth model (developed in many works of mine) is holistic; call it *the well-groundedness model*: actions and intentions are rational on the basis of how well supported they are by the agent's relevant normative grounds, which prominently include reasons for action. In framing a general conception of rational action (one applicable to intention as well), I will consider each of these models in relation to examples that bring out their contours.

I want to begin with an extended example:

> Consider John, a practiced and conscientious retributivist. He believes that he should punish his daughter for talking hours on the phone when she knew she should study. On reflection, he judges that he should deny her a Saturday outing. But a day later, when it comes time to deny her the outing, he looks into her eyes, realizes that she will be quite upset, decides to make do with a stern rebuke, and lets her go. He feels guilty and chides himself. It is not that he changed his mind; he was simply too uncomfortable with the prospect of cracking down. Suppose, however, that he also has a strong standing belief that he must be a reasonable parent and is well aware that the deprivation would hurt the child and cause a rebellious reaction. He might be so disposed that if he had thought long enough about the matter, he would have changed his mind; but that is perfectly consistent with the assumption that if his will were stronger, he would have punished her. Thus, his letting her go may still be incontinent. But it is irrational? I cannot see that it is. (Audi 1990, 276-277)

In evaluating the case, we might first note that John's action is not a case of "passional" incontinence – perhaps the most typical kind – or in any other way tainted by appetitive influences. Moreover, it is backed by, and

hence coheres with, good reasons rooted deep in his character, for instance his desire not to hurt his daughter and not to provoke a rebellion so severe as to undermine the good moral effect of the punishment. In addition, he may have rational doubts about the retributive view on which his practical judgment is based, though he has not given up that view. But even though his incontinent action accords with a civilized and generally admirable compassionate desire, it *does* go against his standing better judgment and its underlying retributive beliefs and desires. Still, the overall rational basis of that judgment is too narrow and is outweighed by the larger rational considerations producing the incontinent action and apparently rendering it rational.

The example indicates a defect in the executive model: it shows that, and how, the normative authority of a practical judgment may be overridden. This is not to say that it is entirely eliminated; the case requires overriding considerations sufficient to render the action rational, not denying that the practical judgment provides *any* reason. On the instrumentalist model, the example may or may not be conceived so that it undermines the executive model. Take first the undermining case. John's merciful action may maximize the satisfaction of his relevant non-instrumental desires. Since John is conscientious and, we may assume, a loving father, we may suppose that, regarding his daughter, he above all wants what is best for her and has other normal parental desires. All of these together can outweigh the desires that support his retributive judgment (or belief) and the punitive action. In coherence language, the action better coheres with his overall belief-desire system than with his practical judgment. We again have rational action against one's better judgment, contrary to what the executive model requires.

The instrumentalist model does not, however, adequately account for cases like John's. Suppose John had been extremely angry, to the point that his non-instrumental desires would be best satisfied by punishing. We may also suppose that he has an irrational though well rationalized belief that the punishment will be good for his daughter, so that even his non-passional desires in the matter also favor the punishment. Where this combination of desires is responsible for his practical judgment, punishment might both accord with the executive model and also maximize satisfaction of his relevant non-instrumental desires. Administering the punishment, then, like intending it, is instrumentally rational for John – in the sense appropriate to our Humean model. But we need not take it to be rational on balance, and we should reject the instrumentalist model in question. To

be sure, Hume himself stressed the calm passions;[6] with that and related points in mind, one might constrain instrumentalism in the way Brandt (1979) and later writers have done. But the point remains that a broadly instrumentalist desire-satisfaction account of rational action too often gives the wrong results in cases like that of John.

More can be learned from noting that the instrumentalist model (like the others) allows, though it does not require, that John have *reasoned* to his practical judgment. As John's case is described, he need not have reasoned to that judgment or a conclusion favoring punishment. If he had, however, he might well have come to the merciful conclusion that goes best with the rationality of his relenting action. Reasoning about action can reverse initial desire-based behavioral inclinations. Reasoning is indeed perhaps likely to do so when these inclinations discernibly oppose the values deepest in our motivational and cognitive character. In suggesting that John is not rationally required to punish his daughter and that his merciful abstention is rational, I am not denying that he might be *excusable* if he did punish her – an important normative point – but the kind of case the example highlights apparently shows that, overall, practical reason favors the merciful route he took.

These points about the commonly salutary effects of practical reasoning might seem to support the third model. But surely we cannot plausibly claim that the conclusions of practical reasoning, such as practical judgments, much less the actions based on it, are invariably rational. Reasoning from an irrational belief is one factor that can prevent the reasoning in question from giving rational support to an action it favors. The same limitation applies to irrational desire: if, depressed at a professional setback, an author irrationally but predominantly wants to burn a carefully written, competent manuscript, we may reasonably urge delay, and he may rationally agree to it. Nonetheless, the contents of our beliefs and desires can figure in determining the premises of our practical reasoning even when those beliefs and desires are irrational – or at any rate normatively defeated by better grounded elements in our psychology. The fidelity to premises model, then, is unsound. Just as one can be faithful to the wrong person, one can act (or believe) on the wrong premises.

---

[6]   Hume (1739-40/1978, 418) says, e.g.: "Men often act knowingly against their interest: For which reason the view of the greatest possible good does not always influence them ... What we call strength of mind, implies the prevalence of the calm passions above the violent."

Reflection on the fidelity to premises model brings out something quite general. Our premises in a piece of reasoning are not our ultimate grounds for action; our premises themselves need adequate grounds and may still not be justified by such grounds. The well-groundedness model is designed to take account of this point as well as the multiplicity of relevant grounds we may have for action or intention. Might that model avoid the defects we have seen in the others? I believe that it can at least capture what is plausible in them. Unlike the other models, it calls for a holistic assessment of action: a rational action is conceived as one that is based, in the right way, on sufficiently good normative grounds, regardless of whether they figure in premises leading to the making of a judgment that favors that action. These grounds may be numerous and diverse. Actions based on them will at least strongly tend to cohere with them, but this is to say that the model takes account of such coherence considerations; it is not itself a coherence model.

Here are three broad points central to the associated conception of (objectively) rational action:

First, reasons and grounds must meet minimal objective standards (e.g., in terms of both the quality of the agent's evidence for the belief(s) involved and the degree of support relevant grounds provide for the agent's cognition or motivation). It is not enough for rational action that the agent believe the action is best, or believe that it is supported by the best reasons, or that it will yield optimal results, or the like. If the belief or desire is irrational, as where it is produced in a certain way by brain manipulation or by unfounded fears, then it does not imply the rationality of the action (whether the type or the token) for the agent in question.

Second, in order to render an action rational, the reason(s) or ground(s) must motivationally explain the action. A reason we have for an action may provide a *rationalization* for it without explaining why we are performing it, but (on my view and on most other views on the matter), reasons that do not, at least in part, motivationally explain an action do not render its actual performance rational. They are like a vertical column just tangent to the bottom of a bridge span but bearing none of its weight: in both cases, the ground exhibits potentiality for support but provides none of it.

Third, even beyond these requirements, the reason(s) must meet adequately high standards. A reason that renders an action based on it rational must be adequate to bear the normative burden of conferring rationality on the action. There may also be a coherence requirement – or what might better be called an incoherence requirement: incoherence of certain kinds

defeats rationality. The action should be based on grounds sufficiently (if imperfectly) harmonious with the agent's overall framework of beliefs and desires, since action based on reasons – such as certain fleeting emotional desires one disapproves of – discordant with the agent's overall makeup, is, even if not *ir*rational, not clearly rational. To take a case similar to that of John, suppose Maria wants above all to be moral, and that this (presumably rational) desire expresses a carefully considered set of ideals and principles to which she is single-mindedly devoted. A clearly rational action on her part would be one she performs on the basis of this desire and in the reasonable belief that the action is her overall obligation. If, however, her governing desire and her belief that it requires the action were not rational, the action would be rational only in some subjective sense. If, by contrast, they are rational but do not explain the action, which she performs, say, for a foolish reason, then it is not well-grounded by them, though it is rationalizable by appeal to the reasons they express. They make it rational for her *to perform* an action of that type but do not render rational her *performing* an action of that type. Doing what it is rational to do does not entail rationally doing it.

3. Belief, Judgment, and Enkratic Principles

So far, I have distinguished between judging, and simply believing, that one ought to do something. Simply believing that I have my car keys does not imply judging this. Judging that p, by contrast, does entail believing p, at least if we consider making a judgment in the common sense that implies coming to *hold* it (at least for a time), as opposed to the *activity* of judging a proposition, which may of course lead to judging it false and thereby rejecting it. Given the spontaneity and scriptedness of so much of our action, it is important to avoid intellectualism in the theory of action. We reach for a key in order to open a door; no practical reasoning is entailed. We enter a meeting, take a seat, and greet our neighbor as if we had internalized a script; no practical reasoning is needed. How much behavior is thus "scripted" is a contingent matter, but it can be a considerably complex sequence of deeds.

To be sure, even reasoning to a conclusion does not entail *reflecting* on the matter it concerns, nor does every case of reflecting on something entail reasoning about it. But both reflection and reasoning processes can lead

us to new grounds for action, belief, desire, intention, or all four. Some of these grounds are intuitions; some may be propositions inferred from what we believe; some may be emotions. The passage of time allows for thoughts, inferences, and, where there is enough time, reflection. Much can occur in the mind even in a short interval. This is a main reason why principles concerning reasons and rationality must be temporally qualified. To see this, compare two principles. A plausible *moderate enkratic principle* is:

> If, at t, S believes S ought (on balance) to A, but does not intend to A, then, at t, S is prima facie irrational.[7]

This is synchronic and does not entail the following related *diachronic principle of enkrasia*:

> If, at t, S judges, forms the belief, or appropriately considers S's already held belief, that S ought on balance to A, then, at the "next moment," S should *form* the intention to A if S does not have it or, if S does, then S should retain it rather than cease holding the belief.

In practice, then, if, in the important case of considering what to do, we arrive, with or without reasoning or reflection, at the belief that we ought on balance to A, a moment may (and I think typically will) pass *before* we form the intention to A. Here we may immediately see the prospect of A-ing as aversive, and rationality may call for reconsidering the matter rather than forming the intention.[8] Moreover, as the case of John shows, supposing we do form the intention to A, we may still rationally give it up in response to adequate grounds. This is possible even if their good influence does not (as it would naturally tend to) unseat the normative belief favoring A-ing.

---

[7]   For Broome (1999), a conditional like this is material, and the rationality requirement has (as I intend here) wide scope; it is thus that *either* one cease holding the belief or have the intention in question. It is not implied, moreover, that even unrepudiated intentions generate reasons for action.

[8]   This point is not uncommonly overlooked, e.g. apparently by Niko Kolodny in one place: he says, speaking of rational requirements in general (practical as well as theoretical), "When a person satisfies the antecedent of B+, for example ["Rationality requires one to believe that p, if one believes that there is conclusive evidence that p"], if he *then* goes on to form the belief that p, *thereby* complying with B+, he does so on the basis of the evidence he believes there is ..." (Kolodny 2005, 547, my italics).

Similar points hold for action. On the plausible assumption that it is rational to do what it is rational to intend, one might hold a diachronic principle such as this *enkratic principle for action*:

> If, at t, it is rational for S to intend to A, then, at t or as soon as possible after t, it is rational for S to A.

But for the kinds of reasons we have seen, the passage of even a moment can bring some new consideration to S's mind (and, of course, A-ing may be impossible until long after t). The most one can say is that rational intention possessed at a time entails that doing the thing in question *at that time* – which may be a practical impossibility at the first moment of intention formation – is prima facie rational for the agent to the same extent as the agent's intention "directed" toward it.

4. Focal Versus Global Rationality

So far, my concern has been a kind of focal rationality: mainly that of a single action, intention, belief, or desire. But the rationality of persons is a global matter. I take it to be determined by an appropriate integration of theoretical and practical rationality and to have a complex relation to the rationality of actions and of other elements. We can be globally rational (though not *perfectly* so) even if certain of our beliefs, desires, or actions at the time are not rational. Since some of these elements are far more important than others, there is no simple way to determine when irrationality in one kind of element (say in beliefs) counts decisively against one's being globally rational. To consider theoretical rationality here would be a major task. Let us pursue just the relation between practical principles we have considered and global rationality.

I have so far granted the plausibility of a *moderate judgmental grounding principle*:

> If, at t, (1) S holds a practical judgment favoring S's A-ing, and (2) the judgment is at least minimally rational, then, at t, (3) S has prima facie reason to intend to A, and (4) S's not (at t) intending to A is prima facie (practically) irrational.[9]

---

9   The formulation does not imply that the reason is that one intends to A. The reason is likely a ground of that intention, e.g. relieving one's pain.

Suppose the judgment is *irrational*. May an irrational practical judgment, if strong and held on reflection, have such normative authority? This is debatable. Granted, other things being equal, one is better *integrated* if, given even an irrational practical judgment favoring A-ing, one intends to A, rather than lacking this intention. If integration (which I take to be largely a matter of coherence) is by itself a ground for intention, the suggested principle apparently holds. But, as the case of the retributive father apparently shows, our achieving integration of this narrow kind is neither necessary for global rationality nor sufficient for the (overall) focal rationality of action.

The matter of how to assess practical rationality in relation to theoretical rationality and to global rationality raises a further problem. The issues central in this paper concern the normative authority of judgments and beliefs that I am taking to be truth-valued or, in any case, otherwise appraisable from the point of view of theoretical reason. Insofar as cognitive matters are crucial for practical rationality, whether of intentions or actions themselves, it would seem that, in deciding what to do and, where the possibility of action is not at hand, what to plan or intend, a rational agent might well want (on balance) to give priority to beliefs – especially if they provide grounds for normative judgments and other beliefs. This implies that it will be normal and indeed common to arrive at a judgment or belief as to what we ought to do *and then* either do it or form an intention to do it. This pattern, however, often contains enough of a time gap to enable us to savor the prospect of the favored action, to recall similar cases, or to see how some principle or standard we hold bears on the matter. In even a short time, the degree of support a practical judgment provides for the action it favors may diminish, as with the retributive father portrayed earlier. We may think of a conflicting normative principle, realize we have a strong competing desire, or feel a sudden aversion to the prospect of A-ing. Such changes may result in its ceasing to be rational for us to perform (or intend) the action in question.

To be sure, much of our behavior, including much that is rational, is automatic or scripted or both. But much is not, and there we must be wary of principles of practical or theoretical reason that do not do justice to the multitude of grounds for action or belief that determine focal rationality and to the importance of temporally restricting our normative principles.

5. Rationality as Responsiveness to Grounds

The view that rationality is a matter of responsiveness to reasons is now widely held in some version (e.g. by Scanlon 1998, 17-49),[10] and, for a broad notion of reasons, it has much plausibility. If reasons are taken to be facts, as is common, the view seems too narrow (as is argued by, e.g., Parfit 2011, Chap. 1). That reasons need not be so conceived I have argued in Audi (2010); but even on the wider view proposed there, reasons are still, as on the factivity views about reasons, the kinds of things that represent contents of propositional attitudes and are expressible in *that*-clauses – though not only in those. John's reason for punishing the child, if he did, would have been (e.g.) *that* it was needed for her rectification (something he believed) or *to rectify* her conduct (something he wanted). But did he have a reason for the merciful action of (say) reprimanding her instead? And must he have had some reason for the action, if it is to be rational?

In answering these questions, I will assume that (normative) reasons are a subclass of grounds. If John's reason for not punishing his daughter was to prevent harm to her, then it was also a (normative) ground for that; and, related to this fact, his action itself was (psychologically) grounded on a desire, or perhaps intention, to prevent that harm. To be sure, if my reason for buying a ticket to Boston is to attend a conference there, it might be odd to describe the action as my buying it on the ground that it is necessary for attending a conference there. But I am not making a synonymy claim about reasons and grounds, and this point may in any case be simply pragmatic: my view is that a normative reason that explains action in some sense grounds it, not that 'reason for' is synonymous with 'ground for'.

For cases of belief in which grounds for belief are not properly considered reasons for it, think of perceptual belief. Suppose I see smoke billowing from a distant hillside. In such a case I will immediately believe something is burning there. My ground, both normative and psychological, for the belief is *my seeing smoke*. Suppose, however, I am talking by telephone to someone far away who has conflicting information and asks my reason for believing something is burning on that hillside. I will likely say *that I see*

---

[10]   The factivity view is not new: "Reasons are what we mean to reason from, and reasons are facts ... sailors who, believing that the earth is flat, declined to sail with Columbus had *in* that belief no reason to decline: since the earth is not flat, its being flat was no reason" (Stampe 1987, 337).

smoke billowing up from it. This that-clause does express *a* reason for my belief, but this smoke-belief is itself grounded in my seeing the smoke; and the belief is not the ground – at least not the *original* ground – of my believing that there is fire. We may also say that my visual perception of smoke is the reason *why* I believe that. But not every reason *why* someone holds a belief is a normative ground or indeed a ground at all, as brain manipulation illustrates. It can cause a belief without being even a psychological ground of it, much less a normative one; and seeing smoke – the perception itself – would not normally be called a reason for my belief.

A supporting point here is that normative reasons are always and naturally expressible in that-clauses, which have truth-valued contents, whereas grounds, such as sensory experiences and memory impressions, need not be so expressible but can still justify truth-valued attitudes. They can do this, moreover, by conferral rather than transmission. They thus do not invite the regress or circle encountered if one supposes that propositional attitudes are justified only by elements of the same truth-valued kind.

Recall John, the retributive father. May we say that John has one or more good grounds for withholding punishment of his daughter? That is the view I have defended by citing his normatively adequate grounds for this. Some of those grounds may be expressed in his beliefs and may be reasons for one or another action toward her, but some of his grounds might not be naturally described as reasons. Consider his sense of her fear and vulnerability as he looks into her eyes. This sense may arouse a strong desire not to hurt her or even a normatively relevant emotion, such as an empathic fear of hurting her. These phenomenal elements in turn might in some direct way lead him to pause, but they might also lead to a belief (not necessarily formulated or otherwise manifested in consciousness) that punishment would hurt her. That belief, together with the desire not to hurt her, could lead to pausing and, together with other cognitions or supporting emotions or both, might also lead to finding a reprimand best in the situation and to making a practical judgment favoring that alternative action.

Here, as in many cases, the executive and well-groundedness models might agree. We are psychologically so constructed that, commonly, our practical judgments do appropriately reflect our overall system of grounds. Their often doing so is certainly implied by a high degree of global rationality in the person in question. But even when we take time to reflect, we can make a practical judgment which conflicts with our most important grounds, beliefs, and desires relevant to the action. If these elements, in an appropriate

way, then determine action or intention that is against our better judgment, we may thereby exhibit weakness of will without irrationality.

Again, we find that at best the *enkratic* requirement is only a prima facie one and applies only synchronically, at a given time. One could put the point this way, echoing the moderate *enkratic* principle: at t, there is prima facie reason not to be in a *state* in which one believes or judges that on balance one ought to A, yet does not intend to A. Compare the counterpart for belief, a *moderate evidential grounding principle*:

> At t, there is prima facie reason not to believe that there is conclusive evidence for p, yet fail to believe p at t.

In different cases – partly depending on how rational the relevant intention or belief is – the prima facie reason may be stronger or weaker. Neither principle is on a par with a *coherence requirement* – or, better, an *incoherence requirement* – that arguably expresses a necessary condition for overall theoretical rationality, at least at its highest levels:

> At any given time, t, one ought not to believe p and that p entails q, but also that not-q.

Even that requirement, however, is synchronic. Because of what can happen with the passage of time, we should not accept the superficially similar *diachronic modus-ponential principle* that

> If, at t, one believes that p and that p entails q, then, at the "next moment," one should infer that q (in the assenting way that entails forming the belief that q).

Nor should we accept – as some have apparently done in considering Kant's hypothetical imperative – the *instrumentalist volition principle* that

> If, at t, one *wills* (say in deciding or forming an intention) to bring about an end, one should then (at the "next" moment or as quickly as possible) will any action one takes to be within one's power and an indispensably necessary means to this end.[11]

---

[11]   The hypothetical imperative is informatively discussed by Korsgaard (1997), though sometimes in versions that lead one to wonder whether she might have in mind a version of the instrumentalist volitional principle.

The *enkratic* requirement and similar ones have a plausibility we must account for; but in the versions we should accept, they must be understood as synchronic, and they should be applied to appraising rationality only in the wide context of a theory of an agent's overall grounds for action.

It should be added here that the well-groundedness view does not *depend* on distinguishing, as I do, grounds from reasons. If reasons encompass what I call grounds, we can speak instead of an adequate reasons view. But distinguishing them enables us to do better justice to the overall appraisal of rational action and, especially, rational belief, which may well be, in a certain way, a more basic normative notion than rational action. Our beliefs respond to the world more directly than our (intentional) actions. Beliefs may be rational when they rest on an experiential ground that is at best misleadingly called a reason. Action is essentially belief-guided in a way that belief is not action-guided – nor desire-based or even goal-based;[12] and, accordingly, rational action is (at least roughly) action for an adequate reason. This point is among the reasons for the plausibility of the *enkratic* requirement in the first place, since, commonly, judgments that one ought on balance to A are based on one or more grounds that support A-ing. If, however, we take normative grounds simply to be kinds of normative reasons, the points I have made about the holistic character of rationality are unaffected and the proposed partial account of rational action can be redescribed.

*   *   *

A prominent though fallible indication of practical rationality is acting on one's normative beliefs, especially when one expresses them in a judgment as to what one should do here and now. But the normative force of practical judgment is not intrinsic to it; that force depends on grounding for such judgments in rationality-conferring elements in the agent. These include certain of the agent's beliefs, desires, and experiences at or near the time – whether reflective, perceptual, or emotional – as well as the precipitate of past experiences accessible to the agent through memory and, potentially, an influence on conduct. Rationality admits of degrees, and in

---

[12]   The metaphorical idea that belief aims at truth concerns how beliefs are to be evaluated and should not be allowed to cause assimilation of beliefs to actions or even intentions.

highly rational agents there is a significant degree of coherence between overall grounds for action and self-addressed moral judgments, and indeed other cognitions expressing beliefs about what one ought to do. But judgments and beliefs may influence us even if they are, overall, unjustified or even irrational. A rational agent may often give up ill-grounded cognitions; but it is also possible that, under the pressure of time which life so often exerts, the influence of one's overall grounds for action may, as with the retributive father, fail to reverse an ill-grounded judgment one holds, yet still normatively outweigh such a judgment present in consciousness at the moment. These grounds may then lead to a rational action against one's better judgment. This action – and even the intention to perform it – will exhibit weakness of will at the time, though not necessarily a constitutionally weak will as an element in the agent's character. The global rationality of agents may override a directive of their will at certain times, and rational action, like rational belief, cannot be adequately appraised simply on the basis of any single judgment, or even any single piece of reasoning, that favors the action. Rationality is a kind of well-groundedness, and there are many kinds of normatively significant grounds and kinds of paths leading from these to rational action.[13]

References

AUDI, R. (1990): Weakness of Will and Rational Action. *Australasian Journal of Philosophy* 68, No. 3, 270-281.

AUDI, R. (2010): Reasons for Action. In: Skorupski, J. (ed.): *The Routledge Companion to Ethics.* London – New York: Routledge, 275-285.

BRANDT, R. B. (1979): *A Theory of the Good and the Right.* Oxford: Oxford University Press.

BROOME, J. (1999): Normative Requirements. *Ratio* 12, No. 4, 398-419.

BROOME, J. (2010): Rationality. In: O'Connor, T. – Sandis, C. (eds.): *A Companion to the Philosophy of Action.* Oxford: Wiley Blackwell, 285-291.

Hume, D. (1739-40/1978): *A Treatise of Human Nature*. Niddich, P. (ed.), 2nd ed. Oxford: Oxford University Press.

Kolodny, N. (2005): Why be Rational? *Mind* 114, No. 455, 509-563.

Korsgaard, C. M. (1997): The Normativity of Instrumental Reason. In: Cullity, G. – Gaut, B. (eds.): *Ethics and Practical Reason*. Oxford: Oxford University Press.

Parfit, D. (2011): *On What Matters*. Vol. I. Oxford: Oxford University Press.

Scanlon, T. M. (1998): *What We Owe to Each Other*. Cambridge (Mass.): Harvard University Press.

Stampe, D. (1987): The Authority of Desire. *Philosophical Review* 96, 355-381.

Rational *Akrasia*

JOHN BRUNERO

Department of Philosophy. University of Missouri – St. Louis. 599 Lucas Hall
1 University Blvd. St. Louis. Missouri 63121. USA
bruneroj@umsl.edu

ABSTRACT: It is commonly thought that one is irrationally akratic when one believes one ought to *F* but does not intend to *F*. However, some philosophers, following Robert Audi, have argued that it is sometimes rational to have this combination of attitudes. I here consider the question of whether rational *akrasia* is possible. I argue that those arguments for the possibility of rational *akrasia* advanced by Audi and others do not succeed. Specifically, I argue that cases in which an akratic agent acts as he has most reason to act, and cases in which an akratic agent achieves a kind of global coherence he wouldn't have achieved had he instead formed intentions in line with his best judgment, do not establish the possibility of rational *akrasia*. However, I do think that rational *akrasia* is possible, and I present two arguments for this thesis. The first argument involves a case in which one is incapable of revising one's belief about what one ought to do, where one also acknowledges this belief to be insufficiently supported by the evidence. The second argument involves a case in which one rationally believes that one ought to have an akratic combination of attitudes.

KEYWORDS: *Akrasia* – Audi – Broome – *enkrateia* – Enkratic Requirement – rationality – rational requirements.

It's commonly thought that it's irrational not to form intentions in accord with one's best judgment. More precisely, it's commonly thought that if one believes one ought to *F*, but does not intend to *F*, then one violates some requirement of rationality.[1] For instance, I violate some requirement

---

[1] Perhaps it's best to avoid speaking of one's "best judgment," and to instead speak of one's beliefs about what one ought to do, when formulating a requirement governing

of rationality when I believe I ought to take the day off, but don't intend to take the day off. I violate a requirement of rationality because my attitudes – specifically, the combination of my believing I ought to take the day off and my not intending to do so – fail to cohere with one another. And this is so regardless of whether my belief is true. Even if it's not the case that I ought take the day off, and I really ought to go to the office instead, I still violate some rational requirement in having this combination of attitudes.

What requirement is it that one violates? John Broome (2010, 290) suggests it's the following:

> *Enkrasia*: Rationality requires of *N* that if
> (1) *N* believes at *t* that she herself ought to *F*, and
> (2) *N* believes at *t* that, if she herself were then to intend to *F*, because of that, she would *F*, and
> (3) *N* believes at *t* that, if she herself were not then to intend to *F*, because of that, she would not *F*, then
> (4) then (*sic.*) *N* intends at *t* to *F*.

The second and third clauses are meant to ensure that one believes that one's *F*-ing depends on one's intending to *F*. After all, there may be nothing irrational about believing one ought to *F* and not intending to *F* when one also believes that one will *F* regardless of what one intends, or when one believes one's intention to *F* would be ineffective in bringing it about that one *F*s.

Broome's requirement is a "wide-scope" requirement in that the logical scope of "requires" ranges over the entire conditional. Since it does so, this

---

akrasia and *enkrateia*. On one reading of "best judgment," one's best judgment is "one's judgment of what is best." But such beliefs about what is best don't seem relevant to the Enkratic requirement, since one could believe some option A is the best option without being irrational in not intending to A. For instance, a satisficing consequentialist might believe he ought to either A or B, and that A-ing is the best option, but yet not be irrational in intending to B instead. Thanks to Nora Heinzelmann for this point. On another reading of "best judgment," one's best judgment is one's judgment made in the best (or at least very good) conditions – with proper reflection, freedom from bias or self-interest, etc. But that can't what's be relevant to the Enkratic requirement either since even our non-ideally-made judgments about what we ought to do are such that we could be irrationally akratic with respect to them. I suspect that, at least with regard to discussion of the Enkratic requirement, what people have in mind by one's "best judgment" is simply one's beliefs about what one ought to do.

requirement doesn't require of the akratic agent that she make some specific change in her attitudes. Rather, what is required of her is that she *either* intend to *F*, *or* not believe that she herself ought to *F*, *or* revise one of her beliefs (in (2) or (3)) about the significance of her intention to *F*. Proceeding in any of these ways would allow her to escape from the state of having an incoherent combination of attitudes.

Although it is commonly thought that *akrasia* is always irrational, some philosophers have recently challenged this.[2] These philosophers argue that there are, or at least could be, cases of "rational *akrasia*" – this is, cases in which one fails to have intentions that cohere with one's beliefs about what one ought to do, but in which one doesn't fail to do as rationality requires.[3]

In this paper, I'll consider whether there are any genuine cases of rational *akrasia*. Specifically, I'll consider whether any of the examples discussed in the literature show that *Enkrasia* is false – that one could believe one ought to *F*, not intend to *F*, have the relevant beliefs about the relationship between one's intending to *F* and one's *F*-ing, and yet not fail to do as rationality requires. I'll argue (§1) that none of the examples or arguments discussed in the literature show that rational *akrasia*, in this sense, is possible. However, I'll also introduce two new cases (§2, §3) that I think *are* genuine instances of rational *akrasia,* in this sense. In these two cases, it is not irrational for one to have the combination of attitudes that *Enkrasia*, or a related requirement, prohibits one from having.

1.

I'll start by focusing on Robert Audi's (1990) argument for the possibility of rational *akrasia*, primarily because he was the first to present a sus-

---

[2]   Audi (1990), MacIntyre (1990), Arpaly (2003), Jones (2003), Tappolet (2003). Coates (2012) challenges the view that epistemic *akrasia* – believing contrary to one's best judgment – is always irrational.

[3]   Although those who defend the possibility of rational *akrasia* tend to speak of the rationality of *acting* contrary to one's best judgment, the arguments they give would also support the rationality of *intending to act* contrary to one's best judgment. And so these arguments would have a clear bearing on *Enkrasia*, which concerns one's intending to *F*, rather than one's *F*-ing. Robert Audi (1990, 272-273), for instance, acknowledges that what he claims regarding the rationality of weak-willed action also extends to the rationality of weak-willed intention.

tained argument for this thesis, but also because other, more recent, attempts to argue for this thesis can be viewed as attempts to follow up on different strains of Audi's argument. Audi's argument focuses on the following example: John believes that he ought not allow his daughter to go out on Saturday night because he thinks that prohibiting her from going out would be proportionate retribution for her talking on the phone when she should have been studying. But when John, a committed retributivist, is about to tell his daughter about her punishment, he "looks into her eyes, realizes that she will be quite upset, decides to make do with a stern rebuke, and let her go" (Audi 1990, 276). John, we are assuming, doesn't change his mind about what he ought to do; we could even suppose that, holding onto his belief, he criticizes himself after the fact for being weak-willed when it came time to administer the punishment. So, we have here a clear case of *akrasia*. But is John's *akrasia* irrational?

Audi argues that our answer to this question will depend on how we fill out the details of John's case. Suppose there are other elements of John's psychology that would cohere well with his not punishing his daughter, such as a belief that doing so would be hurtful and a desire not to be hurtful, and a belief that doing so would likely provoke his daughter to rebel and a desire not to provoke rebellion. These desires and beliefs constitute "good reasons rooted deep in his character" (Audi 1990, 277). And John's belief that he ought to punish his daughter doesn't itself cohere with John's character, interests, and ideals; it is, Audi assumes, an irrational belief for him to hold. Given these features of John's psychology, John's acting *against* his best judgment would better cohere with his overall psychology than would his acting in accordance with his best judgment. This example is presented in support of Audi's main thesis: "that rationality must be holistically conceived and that when it is, some incontinent actions may be seen to be rational" (Audi 1990, 280).

Audi's remark that John's failure to follow through in punishing his daughter is grounded in "good reasons," and his related remark that John's behavior "accords with a civilized and generally admirable compassionate desire," are not relevant to the question we are asking here (cf. Audi 1990, 277). Rationality has to do with coherence among one's attitudes. But whether there is good reason for someone to hold some particular attitude isn't relevant to how well or badly one's attitudes cohere. Consider again the example with which we started. I falsely believe that I ought to take the day off, and, akratically, intend to go to the office instead, where this is

precisely what I ought to do. The fact that there are good reasons to intend to go to the office doesn't render me any more coherent in holding the attitudes I hold, and so won't help establish the possibility of rational *akrasia*.[4]

For similar reasons, we can also set aside those attempts to argue for the possibility of rational *akrasia* by showing how *akrasia* could lead one to act as one has an *internal* reason to act (cf. McIntyre 1990). Internal reasons, very roughly, are reasons that an agent has to perform some action where an agent could come to be motivated to perform that action by a process of practical reasoning starting from his "subjective motivational set" – that is, his set of desires, values, commitments, etc. (Williams 1981). On Williams's conception of internal reasons, there could be an internal reason of which one is unaware, perhaps as a result of a false belief one holds. Suppose I falsely believe that going to the office will be unproductive when in fact it will be quite productive. In this case, there is an internal reason to go to the office of which I'm unaware. When, against my better judgment, I go to the office, I act as I in fact have good (internal) reason to act. But that doesn't make my intending to go to the office cohere any better with my other attitudes. I'm still criticizable as irrational, even though I act as I have an internal reason to act.

---

[4]   A parallel point holds for attempts to establish the possibility of rational *epistemic akrasia* by pointing to cases in which there are good reasons to have a belief even though one believes that one ought not hold this belief. See, for instance, Coates (2012). If we understand rationality as we've been understanding it in this paper, as a matter of coherence among one's attitudes, then whether there's good reason – that is, sufficient evidence – to have some belief isn't relevant to the question of whether one is *rational* in holding it. One would still be incoherent insofar as one has the combination of believing one ought not believe that P and believing that P.

However, in the related case in which one *believes* that there's sufficient evidence for the belief that P, then one does have some attitude which has a potential bearing on one's rationality. But this attitude won't help establish the possibility of rational epistemic *akrasia*. Rather, it seems to introduce a *second* violation of a rational requirement. In addition to the epistemic *akrasia* (believing one ought not believe that P and believing that P), one would also have inconsistent beliefs: one would believe both that one ought not to believe that P and that there's sufficient evidence that P. If we suppose that one also believes that if there's sufficient evidence that P, then it's not the case that one ought not to believe that P, then one has inconsistent beliefs.

So, we'll set aside Audi's observations of how akratic intentions could lead one to act as one has reason to act.[5] Indeed, it's clear that Audi doesn't intend to rest his argument on such observations. His argument, rather, rests upon a different observation: that an akratic action, although involving a local incoherence between the action and one's best judgment, may cohere with one's psychology *taken as a whole*. Consider the passage immediately preceding Audi's introduction of the example of John:

> ...even in acting reflectively we may not take adequate account of our overall perspective, our perspective as determined by certain of our basic beliefs and desires, especially those crucial in our world view. Could there be, then, an action against one's better judgment which, through its accord with (the relevant parts of) our overall perspective, *is* rational? If so, we should reject the common assumption that incontinent actions are all irrational. (Audi 1990, 276)

Here, what matters is coherence: a case of *akrasia*, while involving local incoherence, could achieve global coherence – that is, accord with one's "overall perspective" – and is therefore rational.

However, it's not clear at all how the conclusion that "we should reject the common assumption that incontinent actions are all irrational" would follow. I agree that Audi's example shows that someone's acting akratically can be *more rational* than someone's acting in accordance with one's best judgment. But it doesn't follow from this that not all akratic actions are irrational. After all, the claim that one way of proceeding is more rational than another is a comparative assessment. But it could very well be that both ways of proceeding are irrational, with one of those being more rational than the other. (In the same way, one option could be better than another even though both are bad options.)

---

[5]   A similar line of argument is presented in Tappolet (2003, 115), who observes that one's emotions "can, it seems, make us more rational, in the sense of allowing us to track reasons which we have but which we've neglected in our deliberation" even when those emotions run contrary to one' beliefs about what one ought to do. Likewise, Jones (2003, 181) starts from the observation that "emotions sometimes key us to the presence of real and important reason-giving considerations without necessarily presenting that information to us in a way susceptible of conscious articulation and, sometimes, even despite our consciously held and internally justified judgment that the situation contains no such reasons". These are important observations, but they do not establish the possibility of rational *akrasia* in the sense in which we're interested here.

Is there any justification for claiming that John's *akrasia* is rational? Audi does concede that akratic action is irrational to some degree, but he maintains that "without being rational to the highest degree, an action may still be rational on balance" (Audi 1990, 280 and 275, respectively). Presumably, the idea is that there is more to be said in favor of the rationality of John's *akrasia* than against it. Although the local incoherence speaks against John's *akrasia*, the more significant achievements in global coherence speak in favor of it, and so, on balance, John's *akrasia* is rational.

But this line of reasoning seems to simply overlook that there was *another alternative* open to John – namely, his ceasing to believe he ought to punish and not intending to punish – and that this alternative has all of the advantages of the way he actually proceeded and none of the costs: it achieves global coherence without incurring any local incoherence. It seems inappropriate to assess the rationality of John's *akrasia* without considering this alternative available to him.

Consider an analogy: suppose I've parked my car on the train tracks and there are two ways I could escape this dangerous situation: I could put the car in drive and go forward without incident or put it in reverse and go backwards, running over your dog. I need to remove my car from the tracks very soon to avoid the oncoming train. Now, it would be a very poor defense of my driving in reverse were I to say to you, "Well, there are some considerations speaking against going in reverse – the injury to your dog – but other, more significant, considerations counting in favor of it – namely, that it allowed me to avoid getting hit by the train – and so my actions were, on balance, reasonable." That's a poor defense because there is another, equally good, way to achieve those same benefits without incurring any costs: driving forward. Likewise, there are two ways that John could achieve the benefits of global coherence: by being akratic or by revising his belief. And, for similar reasons, it's a poor defense of his *akrasia* to argue that although it involves some degree of irrationality, its gains with respect to global coherence render it, on balance, rational. Both defenses involve assessing an option without considering relevant, superior alternatives.

I'm not disputing the plausible claim that there are two ways John could proceed – his continuing to believe he ought to punish and not intending to punish, and his ceasing to believe he ought to punish and not intending to punish – that are each *more rational* than his following through and intending to punish. Rather, I'm simply denying this establishes the possibility of rational *akrasia*. And so I'm not expressing any dis-

agreement with Nomy Arpaly's claims in the second chapter of *Unprincipled Virtue*. Arpaly presents an example designed to show that one's acting against (as opposed to in accord with) one's best judgment could increase overall coherence in one's set of beliefs and desires, and she concludes from this example that "there are some cases in which one is *more rational* in acting against one's best judgment than one would be in acting according to it" (Arpaly 2003, 41; emphasis added). She is careful to avoid claiming that there are some cases in which *akrasia* is rational.

Let's sum up the argument so far. I have been considering the question of whether rational *akrasia* is possible – that is, whether one could believe one ought to F, not intend to F, and have the relevant beliefs about the relationship between one's intending to F and one's F-ing, and yet be rational. I have argued that none of the arguments in the current literature establish that it is. Some of those arguments appeal to the ways in which akratic action may involve one's acting for good reasons. But these considerations can be dismissed as irrelevant since we are understanding rationality as a matter of coherence among one's attitudes. Other arguments appeal to the global coherence that could be gained by one's acting against one's best judgment. But such arguments can only establish the comparative conclusion that one's intending akratically may be *more rational* than one's intending in line with one's best judgment, not the conclusion that rational *akrasia* is possible.

2.

I'll now argue that rational *akrasia* is possible. I'll argue that there are some instances in which it would be rational for one to believe one ought to F, not intend to F, and have the relevant beliefs about the relationship between one's intending to F and one's F-ing.

Let's start by considering a variation on Audi's example. Jack, like John, believes that he ought to punish his daughter, and this belief doesn't cohere well with many other beliefs and desires he has. And, we'll assume, Jack believes that there isn't sufficient evidence for his belief that he ought to punish her.[6] But Jack knows that he's unable to change this belief. He

---

[6] It's not clear whether Audi's John has this belief. Audi claims that the belief is irrational, but doesn't specify whether John believes that his belief that he ought to punish isn't supported by sufficient evidence.

knows that even if he were to review all the convincing arguments against punishing her, and remind himself that there's not sufficient evidence for his belief, he would continue to believe that he ought to punish his daughter.[7] Perhaps after extensive therapy, he'll cease to believe this. But he also knows that by then it would be too late, since the decision about whether to punish her needs to be made now.

This example differs from Audi's example of John. In Audi's example, we resisted the conclusion that John's *akrasia* was rational since there was a better option available to John – namely, his revising his belief – that yielded all the gains in global coherence without any local incoherence. But Jack knows this isn't an option for him. (Jack's situation would be analogous to the case in which one has driven on the tracks but, after a mechanical failure prevents one from putting the car into drive, one realizes that backing up is the only way to avoid the train.) If, in light of this knowledge, Jack decides not to punish his daughter, Jack would have the combination of attitudes prohibited by *Enkrasia*, but yet Jack's having this combination wouldn't be irrational.

Note that I'm not claiming that Jack is fully rational. After all, a fully rational person would not hold beliefs he believes to be insufficiently supported by the evidence. Such a combination of attitudes seems to be a violation of an epistemic analogue of *Enkrasia* – a prohibition on believing P when one believes there's insufficient evidence that P. But it doesn't follow from Jack's not being fully rational that Jack's *akrasia* is irrational.[8] For one thing, we can locate the specific source of Jack's irrationality and it's not his *akrasia*, but his *theoretical* irrationality. For another, Jack's not intending to punish his daughter seems to be a rational way of *responding to* this irrational, recalcitrant belief.

Even though a fully rational version of Jack wouldn't violate *Enkrasia*, it seems wrong to accuse *Jack* of irrationality in violating *Enkrasia*. We are familiar with other examples like this in the philosophical literature, where

---

[7]   For discussion of a similar case involving an unalterable belief, in the context of a debate about whether rational requirements are normative, see Setiya (2007, 656).

[8]   It might be inappropriate to call Jack's case one of *akrasia*, since it's not clear that Jack is displaying any *weakness* in his will. But that doesn't matter for my purposes here. All that I wish to argue for is that it's a case in which one believes one ought to *F*, does not intend to *F*, and one has the relevant beliefs about the relationship between intending to *F* and *F*-ing, but yet it seems wrong to say one violates a requirement of rationality in having this combination of attitudes.

what one's fully rational self would do differs from what it would be rational for one's actual self to do, given the ways in which one's actual self falls short of full rationality. Consider an example discussed in another context by Michael Smith (1995) who himself borrows the example from Gary Watson (1975) (see also Railton 1986). A squash player who is aware of his violent uncontrollable anger suffers a crushing defeat to his opponent. Rather than approach him for the customary post-game handshake, he decides to forgo the handshake and leave the courts immediately, for fear that his anger might get the best of him and he'd do something he'd regret. Although a fully rational version of the squash player would have no problem being a good sport and shaking his opponent's hand, this is not the rational course of action for the squash player himself. Rather, given his irrational anger, which he can do nothing about, the rational course of action is to leave the courts immediately.

It's important for an account of rationality to consider such cases. After all, we employ rational requirements when we advise and criticize others. And, an account of rationality that didn't consider the ways in which we fail to be fully rational would end up giving bad advice, and inappropriate criticism. For instance, an account which criticizes our angry squash player for simply walking away issues inappropriate criticism, and an account which advises him to walk up to his opponent issues pernicious advice.

The example of Jack is similar. Although a fully rational version of Jack wouldn't be akratic (since he would have given up his normative belief), Jack's violation of *Enkrasia* is a rational response to an irrational belief that he knows he can't change. An account of rationality that criticized his *akrasia,* not simply his theoretical irrationality, would be issuing inappropriate criticism. And an account that advised him to follow through on his belief that he ought to punish – a belief he recognizes to be irrational, but can't change – would be issuing bad advice.

The example of Jack, I've argued, establishes the possibility of rational *akrasia*. But one might wonder whether there are any actual cases like this. Although I'm arguing only for the thesis that rational *akrasia* is possible, I do think there are some actual cases that are at least plausible candidates for being cases of rational *akrasia* along these lines. People with Obsessive-Compulsive Disorder (OCD) often have unwanted thoughts and behaviors that they can't control, including, sometimes, certain normative beliefs. For instance, one might believe that one really ought to check once more whether the stove is turned off. But, people who have OCD have, at some

point, recognized these obsessions or compulsions as "excessive or unreasonable" (see American Psychiatric Association 2000, Ch. 7, §300.3, B). A person with OCD who recognizes that he can't help having such unreasonable normative beliefs may be able to exercise a greater degree of control over his actions, and actively resist acting upon what he believes he ought to do. And such resistance may lead him to have the combination of attitudes prohibited by *Enkrasia*; it may, for instance, lead someone to believe he ought to check the stove one more time and not intend to check the stove one more time. Such cases, since they involve rational responses to one's irrationality, are, I think, cases of rational *akrasia*, in the sense that one has a combination of attitudes prohibited by *Enkrasia*, yet is not irrational in having that combination. However, since there are questions about how exactly to understand the psychology of those who suffer from anxiety disorders such as OCD, and since a full treatment of the relevant issues in philosophical psychology would take us well beyond the scope of this paper, I'll limit myself to claiming that such cases are at least plausible candidates for being actual cases of rational *akrasia*.

One might challenge the thought that there is a significant difference between the cases of Jack and John, provided that we understand John's case in a certain way. Jack believes he has insufficient evidence for his normative belief (that is, his belief that he ought to punish his daughter), but he knows that he's unable to change it. But John might be such that, although he knows it's possible for him to revise his normative belief, such a revision would be *rationally inaccessible* to him, in the sense that he doesn't see a rational basis on which to revise this normative belief. Even though his normative belief fails to cohere with many of his other beliefs and desires, he may not be in a position to appreciate this incoherence and revise his normative belief in light of these other beliefs and desires. We might suppose that John believes (falsely) that his belief that he ought to punish his daughter coheres well with the rest of his beliefs and desires, thereby making the revision of his belief that he ought to punish his daughter rationally inaccessible.[9]

This does bring the example of John closer to that of Jack in some important ways, since in saying that the revision of one's normative belief is rationally inaccessible, we come closer to saying that it's impossible. How-

---

[9]   Thanks to Robert Audi for remarks which suggested to me this way of understanding the example.

ever, much more would have to be said about the notion of rational inaccessibility than can be explored here. But, putting this aside, I do think we would do better to work with the example of Jack instead. For one thing, if we specify that John believes that his belief that he ought to punish his daughter coheres well with the rest of his beliefs and desires, it's less clear that there would be significant gains in global coherence in his being akratic. In being akratic, he would be acting contrary to a normative belief that, *by John's own lights*, coheres well with the rest of his beliefs and desires. So, it's less clear that he acts in accord with his "overall perspective" in being akratic. For another thing, it's open for one to insist that there's still a better option open to John: revising *both* his belief that he ought to punish her *and* his belief that this belief coheres well with his other beliefs and desires. (John is still *capable* of coming to see how his belief that he ought to punish her doesn't in fact cohere with his other beliefs and desires, and revising his beliefs in light of this discovery, thereby removing any incoherence.)

But, most importantly, in specifying the example of John along these lines, we haven't yet said anything to establish that the local incoherence (his believing he ought to punish but not intending to punish) isn't irrational. But the example of Jack *does* include a feature which licenses us to claim that this local incoherence isn't irrational: the local incoherence is the result of Jack's *rationally responding to* a normative belief he knows he cannot change, and hence isn't irrational.

In summary, Jack has the combination of attitudes prohibited by *Enkrasia* (believing one ought to *F*, not intending to *F*, and having the relevant beliefs about the relationship between one's intending to *F* and ones *F*-ing), but yet his having this combination of attitudes is not irrational.

3.

I'll now turn to another way in which *akrasia* can be rational. Suppose that one believes that one ought to have a combination of attitudes prohibited by *Enkrasia* – that is, one believes that one ought to believe one ought to *F* and not intend to *F* (while having the relevant beliefs about the relationship between one's intending to *F* and one's *F*-ing). Admittedly, this is a peculiar case: it's a case where violating *Enkrasia* would be a way of coming into conformity with one's belief about which attitudes one ought

to have. In this section, I'll argue that in such cases, it may be rational to go against one's better judgment.

Let's consider the applicable requirements in this case by considering an example. Suppose Jill has the following combination of attitudes: she believes she ought to register for the conference, she doesn't intend to do so, but she believes that she ought to have the following combination of attitudes: believing she ought to register and not intending to do so. Rationality has to do with conflicts among our attitudes, and it seems that there are two relevant conflicts, or potential conflicts, that rational requirements might here govern:

(i.)   the conflict between her believing she ought to register and her not intending to register, and

(ii.) the conflict between her believing she ought to have the combination *believing she ought to register and not intending to register* and her not having the combination *believing she ought to register and not intending to register*.

If Jill resolves her *akrasia* with respect to her belief that she ought to register, she'll then be flouting her best judgment: she won't have the combination of attitudes she believes that she ought to have.

Which requirements govern these conflicts? The following two wide-scope requirements seem applicable:

(R1) Rationality requires that (if Jill believes she ought to register, then she intends to register).

(R2) Rationality requires that (if Jill believes she ought to have the combination *believing she ought to register and not intending to register*, then she have the combination *believing she ought to register and not intending to register*).

The second of these requirements might seem a bit unusual since it concerns an agent's belief about a *combination* of attitudes, as opposed to some specific attitude. But we want to be able to criticize *akrasia* with respect to beliefs about combinations of attitudes as well. For instance, I believe I ought not both intend to drink tonight and intend to drive tonight, though I don't think there's anything wrong with having one of these intentions without the other. Were I to go on to have this combination of attitudes, without revising my belief, I would be irrationally akratic. Or I

might have the belief that I ought to have the combination of intending to fill out the application and intending to pay the application fee, while believing there's no point in doing one without the other. We want to be able to criticize one's *akrasia* with respect to this belief as irrational as well.[10]

The second of these requirements might seem unusual for another reason: the agent doesn't believe that her having this combination of attitudes is under her control in the same way in which her actions are under her control.[11] Recall that Broome's *Enkrasia* applies only when the agent has beliefs – specifically, the beliefs (2) and (3) – about how her F-ing depends upon her intending to F. But it's not plausible to assume that Jill, or any rational person, would have similar beliefs regarding this combination of attitudes; she wouldn't believe that her having this combination of attitudes depends on her intending to have them. She knows she can't have this combination of attitudes "at will".

However, I don't think that such beliefs are necessary components of rational requirements in general. Consider, for instance, that it's irrational for one to believe one has conclusive evidence that P, but yet not believe P. Here, one's attitudes fail to cohere. One would be in violation of the rational requirement that Niko Kolodny (2005, 521) has formulated as:

(B+) Rationality requires one to believe that P, if one believes there is conclusive evidence that P.[12]

But, for this requirement to apply, it's not necessary that one think one's believing that P depends upon one's intentions. And that's a good thing, since it's doubtful that we hold our beliefs "at will" – that is, it's doubtful that our having them depends on our intending to have them in the same

---

[10] In these examples, one believes that one ought to have, or not have, a combination of attitudes only because one believes that one ought to perform, or not perform, the relevant combination of actions. For instance, one believes that one ought not have the combination of intending to drink and intending to drive only because one believes one ought not drink and drive. But, as we'll see below, it's possible for one to believe one ought to have a certain combination of attitudes for state-given reasons as well.

[11] I was helped here, and in the next paragraph, by objections from Jonathan Way.

[12] Kolodny argues for a narrow-scope interpretation of this requirement and the related "core" requirements below. Others, including myself, have argued, against Kolodny, that they should be understood as wide-scoped. See Brunero (2010) and Broome (2007). I'll here assume the wide-scope interpretation is correct.

way that our *F*-ing (usually) depends on our intending to *F*. So, I'm not worried by the fact that R2 doesn't require the agent to believe she can have the combination of attitudes "at will".

The requirement R2, however, is not an application of Kolodny's B+, since R2 concerns Jill's believing that she *ought* to have a certain combination of attitudes. It doesn't concern her believing that she has *conclusive evidence* for anything. It's possible that she believes she ought to have this combination of attitudes just because of the pragmatic benefits that come from having them, not for any evidential reasons. However, it's worth noting that Kolodny takes B+ to be a more specific application of the "core requirements" of rationality, which he formulates in (2005, 524) as follows:

(C+)   Rationality requires one to have *A,* if one believes that one has conclusive reason to have *A*; and

(C-)   Rationality requires one not to have *A* if one believes that one lacks sufficient reason to have *A*.[13]

And these more general requirements aren't restricted to beliefs about conclusive evidence for beliefs. Also, such requirements are in place regardless of *why* it is that one comes to believe one has conclusive reason to have *A* – whether one has this belief for the "right kind" or the "wrong kind" of reasons. For instance, if Pascal believes that he has conclusive reason to believe in God only because of the pragmatic benefits of believing in God, but doesn't believe in God, his attitudes fail to cohere: he fails to believe in line with his beliefs about what he has conclusive reason to believe. In the same way, even if Jill believes she ought to have this combination of attitudes for

---

[13]   I'll avoid dealing with the complicated question of the relationship between the requirements of rationality and the principles of correct reasoning. This would take us too far afield. Andrew Reisner has pointed out to me that *Enkrasia* concerns one's belief that one ought to ɸ, as opposed to the belief that one ought to *intend to* ɸ, which might make for some important differences between *Enkrasia* and Kolodny's "core requirements" – which concern beliefs about the attitudes, not actions, that we ought to have – especially when it comes to the correct reasoning associated with these requirements. Despite these differences, there is a common feature in that both Kolodny's "core requirements" and *Enkrasia* require a coherence between one's normative beliefs (one's belief that one ought to have *A*; one's belief that one ought to *F*, respectively) and the attitudes relevant to the contents of those beliefs (*A*; an intention to *F*, respectively).

the "wrong kind" of reasons, R2 would still apply to her.[14] So long as she believes that she ought to have this combination of attitudes, she would be irrationally incoherent in not also having this combination of attitudes. (Also, keep in mind that R2 is a wide-scope requirement. Jill could comply with R2 by giving up her belief that she ought to have the combination of attitudes. She might come to see that she doesn't have good grounds for her belief that she ought to have this combination of attitudes, and then abandon that belief.)

Kolodny's "core requirements" concern single attitudes ("to have A", "not to have A"). However, for the reasons given above, I think we should understand these requirements so that they apply not only to one's beliefs about which specific attitudes one ought to have, but also to one's beliefs about which combinations of attitudes one ought to have. There doesn't seem to be a good reason for insisting that one can be irrationally akratic with respect to the former beliefs but not with respect to the latter. So, I think there is a *prima facie* case for thinking that R2 is a genuine requirement of rationality.

My argument here for the possibility of rational *akrasia* starts from the premise that the belief that one ought to have the combination *believing one ought to F and not intending to F* need not be an irrational belief. In other words, there are some contexts in which one's having this belief would be rationally permissible. Consider an example. Suppose you find yourself to be akratic: you believe you ought to register for the conference but don't intend to do so. Now suppose some eccentric billionaire arrives and offers you a significant reward if you continue to have this precise combination of attitudes, or promises a significant punishment if you resolve your *akrasia*. It seems you now have good reason to think you ought to have the akratic combination. (In case it matters, suppose also that you think it's possible for you to hold onto this combination.) In this case, it doesn't seem irrational for you to believe that you ought to have this combination. Perhaps this belief is false, but it's not irrational for you to hold it.[15]

---

[14] I'm here taking beliefs about which attitudes one "has conclusive reason" to have and beliefs about which attitudes one "ought" to have, to be equivalent. If you think there are significant differences, you could make the appropriate substitutions in the requirements.

[15] In this example, the reason one takes there to be for holding the combination of attitudes is a *state-given* reason. There is a question about whether state-given reasons are

There might be other examples. Consider John Perry's research into "structured procrastination," for which he won an Ig Nobel Prize. Perry (1996) notes that when we procrastinate, we usually don't do so by doing nothing whatsoever, but instead by doing other tasks. The goal of structured procrastination is to structure that procrastination so that one gets important things done when one is procrastinating. One needs to first convince oneself, through some self-deception, that some task is really important when it's actually not, and then, in procrastinating with regard to that task, get a lot of other more important things done. For instance, in procrastinating in registering for the conference, one might get out letters of recommendation, catch up on email correspondence, and write a paper for another conference. If I'm understanding Perry's idea correctly, the structured procrastinator might believe, at least in his more reflective moments, that the combination *believing he ought to register and not intending to do so* is a combination he ought to have; he recognizes that it "channels" his procrastination toward better results, so that the procrastinated task that doesn't get done isn't as important as the tasks accomplished while procrastinating. For someone who guides his life and work by the ideals of structured procrastination, like Prof. Perry, it might not be irrational to believe that the combination *believing he ought to complete some task and not intending to complete that task* is a combination he ought to hold. Again, it might be a false belief, but it doesn't seem to be an irrational one.

So, it's not always rationally impermissible for one to believe that one ought to have the combination *believing one ought to F and not intending to F*. In other words, the narrow-scope claim, "Rationality requires one not believe that one ought to have the combination *believing one ought to F and not intending to F*," is not true.

But here's the problem: it seems that R1 and R2 together entail this narrow-scope claim. So, if the narrow-scope claim is false, then either R1 or R2 must be false. Consider Jill. Note that the only way that Jill can comply with both R1 and R2 is by giving up her belief that she ought to

---

genuine reasons or not. But I need not consider that here. All that I'm aiming to establish here is that it is sometimes rationally permissible for one to believe that one ought to have an akratic combination of attitudes. Even if state-given reasons are not genuine reasons, someone who thinks that they are could rationally believe that he ought to have an akratic combination of attitudes. Perhaps he has a false theory of reasons, but it doesn't follow from this that his belief that he ought to have an akratic combination is irrational.

have the akratic combination of attitudes. However she resolves her *akrasia* with respect to her belief that she ought to register – whether by intending to register or by not believing she ought to – she violates R2. But if she doesn't resolve her *akrasia*, she violates R1. So, the only way of proceeding that *doesn't* involve failing to do what rationality requires is giving up her belief that she ought to have the akratic combination of attitudes (as well as resolving, in some way, her *akrasia* with respect to her belief that she ought to register). And so rationality requires that she give up her belief.

An analogy with legal requirements might help here. Suppose you are again driving your car, and you are stopped at a red light, with pedestrians walking behind your car. You are legally required not to drive in reverse, since doing so would harm the pedestrians. You are legally required not to drive forward, since the light is red. The only way you can comply with these two requirements is by staying put, and so you are legally required to stay put. Jill's situation is much the same: the only way she can comply with two rational requirements, R1 and R2, is by giving up her belief that she ought to have the akratic combination of attitudes. And so that's what she is rationally required to do.

Since "is required that *p*" logically behaves like "is obligatory that *p*" we could also utilize Standard Deontic Logic (SDL) to show how these two wide-scope claims entail the problematic narrow-scope one (see McNamara 2010, §2.1). The axioms of SDL are somewhat controversial. If you don't find SDL acceptable, then I would urge you to put weight instead on the argument in the previous paragraphs. But if you do find SDL acceptable, then you should also accept that R1 and R2 together entail the implausible narrow-scope claim that Jill is rationally required to give up her belief that she ought to have the akratic combination.

I'll use "R" for "rationality requires", "BO" for "you believe you ought to", "I" for "you intend to", and "ϕ" for an action. The only axiom of Standard Deontic Logic that we'll need here is a distribution axiom for rational requirements:

(R-K) $\mathrm{R}(p \rightarrow q) \rightarrow (\mathrm{R}p \rightarrow \mathrm{R}q)$

I'll start with R1 and R2, and show how they together entail the implausible narrow-scope claim: $\mathrm{R}{\sim}\mathrm{BO}(\mathrm{BO}\phi \ \& \ {\sim}\mathrm{I}\phi)$.

R1) $\mathrm{R}(\mathrm{BO}\phi \rightarrow \mathrm{I}\phi)$
R2) $\mathrm{R}(\mathrm{BO}(\mathrm{BO}\phi \ \& \ {\sim}\mathrm{I}\phi) \rightarrow (\mathrm{BO}\phi \ \& \ {\sim}\mathrm{I}\phi))$

3) $R(\sim(BO\phi\ \&\ \sim I\phi) \rightarrow \sim BO(BO\phi\ \&\ \sim I\phi))$ Contraposition, from 2

4) $R\sim(BO\phi\ \&\ \sim I\phi) \rightarrow R\sim BO(BO\phi\ \&\ \sim I\phi)$ R-K, from 3

5) $R(\sim BO\phi \lor I\phi)$ Cond.-Disj. Equivalence, from 1

6) $R\sim(BO\phi\ \&\ \sim I\phi)$ DeMorgan's, from 5

7) $R\sim BO(BO\phi\ \&\ \sim I\phi)$ Modus Ponens, from 4,6

And (7) is just the implausible narrow-scope claim.[16]

Since it may be the case that (7) is false – it may be, as argued above, that Jill is rationally *permitted* to hold this belief – we need to concede that, in such cases, either R1 or R2 is false. I won't take a position on *which* of these two requirements we should reject; rather, I only wish to argue that we must reject *one* of them, so as to avoid the problematic conclusion, (7). If we must reject one of these requirements, then it follows that there is at least one case in which rationality permits one to go against one's best judgment.

There may be an even easier route to establishing the possibility of rational *akrasia*.[17] The above arguments aim to show that from R1 and R2 we can derive the implausible claim that rationality *requires* Jill not to believe that she ought to have the akratic combination of attitudes. But an equally implausible claim is that Jill lacks the *property* of (full) *rationality* insofar as she has the belief that she ought to have the akratic combination of attitudes.[18] And it's very easy to show that if R1 and R2 are both true, then Jill lacks the property of rationality insofar as she has the belief that she ought to have the akratic combination of attitudes. Here's the argument. Suppose she doesn't give up her belief that she ought to have the akratic combination of attitudes. There are two possibilities: she either has the akratic combination of attitudes or she doesn't. If she does have it, she violates R1, and hence lacks the property of rationality. If she doesn't have it, she violates R2, and hence lacks the property of rationality. So, insofar as

---

[16]   Michael Titelbaum has independently presented a similar line of reasoning in "How to Derive a Narrow-Scope Requirement from Wide-Scope Requirements" (ms). Titelbaum's formal proof doesn't rely on SDL, and so may be preferable if you're not a fan of SDL.

[17]   I was here helped by some remarks from John Broome.

[18]   For the distinction between the "property" sense and the "code" sense of "rationality," see Broome (2007, §2). I'll here follow Broome in thinking that you must satisfy all the requirements of the code of rationality in order to have the property of rationality (where this is short for "full rationality").

she has the belief that she ought to have the akratic combination of attitudes, she lacks the property of rationality.[19] And that's implausible. It's implausible to claim that she must give up this belief to be rational, as I've argued above. To avoid this implausible result, we must reject either R1 or R2. In other words, we must concede that, in some cases, rationality permits one to go against one's best judgment.

4.

In summary, I've objected to previous attempts to argue for the possibility of rational *akrasia*. In showing how an akratic intention could lead one to act as one has reason to act, or how it could allow one to achieve global coherence while incurring the expense of local incoherence, one does not succeed in showing that having attitudes contrary to one's best judgment is rationally permissible. However, I've presented two arguments for the possibility of rational *akrasia*. First, I've argued one's *akrasia* could be rational when it is a rational response to some recalcitrant normative belief one believes one ought not have. Second, I've argued one's *akrasia* could be rational when one believes, not irrationally, that one ought to have the akratic combination *believing one ought to F and not intending to F*. In these two cases, it's rationally permissible to form attitudes contrary to one's best judgment. Admittedly, these are unusual cases. But one should expect such cases to be unusual given the ubiquity of the assumption that *akrasia* is necessarily irrational. However, as I've argued, that assumption is nonetheless mistaken: it is possible for *akrasia* to be rational.[20]

---

[19]   And by giving up this belief, while also avoiding the akratic combination of attitudes, she would violate no requirements of rationality, and then have the property of rationality, at least insofar as the attitudes mentioned in this example go.)

References

American Psychiatric Association (2000): *Diagnostic and Statistical Manual of Mental Disorders.* 4th ed. Text Revision, Washington, D.C.

Arpaly, N. (2003): *Unprincipled Virtue.* Oxford: Oxford University Press.

Audi, R. (1990): Weakness of Will and Rational Action. *Australasian Journal of Philosophy* 68, No. 3, 270-281.

Broome, J. (2007): Wide or Narrow Scope? *Mind* 116, 359-370.

Broome, J. (2010): Rationality. In: O'Connor, T. – Sandis, C. (eds.): *A Companion to the Philosophy of Action.* Blackwell, 285-292.

Brunero, J. (2010): The Scope of Rational Requirements. *Philosophical Quarterly* 60, 28-49.

Coates, A. (2012): Rational Epistemic Akrasia. *American Philosophical Quarterly* 48, No. 2, 113-124.

Jones, K. (2003): Emotion, Weakness of Will, and the Normative Conception of Agency. *Royal Institute of Philosophy Supplement* 52, 181-200.

Kolodny, N. (2005): Why be Rational? *Mind* 114, 509-563.

Kolodny, N. (2007): State or Process Requirements? *Mind* 116, 371-385.

McNamara, P. (2010): Deontic Logic. In: Zalta, E. N. (ed.): *The Stanford Encyclopedia of Philosophy* (*Fall 2010 Edition*). Available: http://plato.stanford.edu/archives/fall2010/entries/logic-deontic/.

MacIntyre, A. (1990): Is Akratic Action Always Irrational? In: Flanagan, O. – Rorty, A. (eds.): *Identity, Character, Morality.* Cambridge: Cambridge University Press, 379-400.

Perry, J. (1996): How to Procrastinate and Get Stuff Done. In: *The Chronicle of Higher Education.* Washington, D.C. (23 February 1996, 15 January 2012) Available: http://chronicle.com/section/About-the-Chronicle/83

Railton, P. (1986): Moral Realism. *Philosophical Review* 95, No. 2, 163-207.

Setiya, K. (2007): Cognitivism about Instrumental Reason. *Ethics* 117, No. 4, 649-673.

Smith, M. (1995): Internal Reasons. *Philosophy and Phenomenological Research* 55, No. 1, 109-131.

Tappolet, C. (2003): Emotions and the Intelligibility of Akratic Action. In: Stroud, S. – Tappolet, C. (eds.): *Weakness of Will and Practical Irrationality.* Oxford: Oxford University Press, 97-120.

Titelbaum, M. (ms.): How to Derive a Narrow-Scope Requirement from Wide-Scope Requirements.

Watson, G. (1975): Free Agency. *Journal of Philosophy* 72, No. 8, 205-220.

Williams, B. (1981): Internal and External Reasons. In: *Moral Luck.* Cambridge: Cambridge University Press, 101-113.

Against Normative Judgement Internalism

MIRANDA DEL CORRAL

CONICET – Universidad de Buenos Aires
Puán 480. 1406 Ciudad Autónoma de Buenos Aires. Argentina
mdelcorral@bec.uned.es

ABSTRACT: Normative judgement internalism claims that enkrasia is an ideal of rational agency that poses a necessary link between making a normative judgement, and forming an intention to act according to that judgement. Against this view, I argue that enkrasia does not require the formation of new intentional states; instead, it requires that the agent's intentions do not contravene her normative judgements. The main argument for considering that an intention ought to follow from a normative judgement is the claim that the conclusion of practical reasoning is an intention. I will argue that this account is mistaken: practical reasoning aims at justifying certain actions or intentions, and thus its conclusion is a normative judgement. Defenders of NJI might argue, though, that intentions ought to follow from our normative judgements, because of certain requirements affecting not only practical reasoning, but rational agency. I argue that this conception of enkrasia is too demanding. Enkrasia, I suggest, is better understood as a restriction over our intentions: they ought not enter into conflict with our judgements.

KEYWORDS: Akrasia – enkrasia – normative judgements – normative requirements – practical reasoning.

Introduction

Akrasia is one amongst various rationality failures an agent can commit. The very possibility of irrationality posits the existence of certain "gaps" (Searle 2001), or "breaks" (Hinchman 2009), in the transition between the steps of rational agency. At least two gaps have been identified. The first of

them lies between the agent's deliberation and her choice. To make a choice means to make the decision to act. In this sense, judgement and choice are two distinct kinds of decisions. Judging that we should φ is equal to deciding that we should φ; choosing to φ is deciding to φ, this is, forming the intention to φ (Holton 2009). The second gap in rational agency takes place between the agent's intention to perform an action in the future – this is, her choice to act – and the actual performance of the action. This gap allows for a different kind of irrationality: the agent intends to φ, but intentionally does not-φ. Following this distinction between the two kinds of gaps, Holton argues for two different kinds of irrationality: 'akrasia' and 'weakness of will'; similarly, Hinchman distinguishes between 'incontinence' and 'weakness'. Rationality governs agency through requiring an appropriate relation between reasons, judgements, intentions and actions. Particularly, enkrasia governs the first of the two gaps: it states that rationality requires that the agent's judgements and her intentions stand in an appropriate relation.

Traditionally, akrasia has been conceived as the failure to do, or at least to intend to do, what one judges one ought to do, all things considered:

> [for traditional conceptions] an agent who decisively judges it best to A is thereby rationally committed to A-ing, in the sense that (as long as the judgment is retained) the uncompelled, intentional performance of any action that he believes to be incompatible with his A-ing would open him to the charge of irrationality. (Mele 1995, 71)

Akrasia constitutes a violation of certain rational requirements, i.e. the norms of rationality governing agency; particularly, akrasia is the violation of enkrasia. This requirement states that, in order to be rational, an agent's normative judgements and intentions must stand in an appropriate relation. The question, then, is what the best way to formulate this appropriateness would be – what exactly rationality requires from us.

1. Enkrasia and Normative Judgement Internalism

According to Hinchman (2009), there are three different species of internalism; two of them correspond to the first gap of rational agency – between normative judgements and intentions – and the last one corresponds to the second gap – between intentions and actions:

> We might call 'judgment internalism' the thesis that your all-things-considered judgment about what you should do bears an internal relation to your choice or intention to act accordingly. And we might call 'volitional internalism' the thesis that your choice or intention bears an internal relation to such a judgment. To call these relations 'internal' is to say that they're in some sense necessary or non-empirical [...] 'Resolve internalism' [is] the thesis that intending, resolving or otherwise willing to φ bears an internal relation to actually φ-ing (or at least attempting to). (Hinchman 2009, 396)

Let us postpone the third requirement ('resolve') for the moment, and focus on the first and second requirements. Judgement internalism states that an agent ought to intend to do what she judges best. It is therefore irrational to judge that one ought to φ and, at the same time, not to intend to φ. Volitional internalism, on the other hand, states that the agent's intention to φ ought to stand in an appropriate relation to her judgements about whether she ought to φ. These two kinds of internalism refer to two different links in the first gap, 'forwards' and 'backwards' (Hinchman 2009, 424). The difference between them is often neglected; in fact, intending to φ while believing that one ought not to φ constitutes a violation of both these kinds of requirements. Although Hinchman does not further develop this double distinction, I believe he pinpoints a relevant feature of the first gap in rationality: that rational agency is not necessarily a one way process, and therefore rationality ought to cover both the forwards and the backwards process. We frequently revise our intentions, not only in the light of new information, but also to check whether they interfere with our other plans. Also, we receive advice, and we are given new reasons we had not considered before. In fact, changing our minds is also subject to the requirements of rationality. Let us now focus on judgement internalism.

Normative judgement internalism[1] (NJI) postulates an internal link between a normative judgement and its subsequent intention: "Necessarily, if one judges anything of the form 'I ought to φ', then one also has a general disposition to intend to do whatever one judges that one ought to do" (Wedgwood 2007, 28). Following NJI, rationality requires one not to be

---

[1]   The label 'normative' does not add any significant constraint to 'judgement internalism', given that only all-things-considered practical judgements are necessarily connected with intentions. I have chosen to employ the terminology used by Wedgwood (2007) because I will focus on his formulation of the requirement.

akratic in the traditional sense pointed out above: "that you not judge, all things considered, that you ought to φ while failing to choose or intend to φ" (Hinchman 2012, 1).

Much of the recent debate on rational requirements has focused on the logical relation between the antecedent ('I ought to φ') and the consequent ('I intend to φ'). In broad terms, there are two opposing alternatives: wide-scope and narrow-scope formulations.[2] The distinction goes as follows:

> **Enkrasia** (Narrow-scope):
> If you believe that you ought to φ, then you are rationally required to intend to φ.

> **Enkrasia** (Wide-scope):
> Rationality requires that [if you believe that you ought to φ, then you intend to φ].

The difference between narrow and wide-scope enkrasia lies in the possibilities that an agent has available when she finds herself in a situation of irrationality. Suppose that an agent realises that she believes she ought to φ, and nevertheless she does not intend to φ. Following narrow-scope enkrasia, she is required to intend to φ. However, from a wide-scope perspective, the agent has a choice: she can either form the intention to φ, or revise (and ultimately, abandon) her judgement that she ought to φ. In both scenarios, the agent is considered to be in an irrational state when she judges she ought to φ while not intending to φ. The narrow-scope formulation of enkrasia is a process-requirement:[3] it tells the agent what to do in order to avoid irrationality – namely, to form the intention to do what she judges best.[4] Conversely, wide-scope enkrasia is a state-requirement: it only

---

2 This debate has generated a growing amount of articles; the beginning of the discussion can be found in Broome (1999; see also 2007); this discussion has been continued by Kolodny (2005), Setiya (2007), Brunero (2010), Way (2010) and Hinchman (2012), amongst others.

3 Rationality may be regarded as a state-requirement or as a process-requirement (Kolodny 2005). While state-requirements are requirements over the agent's actual attitudes, which are required to be consistent, process-requirements demand a certain response from the agent, a change in her attitudes.

4 However, as Lord (2011) points out, exiting from the requirement – i.e. making it no longer apply to us through denying the antecedent – is not a form of complying with it, but it does not violate it either. Therefore it is also possible, under a narrow-scope

demands avoiding incoherence, either through the formation of an intention, or through the denial that one ought to φ. Wide and narrow formulations mainly differ in the possibility of detaching the conclusion. What is important for our purpose is that, under both of these formulations, it is irrational not to intend to φ while judging one ought to φ; thus, my critique to NJI applies to both narrow and wide-scope formulations.

I will now examine two different arguments supporting NJI over alternative conceptions of the enkratic requirement. First, it has been argued that enkrasia governs the transition from certain premises (reasons or normative judgements) to a practical conclusion (an intention): this is what practical reasoning consists in. Against this, I will argue in the next Section that the conclusion of practical reasoning is not an intention, but a normative judgement. The second argument for NJI aims to overcome this challenge, and will be explored in Section 3. Even if the conclusion of practical reasoning is not an intention, it could be argued, the enkratic requirement as defined by NJI may still be an appropriate formulation of the relation between judgements and intentions, insofar as a violation of enkrasia commonly leads to be considered irrational. Against this argument, I will suggest some examples in which the enkratic requirement is not fulfilled, and nonetheless the agent's irrationality is not straightforward. In Section 4, I suggest that volitional internalism can overcome the challenges to NJI presented in this paper, while preserving a normative relation between judgements and intentions.

2. The argument from practical reasoning

Practical reasoning contrasts with theoretical reasoning in the following way: while the former is directed towards action, the latter aims to elucidate how the facts stand. This general remark has led a majority of philosophers to claim that the conclusion of practical reasoning is an intention.[5]

---

requirement, to change one's mind about what one ought to do in order to avoid irrationality.

[5]   Some philosophers argue for a stronger claim: that the conclusion of practical reasoning is an action (Dancy 2004; Tenenbaum 2007), or either a decision or an action (Alvarez 2010). However, defendants of this view would also accept that intention is involved at some point in the process: they refer to an *intentional* action. See Streumer (2010) for an overview of the problem.

Practical reasoning, according to Broome (2001), can be either instrumental or normative. Instrumental reasoning consists of an inference on behalf of an agent, in which she takes her intentions and beliefs as premises, and reaches a certain intention as the conclusion of this reasoning process (Broome 2001, 176):

Premise 1:   (I)   I am going to leave the next buoy to starboard.
Premise 2:   (B)   In order to leave the next buoy to starboard, I must tack,
Conclusion: (I)   I shall tack.

This example represents a piece of instrumental practical reasoning. Normative reasoning, on the other hand, involves normative judgements as premises (Broome 2001, 181):

Premise 1:   (B)   I ought to tack.
Conclusion: (I)   I shall tack.

In both cases, the agent would conclude her practical reasoning with the formation of an intention; instrumental reasoning involves other intentions and beliefs, while normative reasoning also includes an all-things-considered normative judgement. Instrumental reasoning may include normative judgements, but they are conditional on the agent's goals: if the agent does no longer intend to leave the next buoy to starboard, then she ought not to tack. On the contrary, normative judgements depend on the agent's reasons, which would be independent of her intentions.

This is the strong thesis of NJI: normative judgements ought to motivate; or, to put it differently, using Broome's terminology, rationality requires that, if an agent believes that she ought to φ, all things considered, then that agent is motivated to φ (which is a prerequisite to intending to φ). This motivation can, of course, enter into conflict with other motivations the agent has. NJI states that normative judgements should, at least, have some motivational force. According to the stronger version of this thesis, that is, that normative judgements should motivate the agent more than whatever other motivations she has, akrasia would be rendered impossible.[6]

---

[6]   For a discussion on the stronger version of NJI, see Wedgwood (2007, 26).

What about second and third-person beliefs and normative judgements? Broome argues that reasoning about what another person will do constitutes theoretical reasoning, because its conclusion is a belief (Broome 2001, 177):

Premise 1: (B) Leslie will leave the next buoy to starboard.
Premise 2: (B) In order to leave the next buoy to starboard, Leslie must tack.
Conclusion: (B) Leslie will tack.

I fully agree with Broome that this example illustrates a case of theoretical reasoning, taking two premises as evidence for believing that Leslie will, in fact, tack. Now, let us turn to normative reasoning: is it theoretical as well? If I believe that Leslie has a normative reason to tack, and judge that, all things considered, Leslie ought to tack, is my reasoning theoretical? My aim is to show that it is not: reasoning about what others ought to do is a form of practical reasoning, just as much as reasoning about what I ought to do is. However, third-person reasoning does not conclude in an intention – I cannot intend you to do something.[7] Therefore, I will suggest, the conclusion of practical reasoning is a normative practical judgement.

2.1. Second and third-person practical reasoning

We frequently give and ask for advice, express our opinions about what someone else should do, and criticise others for not behaving as they ought to. Imagine, for instance, that a father advises his son: "Son, you should study law rather than philosophy: it is much better paid." The reasoning this father has done regarding his son's academic future is similar to the one he would have done when assessing whether to study philosophy himself. However, it could be argued that reasoning about what someone else should do is in fact theoretical, and not practical. For example, Álvarez (2010) states that "practical reasoning presupposes a goal in the person who engages in the reasoning, which is precisely the thing wanted and what gives the point of the reasoning and of the action to which the reasoning leads" (Álvarez, 2010, 367). Thus, following Álvarez, practical reasoning serves a practical goal of the agent: an agent wants to φ, reasons in order to

---

7   I can intend to persuade, coerce, suggest, etc. you to perform an action, but I cannot intend to perform your action as you.

know what she ought to do in order to attain φ (let us suppose that she believes she ought to χ), and concludes her reasoning with the formation of an intention to χ. Hence, she identifies practical reasoning with instrumental reasoning, although the process may include normative judgements as premises as well.

Two objections can be raised against the claim that practical reasoning is necessarily prompted by a need to choose the path of action that will lead the agent to the achievement of her goal. The first of them is suggested by Álvarez herself: "[practical reasoning is driven by the agent's goal] Unless [...] one is just reflecting on how practical reasoning works, or reasoning on someone else's behalf, as a detective might when trying to guess how someone might have acted" (Álvarez 2010, 367). So, there are (at least) two exceptions: exploratory reasoning[8] and second-person reasoning. These two cases are not prompted by the goals the agent is trying to achieve through reasoning; in order to accommodate these two exceptional cases, further clarification on why they are categorised as practical reasoning is needed. Second, it could be argued that theoretical reasoning is also prompted by the agent's goals. We may have the goal to know how things stand just because we want to find out the truth, or because we need certain information in order to make a decision. And yet having this goal does not make a certain reasoning practical. Intuitively, first and second-person practical reasoning share common features that could justify considering both kinds of reasoning to be practical – but, what would these features be?

Offering a comprehensive theory about the differences between theoretical and practical reasoning would largely exceed the scope of this paper, and this task is not needed in order to argue for the claim that reasoning about what oneself or others should do is a form of practical reasoning. It suffices to show that first and second personal reasoning share a common feature, which I take to be the feature that makes reasoning practical in the first place.

Let us start by pointing out a common assumption regarding practical reasoning: it is directed towards action. By engaging in practical reasoning, we aim to know what actions are justified or required, what goals are wor-

---

[8]   Exploratory reasoning refers to reasoning whose conclusion needs not to be believed by the agent. Audi (2006, 92) briefly addresses this kind of reasoning, but I will not analyse exploratory practical reasoning here. Second- and third-person practical reasoning could be included in this category.

thy, and what the best way to achieve them is, amongst others. Practical reasoning takes practical reasons as premises. Different agents with access to different reasons may reach a different conclusion. In a very general and broad sense, a reason is the answer to the question "why?", or what "counts in favour of" doing, believing, or feeling something (Scanlon 1998, 17 ff.). The problem of the ontology of reasons is one of the central points of disagreement in the philosophy of action, so I will not present here all its controversial aspects.[9] However, I believe there are good arguments for the thesis that reasons are facts, but facts are not reasons by themselves: the agent must believe that the fact that she uses as a reason is the case,[10] and she must be willing to use this fact in a practical inference.[11] This means that reasons used in practical reasoning are perspective-dependent: they are reasons for the reasoner.

Both first and third-person practical reasoning use reasons as premises; so does theoretical reasoning. However, practical reasons differ from theoretical reasons in that the former are reasons because of some motivational disposition of the agent, while, in the latter case, accepting something as evidence does not require the agent to be motivated to accept it, as Audi (2006) has argued. A practical judgement states what ought to be done. It is the answer to a practical question. And, Audi argues, making these kinds of judgements is distinctive of practical reasoning. Practical judgements aim to solve a practical problem, guiding the agent's actions. However, this seems to pose a problem for the claim that I am trying to defend: that practical reasoning is not necessarily first-personal. Audi acknowledges that there are differences and similarities between first- and second- (or third-) personal practical reasoning, but does not provide an analysis of what they would amount to.

A straightforward similarity between first and second or third-personal practical reasoning is that they all conclude in a practical judgement, which differs from a mere prediction about what oneself, or others, will do. Several facts can be used as evidence of an agent's future behaviour: her past behavioural tendencies, what we know about her motivations, or what we

---

[9]   See, for a comprehensive overview, Everson (2009).

[10]   Is it then possible that there are reasons for an agent that the agent is not aware of? The answer to this question does not affect our concern in this paper: the enkratic requirement is a subjective requirement of rationality, so I leave this question open.

[11]   I am here following Schroeder (2008).

know about the world. I can predict that my sister will be late today because she is always late, or because I know that the train in which she is coming is late. A practical judgement, on the other hand, states what an agent ought to do. In the case of first-personal reasoning, I employ certain facts as normative reasons because I feel motivated, with variable strengths, to comply with the norms from which these reasons gain their normative status. For example, I could judge that the fact I will be late to an appointment unless I take a taxi is a reason for me to take a taxi as long as I find punctuality valuable. Were I indifferent to being late, I could not use it as a normative reason. The main obstacle to my argument is the following: How are values and motivations expressed in second and third-personal practical reasoning?

There are two main differences between first and second or third-personal practical reasoning (Andreou 2006). First, the amount of information available to each agent differs – not only the information about the world, but also about the other agent's motivations. Second, even if the information were shared, the evaluation mechanisms used to assess the reasons for and against a particular goal or action may differ. Suppose that an agent seeks advice. Judgements about what the advisee ought to do can be derived either from the adviser's reasons and standards, or from those attributed by the adviser to the advisee – what is commonly known as 'putting oneself in the other's shoes'. In both cases, the adviser has to engage in practical deliberation in order to assess what the advisee ought to do. Second and third-personal practical reasoning can take the form of either normative or instrumental reasoning. While normative reasoning examines the normative reasons for a certain action, instrumental reasoning takes an end as given, and aims to find out the best means to achieve it.

Thus, motivations and values do play a role in second or third-personal practical reasoning. They may not play the same causal role in the production of an action, but they play a fundamental role in assessing what facts are used as reasons in a practical inference: this is a distinctive feature of practical, as opposed to theoretical, reasoning.

3. The argument from rational agency

Even if the conclusion of practical reasoning were a normative judgement, it could be argued, it does not follow that NJI is false. What follows

is that enkrasia is not a rational requirement governing practical reasoning, because its conclusion is not an intention. But it may well be the case that enkrasia is a requirement over rational agency. NJI defends that rationality requires an agent to form an intention to φ whenever she judges that she ought to φ; even if her reasoning were correct and complete, she would still be considered irrational if she lacked the intention to φ.

However, there are cases in which requiring an agent to form an intention is too demanding as a condition for rationality. We sometimes deliberate on whether we should do something in particular; and the answer can be positive, as well as negative. If I wonder now whether I should φ, and I conclude that I ought not to, requiring me to form the corresponding intention not to φ seems too strong as a condition for rationality. The distinction between the absence of intention and the presence of a negative intention is often overlooked (cf. Kolodny 2005). Suppose that I am visiting the Prado Museum, and while standing in front of Las Meninas, a paper by Sam Shpall comes to my mind. His paper analyses the agential commitments derived from the belief "I ought to spit on Las Meninas" (Shpall 2011). I deliberate on my reasons to spit on Las Meninas, and I find that I do not have any – in fact, I have many reasons not to spit on that painting. Given my reasons, I judge that I ought not to spit on Las Meninas. I can avoid irrationality by merely not intending to do what I believe I ought not to do.

It could be argued that abstaining from doing something is also an action, and therefore there is a practical link between my judgement that I ought not to spit on Las Meninas and my not spitting on them, which would be an action. Some clarification is needed at this point. The concept of omission is problematic. On its widest conception, every absence of action is an omission, and the number of things we do not do is countless. There are different ways, though, to narrow down the set of omissions that count as actions, such as taking into account the agent's intentions (Clarke 2010), or her degree of reason-responsiveness and self-control (Fischer 1997). These criteria serve different purposes, such as attributing responsibility to an agent for not doing something, or analysing the conditions under which omissions are indeed actions. Let us focus on intentional omissions, for the matter discussed here is whether a negative intention (an intention not to do something) ought to follow from a negative normative judgement. My omission to spit on Las Meninas may be intentional or unintentional. Suppose that it is unintentional: I do not have any intention to

spit on that painting, nor any intention not to spit on it. I had previously judged that I ought not to spit on Las Meninas, and the lack of any intentional state whose content is to avoid or omit spitting on it does not constitute a violation of a rational requirement, apparently. Intentions to perform something that we do not usually need to intend to do are similar in what concerns enkrasia, as Wedgwood argues; we do not need to form an intention to breathe in order to do so, and therefore it is not irrational not to form an intention to breathe, even if we judge we ought to breathe (cf. Wedgwood 2007, 30). It follows that I do not need to form an intention not to spit on Las Meninas in order to avoid irrationality: before wondering whether I should do so, I had no previous intention to perform that action. But let us suppose that I actually had that intention. I then ask myself whether I ought to spit on the art work, and conclude that I ought not to. Intuitively, dropping my previous intention suffices in order to avoid irrationality.

It may be argued that NJI can accommodate these exceptions. Wedgwood discusses the possibility of being rational without forming an intention. He argues that NJI only applies to first-person normative judgements whose content (φ) is something "of the appropriate sort", which means that "φ-ing must be a course of action that is 'manifestly dependent on intention', in a situation 'with no relevant uncertainty'" (Wedgwood 2007, 81). Hence, normative judgements of the appropriate sort meet two requirements. On the one hand, the agent must know that her intention to φ makes a difference to the chances of her φ-ing. Furthermore, the agent must know that she will φ if and only if she intends to (Wedgwood 2007, 30). Not intending to φ while knowing that only by intending to φ will one φ, Wedgwood argues, counts as willingly failing to φ, and this is why the agent who does not form the intention is akratic: she is willingly failing to do what she believes she ought to do.[12] Concerning the second condition for appropriateness, it states that the agent must be sure about what her best option is. If the agent knows that she ought to do either φ or χ, but she does not know which one she ought to do, then there is 'relevant uncertainty'. In this situation, the agent is not rationally required to intend to do neither φ nor χ. Similarly, if the agent judges, all things considered, that

---

[12]   However, Wedgwood's conclusion is problematic. It is not necessarily true that, by lacking the intention to φ, one is willingly not φ-ing. I will discuss this problem in the next section.

she ought to φ, but she is unsure about her judgement, then her judgement is not of the appropriate sort.

The criteria for inappropriateness could be broadened in order to accommodate normative judgements concerning what agents other than oneself ought to do, or what we ought to do given certain hypothetical circumstances, which would be a case of exploratory reasoning. If the criteria for appropriateness are narrowed, the challenges I have presented to NJI disappear. However, the concept of "appropriateness" concerning judgements is quite problematic, and hence the strategy to defend NJI is controversial, at the very least. First, the distinction between appropriate and inappropriate judgements ought to be based on a common feature that either all appropriate or all inappropriate judgements share. Because of this feature, appropriate judgements would be subject to a certain rationality norm (enkrasia, as defined by NJI), while inappropriate judgements would not be subject to that rationality norm.[13] In order to explain its capability to interfere with a rational requirement, this common feature (either to appropriate or to inappropriate judgements) should somehow be related to the enkratic requirement.

Second, Wedgwood's criteria for appropriateness are themselves problematic. The first requirement states that the course of action must be 'manifestly dependent on intention', which means that the agent must believe that she will φ if and only if she intends to φ. However, it is not clear whether one's intention necessarily makes an objective significant difference, or, on the contrary, whether an agent whose beliefs concerning success are wrong is also making a judgement "of the appropriate sort". Furthermore, it would be necessary to quantify, or at least specify to a higher degree, what a "significant difference" amounts to. Training to be a professional athlete makes a difference in the chances of success, but the chances are objectively very low; thus, even if an agent believes she ought to become an athlete, rationality (following Wedgwood) would not require her to form an intention to achieve that goal. On the other hand, the certainty requirement needs further clarification. What level of certainty, self-confidence or trust is needed in order to meet this requirement? Suppose that I now judge that I ought to φ, and at the same time, I know that I am likely to change my mind in the future: do I meet the certainty requirement?

---

[13] They might be subject to *other* requirements, although this possibility has not been explored in the literature.

In sum, I think that Wedgwood's strategy needs further development in order to provide a satisfactory explanation of why, under certain circumstances, rationality does not require the agent to form an intention. I will now suggest a different interpretation of the enkratic requirement that allows us to accommodate a broader scope of judgements.

4. Volitional internalism: enkrasia as a restriction

Judgement internalism is one of the two kinds of normative bridges in the first gap of rational agency pointed out by Hinchman (2009). The second bridge would be volitional internalism, which states that our intentions bear an internal relation to our normative judgements. Not intending to φ while believing that one ought to φ conflicts with judgement internalism, but it does not contravene volitional internalism; holding an intention to φ while believing that one ought not to φ conflicts with volitional internalism, and also contravenes judgement internalism. This asymmetry is often neglected, or explicitly rejected: as pointed out in the previous Section, Wedgwood claims that an agent who knows that only by intending to φ she will φ and, at the same time, does not intend to φ, is willingly failing to φ. However, the relation between the absence of intention to φ and the voluntary failure to φ is not straightforward. In this Section, I argue that the absence of intention does not necessarily lead to acting against one's best judgement, and that the enkratic requirement is better understood as a restriction over our intentional states.

Although akrasia has traditionally been defined as lacking the intention to φ while believing that one ought to φ, there is an alternative formulation of akrasia that is based on the irrationality of contravening volitional internalism. Under this view, akrasia would be defined as acting against one's best judgement, that is, intentionally φ-ing, while judging that one ought not to φ (Audi 1993; Gilead 1999; Mele 1995; Tenenbaum 2010). The reason why akrasia is possible is that the evaluation of our normative reasons for φ-ing may not be in line with our motivation to φ:

> [A]ttributing an action-guiding function to evaluative judgements [...] does not commit one to supposing that the judgements are themselves logically or causally sufficient for the presence of corresponding intentions. [...] There is no motivational magic in the thought content 'My A-ing would be best'. (Mele 1995, 25)

Normative judgements can provide rational guidance even when they lack sufficient motivational force (Audi 2006, 81; Railton 2006). This is not to say that normative judgements are completely independent of our motivations. As argued above, our values and motivations are necessarily connected to what we take to be a practical reason. However, many factors may intervene in the motivational strength of a normative judgement: variations in our motivation across inter-temporal agency,[14] conflicts between first- and second-order desires[15] or depletion of self-control,[16] to name a few.

An agent may judge either that she ought to φ (BOφ), or that she ought not to φ (BO¬φ). Akrasia is the result of a conflict between her judgements and her actual intentions. Given that an agent can intend to do something (Iφ), not to do something (I¬φ), or she can lack the relevant intentional state (¬Iφ ∧ ¬I¬φ), the following combinations may obtain:

	Iφ	Enkrasia		Iφ	Akrasia
BOφ	I¬φ	Akrasia	BO¬φ	I¬φ	Enkrasia
	(¬Iφ ∧ ¬I¬φ)	?		(¬Iφ ∧ ¬I¬φ)	?

Table 1: Combinations of judgements and intentions

There are four uncontroversial cases: two of akrasia, and two of enkrasia – either the agent's intentions are in line (enkrasia) or contradict (akrasia) the agent's normative judgement. Akrasia would consist in intentionally φ-ing, or holding an intention to φ, while believing that one ought not to φ. Conversely, an agent would also display akrasia if she intends not to φ, or intentionally does not φ, while believing that she ought to φ. The rational requirement that is violated by akrasia can be formulated as follows:

---

[14]   For instance, we can predict that we will be glad we did something, as the title of Harman's article suggests (2009), even if we do not feel a present motivation to do it.

[15]   A second-order desire is a desire to have a desire ("I wish I cared more about my health"). Following Frankfurt (1971), second-order desires guide first-order desires; they also control the formation of intentions.

[16]   The process of exhaustion of willpower has been called "ego depletion" (Baumeister – Vohs – Tice 2007).

Enkrasia:
Rationality requires that you do not [intentionally not φ if you believe that you ought to φ].[17]

If akrasia consists in intentionally acting against one's best judgement, the following question arises: Why is it irrational to *intend to act* against our judgements? Suppose that I believe I ought not to φ; nonetheless, I intend to φ. However, I finally drop my intention, and therefore I am not acting against my judgement. Does not merely holding an intention to φ make me irrational, in the light of my normative beliefs? The answer to that question is affirmative. However, the explanation of why this is so requires us to introduce a different rational requirement, governing the second break in rational agency, the gap in the transition from intentions to actions. Following Hinchman (2009), resolve internalism states that our intentions bear an internal relation to our actions. The resolve requirement may be formulated as follows:

Resolve:
Rationality requires that [if you intend to φ, then you intentionally φ].[18]

Resolve, then, is a process-requirement. It states that our intentions ought to have volitional control over our actions: intending to φ commits us to φ-ing. Intentions, as opposite to normative judgements, ought to be motivationally sufficient to initiate action.

---

[17]   This is a wide-scope and positive formulation of enkrasia. If the normative judgement is negative, the wide-scope formulation would go as follows:

Enkrasia (wide−): Rationality requires that you do not [intend to φ if you believe that you ought not to φ].

Enkrasia could also be formulated as being narrow-scoped:

Enkrasia (narrow+): If you believe that you ought to φ, then rationality requires that you do not [intend not to φ].

Enkrasia (narrow−): If you believe that you ought not to φ, then rationality requires that you do not [intend to φ].

Discussing here the scope of rational requirements would exceed the scope of this work; my aim is just to show that the approach to enkrasia I am defending here might, in principle, be formulated in different ways.

[18]   Again, this formulation is wide-scoped; but a narrow-scoped reformulation would also be compatible with the argument presented here.

Through the resolve requirement, we may explain why intending to φ while believing that one ought not to φ is akratic. Suppose I believe I ought not to φ. Because of enkrasia, it is prohibited for me to intentionally φ. However, I intend to φ. Then, rationality requires that I φ intentionally, because of the resolve requirement – otherwise, I would be weak-willed (Holton 2009). Rationality would then be requiring me to intentionally φ and, at the same time, prohibiting me to do so. Both enkrasia and resolve regulate intentions: the former is a backwards process (from intentions to normative judgements, in the form of a restriction) and the latter is a forwards process (from intentions to action, in the form of a positive requirement). Therefore, holding an intention to do something that contravenes our judgements is also a form of akrasia.

Thus, the enkratic requirement is rather a prohibition: we ought to avoid intending to do anything that contravenes our normative judgements. Not intending to do what we believe we ought to do is not necessarily irrational. This claim may seem quite counterintuitive, so let us illustrate it with a few examples.

First, let us bring back our example above concerning Las Meninas. I believe I ought not to spit on it. And yet I do not have any intention not to spit on it, nor any intention to do so. However, I am not contravening my normative judgement: simply by not doing it I am complying with that judgement, i.e. I am not acting against it. Similarly, I do not need to intend not to lie every time I speak, even if I believe I ought not to lie. I agree that judgements to perform something we were going to do regardless of our intentions are uncommon.

So, let us imagine a different scenario. I deliberate about the relation that my sister and I have. I conclude, all things considered, that I ought to meet her more often. However, I know I will organise my agenda tomorrow, so I prefer to wait until then in order to choose when to meet her. I have not formed the intention to meet her, and I do not think I am being akratic (nor that my reasoning is not complete). It may be argued, though, that I have indeed made a choice, and thus formed an intention, to meet my sister: all I have to do is to plan how to do it. However, an intention is a mental state that exerts volitional control over our mental states and our behaviour (Bratman 2009). It is not merely a wish, or a desire, or a belief about what I will do. In this example, I do not form an intention [to meet my sister at some point]; rather, I have suspended choice, which does not necessarily violate enkrasia.

Let us imagine a third scenario. I believe I ought to meet my sister by the end of this week. The week ends and I have not decided whether to meet my sister or not – that is, I have not formed any intention. However, I have been doing many things that are incompatible with meeting my sister, and I am aware of this incompatibility. In this case, I would consider that I am being akratic. The reason is that I am intentionally not doing something I believe I ought to do. Therefore, I am acting against my all-things-considered judgement.

Lastly, let us imagine the following variation of the last scenario. On Monday, I believe I ought to meet my sister by the end of the week, but I have not formed the subsequent intention. On Saturday, I realise that the end of the week is near, and then decide to meet my sister that day. Am I being irrational from Monday to Saturday? No, because I have not acted against my best judgement.

I have tried to show through these examples that it is possible to judge that we ought to do something and, at the same time, not intending to do it, without violating a rationality requirement. Granted, if the normative judgement that the agent makes has the form "I ought to φ now", then there is no room for delaying her decision: either she acts upon her judgement, or she changes her mind about what she ought to do. NJI can only deal with this kind of normative judgements. Volitional internalism, on the contrary, can deal with a much wider range of judgements. Of course, many normative judgements have a deadline, so reaching that date and not having neither intended to φ nor having intentionally φ-ed counts as acting against one's judgement that one ought to φ. Therefore, in many occasions, not intending to φ amounts to willingly failing to φ, as Wedgwood claims. However, there is not a *necessary* connection between the absence of intention and an intentional failure to act, and this is the reason why we cannot base a rationality requirement upon this contingent relation.

Thus, only one combination between intentions and judgements violates enkrasia, and is therefore akratic:

	Iφ	Enkrasia-derived		Iφ	Akrasia
Bφ	I$\neg\varphi$	Akrasia	B$\neg\varphi$	I$\neg\varphi$	Enkrasia-derived
	$\neg$Iφ	Enkrasia-compatible		$\neg$Iφ	Enkrasia
	$\neg$I$\neg\varphi$	Enkrasia		$\neg$I$\neg\varphi$	Enkrasia-compatible

Table 2: Akrasia and three kinds of enkrasia

An akratic agent acts against her best judgement, or intends to do something that contravenes her judgement. Concerning enkrasia, I have made a threefold distinction: enkrasia, enkrasia-derived, and enkrasia-compatible. In fact, neither of these three combinations violates enkrasia; the distinction aims to highlight the different ways there are to comply with a negative requirement, i.e. a prohibition.

Firstly, an agent is enkratic as long as she does not intend to do something she believes she ought not to do, and vice versa: she does not intend not to do something she believes she ought to do. Thus, if we lack the intention to contravene our judgement, we are being enkratic.

Secondly, if an agent intends to φ, then rationality requires that she does not intend not to φ; otherwise, she would have inconsistent intentions. Therefore, an agent who intends to φ and believes she ought to φ is not violating enkrasia – she is not acting against her better judgement. The label 'enkratic-derived' aims to stress that the rationality of acting according to one's judgements is derived from complying with the enkratic requirement: if an agent is making a decision to act, and she believes she ought to φ, she is not permitted to intend not to φ, and thus she only has one choosable path: intending to φ.

Finally, an agent is in a state that is compatible with enkrasia ('enkrasia-compatible') if she does not intend to do something she believes she ought to do. It is compatible with enkrasia because we do not have enough information to know whether she is being akratic. As I have argued above, there are cases in which the absence of intention does not necessarily lead to willingly failing to do what one judges best. If the situation is such, then she is not violating enkrasia, and therefore she is not being akratic. Conversely, if her normative judgement would be contravened if she does not form an intention, as in the third example above, in which the judgement has a deadline for fulfilment, then she would violate enkrasia; as we do not know what kind of judgement the agent has made, her state is in principle compatible with enkrasia.

Conclusion

The aim of this paper has been to argue that Normative Judgement Internalism cannot accommodate certain kinds of normative judgements, and that volitional internalism is an alternative view on the normative require-

ments of rationality that succeeds in accommodating these judgements. It is widely assumed that not intending to act necessarily amounts to willingly failing to act. Against this intuition, I have argued that, although frequent, the relation between the absence of intention and an intentional failure is not necessary, and therefore cannot be formulated as a normative requirement of rationality. Furthermore, volitional internalism does not rely on the assumption that the conclusion of practical reasoning is an intention, which I have argued to be wrong. The enkratic requirement formulated by volitional internalism gathers an important and often neglected thesis in the theories on rationality: that normative judgements do not need to cause an intentional state in order to provide rational guidance.

Acknowledgment

References

ÁLVAREZ, M. (2010): Reasons for Action and Practical Reasoning. *Ratio* 23, No. 4, 355-373.

ANDREOU, C. (2006): Standards, Advice, and Practical Reason. *Journal of Moral Philosophy* 3, No. 1, 57-67.

AUDI, R. (1993): *Action, Intention, and Reason*. Ithaca, NY: Cornell University Press.

AUDI, R. (2006): *Practical Reasoning and Ethical Decision*. London: Routledge.

BAUMEISTER, R. F. – VOHS, K. D. – TICE, D. M. (2007): The Strength Model of Self-Control. *Current Directions in Psychological Science* 16, No. 6, 351-355.

BRATMAN, M. (2009): Intention, Practical Rationality, and Self-Governance. *Ethics* 119, No. 3, 411-443.

BROOME, J. (1999): Normative Requirements. *Ratio* 12, No. 4, 398-419.

BROOME, J. (2001): Normative Practical Reasoning. *Aristotelian Society Supplementary Volume* 75, 175-193.

BROOME, J. (2007): Wide or Narrow Scope? *Mind* 116, 359-370.

BRUNERO, J. (2010): The Scope of Rational Requirements. *The Philosophical Quarterly* 60, No. 238, 28-49.

CLARKE, R. (2010): Intentional Omissions. *Noûs* 44, No. 1, 158-177.

DANCY, J. (2004): *Ethics without Principles*. Oxford: Oxford University Press.

EVERSON, S. (2009): What are Reasons for Action? In: Sandis, C. (ed.): *New Essays on the Explanation of Action*. New York: Palgrave Macmillan.

FISCHER, J. M. (1997): Responsibility, Control, and Omissions. *The Journal of Ethics* 1, No. 1, 45-64.

FRANKFURT, H. G. (1971): Freedom of the Will and the Concept of a Person. *The Journal of Philosophy* 68, No. 1, 5-20.

GILEAD, A. (1999): How is Akrasia Possible After All? *Ratio* 12, No. 3, 257-270.

HARMAN, E. (2009): "I'll Be Glad I Did It" Reasoning and the Significance of Future Desires. *Philosophical Perspectives* 23, No. 1, 177-199.

HINCHMAN, E. S. (2009): Receptivity and the Will. *Noûs* 43, No. 3, 395-427.

HINCHMAN, E. S. (2012): Rational Requirements and "Rational" Akrasia. *Philosophical Studies*, 1-24. doi:10.1007/s11098-012-9993-5

HOLTON, R. (2009): *Willing, Wanting, Waiting*. Oxford – New York: Oxford University Press.

KOLODNY, N. (2005): Why Be Rational? *Mind* 114, No. 455, 509-563.

LORD, E. (2011): Violating Requirements, Exiting from Requirements, and the Scope of Rationality. *The Philosophical Quarterly* 61, No. 243, 392-399.

MELE, A. R. (1995): *Autonomous Agents: From Self-Control to Autonomy*. Oxford – New York: Oxford University Press.

RAILTON, P. (2006): Normative Guidance. In: Shafer-Landau, R. (ed.): *Oxford Studies in Metaethics*. Oxford: Oxford University Press, 3-34.

SCANLON, T. (1998): *What We Owe to Each Other*. Cambridge (MA.): Harvard University Press.

SCHROEDER, M. (2008): Having Reasons. *Philosophical Studies* 139, 57-71.

SEARLE, J. R. (2001): *Rationality in Action*. Cambridge (MA.): MIT Press.

SETIYA, K. (2007): Cognitivism about Instrumental Reason. *Ethics* 117, No. 4, 649-673.

SHPALL, S. (2011): Wide and Narrow Scope. *Philosophical Studies* 163, No. 3, 717-736.

STREUMER, B. (2010): Practical Reasoning. In: Timothy, O. – Sandis, C. (eds.): *A Companion to the Philosophy of Action*. Malden (MA.): Wiley-Blackwell, 244-251.

TENENBAUM, S. (2007): The Conclusion of Practical Reason. In: Tenenbaum, S. (ed.): *New Essays on Moral Psychology*. Amsterdam: Rodopi.

TENENBAUM, S. (2010): Akrasia and Irrationality. In: Timothy, O. – Sandis, C. (Eds.): *A Companion to the Philosophy of Action*. Malden (MA.): Wiley-Blackwell, 274-281.

WAY, J. (2010): Defending the Wide-Scope Approach to Instrumental Reason. *Philosophical Studies* 147, No. 2, 213-233.

WEDGWOOD, R. (2007): *The Nature of Normativity*. Oxford: Oxford University Press.

Intentions, Akrasia, and Mere Permissibility

JONATHAN WAY

Philosophy. Faculty of Humanities. University of Southampton
Southampton SO 17 1BF. United Kingdom
j.way@soton.ac.uk

ABSTRACT: Something is wrong with akrasia, means-end incoherence, and intention inconsistency. This observation has led many philosophers to postulate 'wide-scope' requirements against these combinations of attitudes. But other philosophers have argued that this is unwarranted. They claim that we can explain what is wrong with these combinations of attitudes by appealing only to plausible independent claims about reasons for particular beliefs and intentions. In this paper, I argue that these philosophers may well be right about akrasia but that they are wrong about means-end incoherence and intention inconsistency. While it is plausibly impossible to be akratic while having no specific attitude (or lack of an attitude) that you should not have, it is possible to be means-end incoherent or to have inconsistent intentions while having no specific attitude you should not have. There is thus a strong motivation for accepting wide-scope requirements against means-end incoherence and intention inconsistency which does not apply to akrasia. This result offers support to a view of means-end coherence and intention consistency I have defended elsewhere.

KEYWORDS: Akrasia – intentions – mere permissibility – objectivism and perspectivism about 'ought' – rational requirements.

The akratic agent believes that he should A but does not intend to A. It seems clear that something is wrong with the akratic agent – his attitudes do not fit together in the way that they should. In this respect, akrasia is similar to several other problematic combinations of attitudes. For instance, something is wrong with the agent who has inconsistent intentions,

or who fails to intend what he takes to be the necessary means to an in-
tended end. Again, the attitudes of these agents do not fit together in the
way that they should.[1]

To explain what is wrong with these combinations of attitudes, we
might suppose that having certain attitudes requires you to have, or lack,
certain others. For instance, we might suppose that believing that you
should A makes it the case that you should intend to A, or that intending
an end makes it the case that you should intend what you take to be the
necessary means to that end. But on reflection, this idea does not seem very
plausible. Someone who believes that they should jump over the moon
does not thereby make it the case that this is what they should intend to
do. Someone who intends to assassinate the president does not thereby
make it the case that they should intend to hire an assassin. It is not so
easy to "bootstrap" such requirements into existence.[2]

Many philosophers have taken observations of this sort to motivate the
idea that there are "wide-scope" requirements against the problematic com-
binations – against akrasia, means-end incoherence, and intention inconsis-
tency:

> (Enkrasia Wide): You should not [believe that you should A and not
> intend to A].
>
> (Intention Consistency Wide): You should not [intend to A, believe
> that you cannot both A and B, and not intend to B].
>
> (Means-End Wide): You should not [intend to E, believe that M-ing is
> necessary for E-ing, and not intend to M].[3]

Since these requirements simply prohibit the problematic combinations,
they do not entail that having certain attitudes requires us to have others.
For instance, Enkrasia Wide does not entail that if you believe that you
should A, then you should intend to A. After all, if you believe that you
should A, there are two ways you can come to comply with Enkrasia Wide

---

[1]   I use 'attitudes' in a broad sense in which absences of beliefs and intentions count as
attitudes.

[2]   See Bratman (1987) and Broome (1999) for especially influential versions of this
point.

[3]   Broome (1999) makes this case in an especially clear and powerful way. See Way
(2010) for further references.

– by forming the intention to A or by dropping the belief that you should A. Enkrasia Wide only entails that you should do one or other of these things.

However, other philosophers have argued that we can explain what is wrong with the problematic combinations without accepting wide-scope requirements, and without allowing for objectionable bootstrapping. According to the view I shall call *disjunctivism*, each of the problematic combinations guarantees that you go wrong in some more specific way – e.g. that you believe something you should not or fail to intend something which you should. In the cases of interest here, the disjunctivist claims that:

> (Enkrasia Disjunctive): If you are akratic, then *either* you should not believe that you should A *or* you should intend to A.
>
> (Means-End Disjunctive): If you are means-end incoherent, then *either* you should not intend to E *or* you should not believe that M-ing is necessary for E-ing *or* you should intend to M.
>
> (Intention Consistency Disjunctive): If you have inconsistent intentions, then *either* you should not intend to A *or* you should not believe that you cannot both A and B *or* you should not intend to B.

The disjunctivist holds that each of these claims follows from plausible independent claims about reasons for belief and intention. He then suggests that it is these claims which explain what is wrong with the problematic combinations. For the disjunctivist, the problem with the akratic (means-end incoherent, intention inconsistent) agent is not the way in which he combines his attitudes. It is instead an ordinary failure to conform to the balance of reasons bearing on belief or intention – the kind of failure you might make even if you did not exhibit the problematic combination. So there is no need to posit wide-scope requirements to explain what is wrong with akrasia, means-end incoherence, and intention inconsistency. Ordinary reasons for belief and intention already ensure that something is wrong with these combinations.[4]

---

[4]   Raz (2005) makes this case with respect to means-end incoherence. Kolodny (2007; 2008a; 2008b; 2009) extends the approach to intention inconsistency, akrasia, and other 'coherence requirements'. Lord (2013), Schroeder (2009), and Skorupski (2010, 102) develop broadly similar approaches. Wedgwood (2003) presents a disjunctive account of akrasia.

The disjunctivist faces two tasks. First, it needs to be shown that the problematic combinations do ensure that you go wrong in some more specific way – that each of the above disjunctions of requirements hold. Second, it needs to be made plausible that it is this which explains what is wrong with akrasia, means-end incoherence and intention inconsistency. In this paper I consider the first of these tasks, leaving the second for another time (see Way forthcoming). In section one, I argue that the disjunctivist is plausibly right about akrasia – there is a strong case that akrasia ensures that you believe something you should not or fail to intend something you should. However, I then argue, in sections two and three, that the disjunctivist is wrong about means-end incoherence and intention inconsistency. It is perfectly possible to exhibit these combinations of attitudes without going wrong in any more specific way. If this is right, then there is a strong motivation for accepting wide-scope requirements against means-end incoherence and intention inconsistency which does not apply to akrasia. This result puts pressure on the common assumption that we should give parallel explanations of what is wrong with these combinations. In the final section, I argue that this point offers support to a view of the requirements of means-end coherence and intention consistency I have defended elsewhere.

Two preliminary points are in order. First, it might be thought that, rather than offering a way to dispense with wide-scope requirements, disjunctivism in fact offers a way to vindicate such requirements. After all, given standard deontic logic, the disjunctions of requirements above entail the corresponding wide-scope requirements. However, and leaving aside the point that standard deontic logic is rightly controversial, this seems a mistake. Those who put forward wide-scope requirements do not merely claim that such requirements are true. They claim that these requirements explain what is wrong with the problematic combinations. But this could not be said of wide-scope requirements which are merely trivial consequences of the above disjunctions of requirements. So even if disjunctivism entails wide-scope requirements, it does not entail what we can call the wide-scope *view* (cf. Kolodny 2007, n.18; Way forthcoming).

Second, I assume that what you should do and believe is determined by your reasons – that you should A if you have most reason to A and that it is permissible for you to A if you have sufficient reason to A. However, philosophers sometimes distinguish between 'objective' and 'subjective' reasons, and corresponding senses of 'should'. Similarly, it is common in the literature we are concerned with to distinguish between the attitudes you

should have and those that *rationality requires* you to have. (Sometimes, but not always, this distinction maps onto the former distinction).[5] However, while I am sympathetic to some versions of these distinctions, I will not rely on them in the bulk of the paper.[6] For present purposes, it does not much matter whether wide-scope requirements would be better understood as claims about what rationality requires. Similarly, it will simplify matters to assume that the disjunctivist only recognises one relevant sense of 'should' and 'reason'. With mild circularity, we can identify the relevant sense of 'should' as that which answers the deliberative questions of what to do and believe.[7]

1. Disjunctivism about Akrasia

The akratic agent believes that he should A but does not intend to A. To show that something is wrong with akrasia, the disjunctivist must show that the akratic agent either holds the belief that he should A on insufficient grounds or fails to intend something which he should. So on the assumption that you should intend to A if you should A,[8] what the disjunctivist needs to establish is a kind of *weak infallibilism* about what we should do:

If you permissibly believe that you should A, then you should A.

You *permissibly* believe that you should A if you believe that you should A on the basis of considerations which together give you sufficient reason for

---

[5]   See, e.g. Broome (2005); Kolodny (2005); Parfit (2011); Wedgwood (2003); Schroeder (2004; 2009); Way (2010; 2012).

[6]   One version of the distinction will surface in the final section. I also note points at which such distinctions may be relevant in notes 7 and 10.

[7]   Cf. Kolodny (2007, 232-233). Schroeder (2009) argues that akrasia and means-end incoherence involve a failure to do what you *subjectively* should do. Although I will not be able to discuss this view here, it should be clear to readers familiar with Schroeder's paper that the arguments against disjunctivism about means-end coherence below apply equally to Schroeder's view.

[8]   I do not think this final assumption is true. It fails in cases in which intending to A is not necessary for A-ing. However, the assumption is harmless here, since there need be nothing wrong with cases of akrasia in which you permissibly believe that intending to A is not necessary for A-ing. Cf. Broome (2005, 323).

this belief. In this section I shall argue that weak infallibilism is plausibly true.

The case for weak infallibilism varies depending on whether we accept *objectivism* or *perspectivism* about what we should do and believe. *Objectivism* is the view that what you should do and believe turns on the facts of your situation. *Perspectivism* is the view that what you should do and believe turns on your epistemic position (your 'perspective'). I shall not here adjudicate between these views. Instead, I shall argue that weak infallibilism can be defended on either view.

The route from objectivism to weak infallibilism is straightforward. All objectivists I know of endorse the following:

(Truth Norm): It is permissible to believe that *p* only if *p*.[9]

Weak infallibilism is a trivial consequence of the Truth Norm.[10]

The same holds on some versions of perspectivism. Perspectivists hold that it is permissible to believe only what you are in a good enough epistemic position to believe. However, on an increasingly popular version of perspectivism, you are in a good enough epistemic position to believe *p* just in case you know, or are in a position to know, *p*.[11] Since knowledge is factive, this view also trivially entails weak infallibilism.

---

[9]   For endorsements of the Truth Norm see, e.g. Littlejohn (2010); Shah (2003); Wedgwood (2002); Whiting (2010; and forthcoming).

[10]   It may seem implausible that what is wrong with some cases of akrasia is simply that the akratic agent has a false belief. While this concern falls outside this paper's focus on whether the disjunctivist can show that *something* is wrong with the problematic combinations, it is worth noting that there is more the objectivist can say here. One possibility is to distinguish between what you *objectively* and *subjectively* should do, or what you should do and what rationality requires you to do. On many ways of drawing these distinctions, we will be able to argue from the Truth Norm that the akratic agent does something he *subjectively* should not do, or is irrational (Wedgwood 2003 argues in this way). Another possibility is to argue that what is distinctive of the problematic combinations is not just that you go wrong in some way but that you are in a position to know that this is so (Kolodny 2007; see Way forthcoming for discussion). Either of these approaches will allow the disjunctivist to claim that there is something distinctively wrong with akrasia which does not apply to all cases of false belief.

[11]   See Williamson (2005) for a prominent example. And see Whiting (forthcoming) for discussion and further references.

Things are more complicated on other versions of perspectivism. Weak infallibilism will not be a trivial consequence of views on which you can be in a good enough epistemic position to believe p even when p is false – that is, views which allow for permissible false belief.[12] Nonetheless, I want to suggest that there is still considerable pressure on perspectivists of this sort to endorse weak infallibilism.[13]

My argument for this claim will turn in part on what I shall call the *uniformity thesis*. This holds that we should accept a uniform account of the perspective-relativity of norms of belief and action: we should be perspectivists about norms of belief and action or neither. This thesis is sometimes denied (see, e.g. Feldman 1988b and Skorupski 2010) and I cannot fully defend it here. However, I do want to offer a couple of points in its support.

The first point is that we should not be surprised if the uniformity thesis is true. At a general enough level, we should expect similarities between norms of belief and action. (As Gibbons (2010, 1) puts a related point, 'similarities between practical and theoretical reasons have a built-in explanation: they're both reasons'.) Since the question of whether norms of belief and action turn on your epistemic position occurs at a highly general level, we should not be surprised if this question is answered in the same way in both cases.

The second point is that arguments for and against perspectivism in one domain invariably have analogues in the other. I shall give two examples. First, a standard argument against perspectivism about action turns on

---

[12]   Perspectivists differ about what constitutes your epistemic position and about what it takes to be in a good enough epistemic position to believe something. One way to develop a perspectivist view which allows for permissible false beliefs is simply to allow false beliefs to be included amongst the determinants of your epistemic position – thus consider the view that your epistemic position is constituted by your non-factive mental states. But even if we think that your epistemic position is constituted just by what you *know*, we may still allow for permissible false beliefs if we hold that in order to be in a good enough epistemic position to believe that p, p needs only to be sufficiently probable in light of your epistemic position.

[13]   I am not aware of many perspectivists of this sort who explicitly endorsing weak infallibilism. Kiesewetter (2011, 4) seems to do so, as does Kolodny (2009) and Wedgwood (2003) (although Wedgwood is only a perspectivist about the 'subjective should'). Gibbons (2009, 171-173) and Smithies (2012, 283) defends claims close to weak infallibilism. (Gibbons' argument turns on a wide-scope requirement against akrasia, and so is of little use to the disjunctivist).

the observation that advisors typically take into account what *they* take to be the facts, not just what their advisees take to be the facts, when considering how their advisees should act.[14] To the extent that this observation counts against perspectivism about action, it also counts against perspectivism about belief: advisors also typically take into account what they take to be the facts when considering what their advisees should believe (cf. Thomson 2008, 225). Second, a standard argument for perspectivism about action turns on cases in which the only sensible option is to do what, by your own lights, would not be the thing to do if you knew all the facts. For example, suppose that you have to choose between three envelopes. You know that A contains £70. You also know that one of B and C contains £100 and the other contains nothing but you do not know which is which. Here it seems highly plausible to say that you should take A even though you know that taking A would definitely not be the thing to do if you know all the facts.[15] This observation also has an epistemic parallel.[16] Suppose you are keen to know whether *p* but have no evidence either way. In such a case, it seems highly plausible that you should suspend judgment even though if you knew all the facts – including whether *p* – suspending judgment would not be the thing to do.[17]

I take these points to show that the uniformity thesis has at least has the status of default presumption – it is something which we should accept in the absence of strong reasons to deny it. Since I cannot consider here whether there are such strong reasons, I will henceforth take the thesis for granted.

Given the uniformity thesis, there is pressure on perspectivists who allow for permissible false beliefs to accept weak infallibilism. For given the uniformity thesis, whether or not you should A is determined by the same

---

[14]   Cf. Kolodny – MacFarlane (2010, 119-120) and Thomson (2008, 187-191). See Kiesewetter (2011) for a perspectivist reply to this argument.

[15]   This version of the argument is due to Ross (2012). For alternative versions see, e.g. Jackson (1991) and Kiesewetter (2011).

[16]   I believe I have seen this point made elsewhere. Unfortunately, I have not been able to find it again.

[17]   Indeed, if the objectivist holds that it is always permissible to believe the truth (Shah 2003; Wedgwood 2002; 2003; Whiting 2010), then suspending judgment is never the thing to do. For this sort of objection to objectivism about belief, see Feldman (1988a, 245).

set of considerations as determines whether it is permissible for you to believe that you should A – both are determined by those considerations which fall within your perspective. This point distinguishes questions about what you should do from most other questions. Ordinarily, the considerations which determine whether it is permissible to believe that *p* need not determine whether *p*. Since the question of whether you should A is an exception to this general rule it would not be surprising if, in this special case, truth and permissible belief did coincide.

This point does not guarantee weak infallibilism. Even though the considerations bearing on the questions of whether you should A and on whether it is permissible to believe you should A must both fall within your perspective, different considerations could still be relevant to each question. Perspectivism does not as such rule out the possibility of reasons to believe that you should A which are not reasons to A, or reasons not to A which are not reasons not to believe that you should A.[18] Reasons of this sort could make it permissible to believe that you should A when it is not the case that you should A. So it is certainly possible to accept perspectivism while denying weak infallibilism.

Nonetheless, while this possibility remains open, I would be surprised if many perspectivists wanted to endorse it. If weak infallibilism is false, then there are cases in which it is not the case that you should A but in which it is permissible – there is sufficient reason – to believe that you should A. On the plausible assumption that if there is sufficient reason to believe that you should A, then you are not in a position to know that it is not the case that you should A, this means that there are cases in which we are hidden from normative truths by our permissible normative beliefs. I suggest that this runs counter to the spirit of perspectivism.

The central motivation for perspectivism is the vague but intuitive thought that the normative must be able to *guide* us. That is why perspectivists reject objectivism – they hold that considerations which fall outside our epistemic position cannot perform this function. If the normative is to guide us, they insist, there must be a special relationship between our epistemic position and facts about what we should do. However, if the considerations which determine whether we should A are to guide us then we do not only need access to those considerations. We also need access to their

---

[18] Putative examples of this sort are discussed in the literature on Kearns – Star's (2009) view of 'reasons as evidence'.

normative significance. This is not to say that we cannot make all sorts of mistakes about what we should do. Nor, perhaps, is it to rule out the possibility of unknowable facts about what we should do.[19] But it does seem to rule out the possibility of certain sorts of mistakes about what we should do. In particular, it seems that you should not be prevented from knowing whether you should A by the very strength of your epistemic position with respect to that question. When things are going well, as when your epistemic position gives you strong enough reason to permit the belief that you should A, your normative beliefs should not lead you astray.

While not conclusive, these considerations seem to me to put pressure on those perspectivists who allow for permissible false beliefs to nonetheless endorse weak infallibilism.[20] And as we have seen objectivists and perspectivists who reject the possibility of permissible false beliefs have a far more straightforward route to weak infallibilism. So if what I have argued here is right, both objectivists and perspectivists should accept that akrasia always involves some local failing – a belief you should not have or a failure to intend something which you should. Disjunctivism about akrasia thus looks like a viable view.

---

[19] One reason perspectivists might want to allow for this turns on application of Williamson's (2000) 'anti-luminosity' argument to facts about what we should do.

[20] We might be tempted to think that the perspectivist can offer a more conclusive case for weak infallibilism as follows:

(1) What you should do is determined by considerations which both fall within your perspective and bear on what you should do.

(2) A consideration p falls within your perspective if you permissibly believe that p.

(3) So, when you permissibly believe that you should A, the consideration that you should A falls within your perspective.

(4) The consideration that you should A bears conclusively on what you should do.

(5) So, when you permissibly believe that you should A, you should A.

But while tempting, this argument feels like a cheat. If you should A that must be in virtue of *other* considerations which give you conclusive reason to A. To guarantee weak infallibilism, it needs to be shown that when you permissibly believe that you should A, there must be such other considerations.

2. Disjunctivism about Means-End Coherence and Intention Consistency: The Problem of Mere Permissibility

The means-end incoherent agent fails to intend what he takes to be a necessary means to an end he intends. The inconsistent agent intends to do things which he believes he cannot do together. Can the disjunctivist show that such agents must either have an attitude they should not have or lack an attitude they should have?

In considering this question I shall, until section 3.2, assume objectivism, for simplicity. I shall also restrict myself to cases of means-end incoherence and intention inconsistency in which the beliefs involved are held permissibly (and so, given the first assumption, true). And I shall continue to assume that you should intend to A if you should A, and now also that you should not intend to A if you should not A. Given these assumptions, the disjunctivist can show that there is something wrong with means-end incoherence and intention inconsistency by establishing that the following claims hold:

(1) If you permissibly intend to E and M-ing is necessary for E-ing, then you should M.
(2) If you permissibly intend to A and you cannot both A and B, then you should not B.[21]

The standard way to argue for (1) is to appeal to the familiar idea that reasons for action transmit from ends to means. More precisely, suppose we assume:

(Transmission): If you have reason to E and M-ing is necessary for E-ing, then you have reason to M which is at least as weighty as your reason to E.

It plausibly follows from Transmission that if you have most reason to intend an end, then you have most reason to take the necessary means to

---

[21]   It might seem implausible that what is wrong with some cases of means-end incoherence and intention inconsistency is merely that, e.g. the agent has a false belief. Again though, my focus in this paper is on whether the disjunctivist can show that *something* is wrong with the problematic combinations. It is a further question whether disjunctivists give a plausible explanation of what is wrong with these combinations. Cf. n.10.

that end (cf. Raz 2005; Schroeder 2009; Skorupski 2010). It also plausibly follows that if you have most reason to intend an end, then you should not do anything which is incompatible with achieving that end.

Transmission thus supports (1) and (2) in a fair range of cases. However, as has often been noted, this does not establish that (1) and (2) hold in all cases. The problem is that there are many cases in which you permissibly intend to A but do not have most reason to A. In cases of *mere permissibility*, you have sufficient but not conclusive reason for multiple incompatible options. To take a standard example, Buridan's ass, stuck between two equally attractive bales of hay, has sufficient reason to take the right bale of hay, but also sufficient reason to take the left bale of hay. In cases of this sort, it does not follow from Transmission that if you permissibly intend an end, you should take the means. Nor does it follow that you should refrain from doing anything incompatible with achieving that end. At most, it follows that you have sufficient reason to do these things.[22]

As I say, this problem has often been noted; in the next sections, I shall consider whether it can be solved. For now I want to emphasise two points. First, the problem here is not just one for the Transmission-based explanation of (1) and (2). The more general problem is that cases of this sort look like counter-examples to the basic disjunctivist idea that means-end incoherence and intention inconsistency always involve a more local failing. Suppose that Buridan's ass has inconsistent intentions – he intends to take the left bale and also intends to take the right bale. Something is wrong with this combination. But taken individually, both attitudes look perfectly fine – it is okay to intend to take the right bale of hay and it is okay to intend to take the left bale of hay. The problem only arises when you put these attitudes together.

Second, cases of mere permissibility are ubiquitous.[23] To start with, cases in which, like Buridan's ass, we have *equally weighty* reasons in favour of multiple options are not unusual. As Michael Bratman (1987, 11) re-

---

[22] For this point or the more general problem described below, see, e.g. Kolodny (2007; 2008); Ross (2012); Schroeder (2009); Wedgwood (2011); Way (2012). Bratman (1987) deserves credit for emphasizing the 'importance of Buridan' to the topic of practical reason.

[23] As Joseph Raz (1999, 100) famously put it, 'most of the time people have a variety of options such that it would accord with reason for them to choose any one of them and it would not be against reason to avoid any of them'. Raz calls this the 'basic belief'.

minds us, we face cases of this sort every time we pick from a shelf of cereal packets in the store. Importantly though, cases of equally weighty reasons are far from the only cases of mere permissibility. There are also cases of *incommensurability*, in which the reasons in favour of two or more of our options do not outweigh each other but are not equally weighty. There may be incommensurable reasons in this sense for many people to go to law school or graduate school in philosophy, to visit Salisbury Cathedral or Stonehenge, to listen to the Beach Boys or the Beatles, or for Sartre's famous student to stay home with his mother or to fight for the resistance.[24] There are also cases of mere permissibility in which one of our options is *supererogatory* – morally admirable but not required. For example, it might be supererogatory in this sense to sacrifice next summer's holiday in order to make a large donation to charity.

In all of these cases, there is clearly something wrong with means-end incoherence and intention inconsistency. For instance, something would be wrong if you intended to go on holiday as normal but did not intend to book a ticket, or if Sartre's student intended both to stay home with his mother and also to fight for the resistance. But taken individually, the attitudes involved in these combinations are perfectly acceptable. So there seem to be a wide range of cases in which (1) and (2) are false.

3. Two Disjunctivist Replies to the Problem of Mere Permissibility

However, we should not be too quick to reject (1) and (2). The cases above demonstrate, I think, that *prior* to intending a merely permissible end, you might lack most reason to take the necessary means, or to refrain from pursuing incompatible alternatives. But that is not enough to show that cases of means-end incoherence and intention consistency need involve no local failing. After all, means-end incoherent and intention inconsistent agents do not merely face a choice between merely permissible op-

---

[24]   It is controversial whether such cases are well characterised as cases of incommensurability, as opposed to, e.g. incomparability, parity, or rough equality. (See, e.g. the introduction to Chang 1997, and the essays therein.) For my purposes, this dispute does not matter; my use of the term 'incommensurability' is entirely stipulative. The crucial point is that in cases of this sort the central premise of the so-called "small improvements" argument applies (Chang 1997, 23-27): a small increase in the weight of the reasons in favour of one of the options would not make that the option you should pursue.

tions. In addition, the means-end incoherent agent has chosen to pursue one of those options – and the inconsistent agent has chosen to pursue both! What the disjunctivist must argue is that it is this which makes the difference. The disjunctivist must claim that cases of mere permissibility are only possible prior to intending an end – that once you intend a merely permissible end, the balance of reasons shifts so that now you should take the necessary means, and refrain from incompatible alternatives.[25]

I know of two ways to argue for this conclusion. In what follows I shall consider them in turn and argue that neither succeed in getting around the problem. Even if we grant that intending an end can affect your reasons in the ways suggested, there will still be cases of mere permissibility in which (1) and (2) are false. For brevity, and because it is the more promising case for the disjunctivist,[26] I shall focus on the means-end case.

---

[25]   Kolodny (2008a, 453) also notes that the disjunctivist must argue that it is the intention for the end which makes the difference. However, Kolodny's suggestion in that paper and in his (2007, 252) is not that intending a merely permissible end ensures that you have most reason to take the means. Rather, Kolodny argues that being means-end incoherent makes it the case that you should drop the end. Kolodny's argument for this surprising conclusion has two premises: (i) that if intending to E makes it no more likely that you will E then you should not intend to E and (ii) that intending to E without also intending (what you have sufficient reason to believe are) the necessary means makes it no more likely that you will E. Although I cannot discuss this argument here, I will make two comments. First, although the argument is framed as a response to permissive cases, it actually applies more generally. The argument makes no appeal to there being merely sufficient reason to E. So if it succeeds, the argument shows that, so long as they have sufficient reason for their means-end beliefs, the means-end incoherent should give up their ends. Second, this suggests that the argument shows too much. For surely *sometimes* the solution to means-end incoherence is to decide to take the means. To my mind then, this argument faces an objection not dissimilar to the objection which has traditionally motivated wide-scoping. Just as it is implausible to think that merely intending to E is enough to make it the case that you should intend to M, it is implausible that merely being means-end incoherent is enough to make it the case that you should drop the intention to E. Kolodny's suggestion thus seems to allow a sort of "reverse bootstrapping". Unfortunately, I cannot further explore this matter here.

[26]   The strategies discussed below could be used to argue that if you intend to A and cannot both A and B, then you should not *form* the intention to B. But they could not be used to argue that the agent who does intend to A and also intends to B should *drop* the intention to B. The intentions to A and to B are symmetrical – any difference which the former makes to your reasons will be matched by the latter. So intention inconsistency is an even harder case for the disjunctivist.

3.1 First Strategy: Intentions Provide Reasons for the End

It is sometimes claimed that intending an end gives you an extra reason to pursue that end. If that is so, then even if two of your options are permissible prior to your choosing between them, forming the intention to pursue one of them might change this. Once you form an intention, you now have an extra reason to pursue that end, and so more reason to pursue that end than the alternative. Given Transmission, your reasons to take the necessary means to the intended end will now be stronger than your reasons to take the necessary means to the alternative. Thus you should take the necessary means to the intended end, and will be going wrong if you do not intend to do so.[27]

One question this suggestion raises is *why* intentions provide reasons. One possible answer that is sometimes noted is that there might be some value in resoluteness, which we might understand (no doubt oversimplifying) to be a matter of doing what you intend to do. Another possible answer is that we should think of an intention for an end as *lowering the cost* of pursuing that end.[28] Intending an end, the suggestion goes, is the first step to achieving that end. So once you intend a merely permissible end, you have taken a step towards achieving that end which you have not taken towards achieving the alternative. There is thus less reason against, and so stronger reason for, pursuing the end that you intend. On this view, intending to take the right bale of hay gives you an extra reason to take that bale of hay in just the way that taking one step towards the right bale would.

One feature these ideas share is that the reasons which intentions provide are quite weak. In a way this is a virtue, since it would not be plausible to suggest that intending an end gives you a very strong reason to pursue that end. (This is just the "bootstrapping" objection with which we began). Unfortunately, this feature also prevents the idea from doing the work that the disjunctivist needs. If intending an end provides you with a weak reason to pursue that end, then this reason is presumably capable of breaking ties. So if prior to making up your mind the left and right bale are equally attractive, an intention-provided reason to take the left bale ensures that you now have most reason to take the left. But as I have emphasised, cases of

---

27   This account of cases of mere permissibility is offered by Schroeder (2009).

28   For defence of this idea, see Kolodny (2011, section 4).

this sort are far from the only cases of mere permissibility. In the other cases, the suggestion does far less well. If two options are incommensurable, then a slight increase in the reasons in favour of one of them does not ensure that the reasons now conclusively favour that one. (This is one of the marks which distinguishes a case of incommensurability from a case of equality; cf. n. 24). Thus even if intending to go to law school, rather than graduate school, does give you a slight extra reason to go to law school, this does not ensure that you now have most reason to go to law school – just as a slight increase in your expected salary after law school would not.

The point is even clearer in the case of supererogation. Even if intending to take a holiday gives you a slight extra reason to go on holiday, rather than donating the money to charity, this reason need not be enough to shift the balance of reasons decisively in favour of going on holiday. After all, even if the holiday had been slightly cheaper or slightly more attractive in the first place, it would still have been okay to give the money to charity instead.

I take these examples to show that on plausible versions of the idea that intending an end gives you an extra reason to pursue that end, there will remain cases of mere permissibility which the disjunctivist cannot accommodate.

3.2. Second Strategy: Intentions Strengthen Reasons for the Means

John Brunero (2007) and Niko Kolodny (2011) defend a different explanation of how intending a merely permissible end might give you most reason to take the necessary means to that end. On Brunero and Kolodny's view, intending an end does not give you an extra reason to pursue that end. Rather, intending an end increases the strength of your reasons to take the means to that end. It is not clear whether Brunero and Kolodny take this account to show that intending a merely permissible end will *always* make it the case that you have most reason to take the necessary means to that end, as the disjunctivist requires. Nonetheless, it is worth considering whether it does so.[29]

---

[29]   Neither Brunero nor Kolodny present their account as a defence of disjunctivism about means-end coherence or intention consistency. Brunero accepts the wide-scope view. Kolodny does claim that the ideas he draws on 'explain many of the phenomena taken to be evidence for a rational requirement of consistency in intention' (Kolodny 2011, n. 38). However, the account of intention consistency he defends in his (2008b)

The Brunero/Kolodny account combines three ideas. First, intending an end makes a difference to what the future is likely to hold. Once I intend to go to Boston in the spring, I am more likely to do so. Importantly, I am also more likely to take the partial means to this end – for instance, to buy a ticket, book time off work, travel to the airport, board the plane, and so forth. Second, the likelihood that taking some partial means to an end will help you to achieve that end depends on how likely it is that you will take the other means which together will achieve that end. Thus how likely it is that buying a ticket will help me get to Boston depends on how likely it is that I will also book the time off work, travel to the airport, and so on. Third, the strength of your reasons to take some partial means to an end depends (among other things) on how likely taking that means is to help you achieve that end. (Thus Brunero and Kolodny assume a kind of perspectivism about what we have most reason to do.)

These three ideas support the claim that we often have stronger reasons to take the means to permissible intended ends than to permissible unintended ends. Given the first and second ideas, it is more likely that taking a partial means to an intended end will help you to achieve that end than that taking a partial means to an unintended end will help you to achieve that end. And so given the third idea, we have stronger reasons to take such means. To illustrate, suppose that you intend, with sufficient reason, to go to graduate school in philosophy, although you also have sufficient reason to go to law school. Now that you have this intention, studying for the GRE is more likely to be an effective means to going to graduate school than studying for the LSAT is to be an effective means to going to law school. For since you do not intend to go to law school, you are unlikely to do any of the other things you need to do to get into law school, and so studying for the LSAT would be a waste of time. And since you do intend to go to graduate school, it is quite likely that you will do the other things you need to do to get into graduate school, and so studying for the GRE will not be a waste of time. You therefore have better reason to study for the GRE than to study for the LSAT.

---

appeals to reasons against combinations of intentions of the sort described in section 4 below. He does not explicitly address means-end coherence in the work I am drawing on here. As I note in n.25 the account of means-end coherence he offers elsewhere is very different.

I think that the phenomenon Brunero and Kolodny point to is genuine and important, and that their account of it is promising.[30] Nonetheless, I doubt that the account supports the claim that we always have most reason to take the necessary means to merely permissible intended ends. The problem is that (as both Brunero and Kolodny note) whether or not you intend an end makes no difference to the likelihood that taking a *sufficient* means to that end will be effective – sufficient means are, after all, sufficient. So the three ideas above do not support the claim that you should take the necessary means to an intended end when an alternative is to take a sufficient means to an unintended end.

Consider again the choice between taking a holiday and making a large donation to charity. Suppose that a necessary means to taking the holiday is booking a ticket and that a sufficient means to making the donation is pressing a button which will instantly and irrevocably send the money from your bank account to the charity. The disjunctivist needs to show that if you intend to take the holiday, then you have most reason to book the ticket, and in particular, more reason to book the ticket than to press the button. Clearly though, we cannot defend this claim by arguing that booking the ticket is a more effective means to going on holiday than pressing the button is to making the donation. The latter is as effective as a means can be.

Nonetheless, means-end incoherence would be clearly problematic in this case. Something would be wrong if you intended to take the holiday but did not intend to book the ticket. The Brunero/Kolodny account of

---

[30] Although it is worth noting a worry that Kolodny mentions and attributes to Jay Wallace (2011, 58). We might suppose that what explains why intending an end makes you more likely to take means to that end is that, as a rational agent, you will be disposed to do what you take yourself to have most reason to do. But if that is so, the Brunero/Kolodny account presupposes that there is more reason to take means to intended permissible ends than to unintended permissible ends. It cannot explain why this is so.

Kolodny has a reply to this. He accepts that intending an end makes it the case that you have slightly more reason to pursue that end, on the grounds that intended ends are slightly "cheaper" than unintended ends. This is enough, he suggests, to dispose a rational agent to take the means to intended ends, rather than unintended ends.

However, if what I said in the previous section is right, this reply should not convince. Even if intending an end does increase the strength of the reasons to pursue that end it will often not do so enough to give you most reason to pursue that end.

how intending an end affects your reasons to take the means to that end does not help the disjunctivist to explain why this is so.[31]

4. An Upshot

I have argued that there is a plausible case that akrasia guarantees a local failing but that means-end incoherence and intention inconsistency do not. If this is right, then there is an important motivation for accepting wide-scope requirements against means-end incoherence and intention inconsistency which does not apply to the requirement against akrasia. In turn, this counts against the common assumption that we should expect a uniform treatment of akrasia, means-end incoherence, and intention inconsistency. In this final section, I want to suggest one reason why this is a significant result.

Consider what I have elsewhere called the *intermediate-scope* view of means-end coherence and intention consistency. Like the wide-scope view, this view accepts that means-end incoherence and intention inconsistency need not involve any specific attitude you should not have. But this view rejects the wide-scoper's claim that what is wrong with means-end incoherence and intention inconsistency is explained by wide-scope requirements against these combinations. Instead, the view holds that there are requirements against the combinations of *intentions* involved in these combinations. The view might accept:

(Means-End Intermediate): If M-ing is necessary for E-ing, then you should not [intend to E and not intend to M].

(Intention Intermediate): If you cannot both A and B, then you should not [intend to A and intend to B].

These requirements imply that there is something wrong with cases of means-end incoherence and intention inconsistency which involve true be-

---

[31] The problem here is clearest in the case of sufficient means but may not be limited to them. There may be cases in which you have to choose between a merely permissible risky end and a merely permissible sure-thing. In such cases, the means to the sure-thing may be more likely to succeed than the means to the risky end. So the explanation of why you should take the means to the risky end, if that is what you intend, cannot be that these means are more likely to be effective than the means to the sure thing.

liefs. They do not by themselves imply that there is something wrong with cases of means-end incoherence and intention inconsistency which involve false beliefs. But the intermediate-scope view can be extended to cover such cases. One way to do this is to incorporate an element of the disjunctivist strategy. For example, if we accept objectivism, we could argue that means-end incoherence always involves *either* a belief you should not have *or* a failure to combine your intentions as you should. Another way to extend the view is to distinguish between reasons and rationality and argue that what you are rationally required to do is what, relative to your (perhaps rational) beliefs, you have most reason to do. Given this claim, the intermediate-scope requirements plausibly support the claims that if you (rationally) believe that M-ing is necessary for E-ing, then you are rationally required not to [intend to E and not intend to M], and that if you (rationally) believe that you cannot both A and B, then you are rationally required not to intend both to A and to B. When developed in either of these ways, the intermediate-scope view implies that there is something wrong with all cases of means-end incoherence and intention inconsistency.[32]

The difference between the wide- and intermediate-scope views may seem slight. But as I have argued elsewhere, it turns out to be significant. We can illustrate this point by noting two important challenges that the wide-scope view faces. The first challenge is that it is not clear what grounds wide-scope requirements. For example, it is not clear what makes it the case that you should be means-end coherent. Part of the difficulty here is that standard wide-scope requirements have *unlimited application* – they apply to all agents, in all circumstances. This means that we are severely limited in the resources which we can appeal to, to explain why they hold. In particular, we cannot appeal to idiosyncratic features of particular agents, or particular circumstances, to explain why they apply to those agents in those circumstances (cf. Schroeder 2004, 349 and n.20; Way 2012, 492-493). The second challenge is that it is not clear why we should comply with wide-scope requirements, or what reason we have to do so (Kolodny 2005; Broome 2005). (These challenges are connected, of course, because one way to ground such requirements – I suspect the most promis-

---

[32] The second of these views is defended in Way (2010; 2012). A perspectivist version of the first is defended, with respect to intention consistency, in Kolodny (2008b). Wedgwood (2003; 2011) defends a view of means-end coherence which draws on both strategies.

ing way – is to appeal to reasons to comply with them.) One reason that this challenge is hard is that wide-scope requirements have a peculiar form, so that familiar models of reasons for intention and belief cannot be applied to them. Ordinarily, reasons bearing on intention are reasons bearing on the action intended, and reasons bearing on belief are evidence bearing on the proposition believed. But neither of these familiar kinds of reasons bears on the combinations of belief and intention ruled out by the wide-scope requirements against means-end incoherence and intention inconsistency. It is thus hard to see what a reason to comply with these requirements might look like.

The intermediate-scope view promises to fare better with both of these challenges.[33] First, intermediate-scope requirements do not have unlimited application. They do not apply to everyone, in all circumstances. They apply only to agents in certain circumstances – agents for whom a means is necessary for an end or who face incompatible options. We should thus expect it to be easier to explain intermediate-scope requirements than to explain wide-scope requirements. Second, the intermediate-scope requirements make available simple and natural answers to the question of what reason there is to comply with them. When M-ing is necessary for E-ing, the reason not to [intend to E and not intend to M] is simply that M-ing is necessary for E-ing. And when you cannot both A and B, the reason not to [intend to A and intend to B] is that you cannot do both. These answers parallel the natural answer to the question, 'what reason is there not to intend to A?', when you cannot A – namely, that you cannot A. Third, reasons against combinations of intentions are familiar in other contexts. For example, I might have reason against both intending to take drug 1 and intending to take drug 2 because, although taking either would cure me, taking both would kill me. So on this view, the reasons involved in means-end coherence and intention consistency do not look to be of a radically different kind to the reasons bearing on intentions in other contexts.

Despite this promise, a significant worry about the intermediate-scope view is that it appears not to extend to the requirement against akrasia. What is wrong with akrasia must be either (i) a problem with the belief that you should A, (ii) a problem with the lack of an intention to A, or (iii) a problem with the akratic combination. Since akrasia only involves two at-

---

[33]   See Way (2010; 2012) for a fuller defence of these claims.

titudes, there is simply no room for an intermediate-scope requirement against akrasia.

This may seem like a problem for the intermediate-scope view. We might have thought that if wide-scoping is the right way to go about akrasia, it will be the right way to go about means-end incoherence and intention inconsistency, and that if disjunctivism is defensible about akrasia, it will be defensible about the other two combinations. Either way, the intermediate-scope view must be a mistake. However, if what I have argued in this paper is correct, we should not be so quick to assume a uniform account of the three problematic combinations. There is a strong motivation for accepting requirements against combinations of intentions in the case of means-end incoherence and intention inconsistency which simply does not apply in the case of akrasia. Ordinary reasons for and against individual intentions and beliefs do not explain what is wrong with means-end incoherence or intention inconsistency. But it may well be that such reasons do explain what is wrong with akrasia. So the intermediate-scope view may yet be the way to go.[34]

References

BRATMAN, M. (1987): *Intentions, Plans and Practical Reason*. Cambridge, MA: Harvard University Press.

BROOME, J. (1999): Normative Requirements. *Ratio* 12, 398-419.

BROOME, J. (2005): Does Rationality Give Us Reasons? *Philosophical Issues* 15, 321-337.

BRUNERO, J. (2007): Are Intentions Reasons? *Pacific Philosophical Quarterly* 88, 424-444.

CHANG, R. (ed). (1997): *Incommensurability, Incomparability, and Practical Reason*. Cambridge, MA: Harvard University Press.

FELDMAN, R. (1988a): Epistemic Obligations. *Philosophical Perspectives* 2, 235-256.

FELDMAN, R. (1988b): Objective and Subjective Justification in Ethics and Epistemology. *The Monist* 7, 405-419.

GIBBONS, J. (2009): You Gotta Do What You Gotta Do. *Noûs* 43, 157-177.

GIBBONS, J. (2010): Things That Make Things Reasonable. *Philosophy and Phenomenological Research* 81, 335-361.

JACKSON, F. (1991): Decision-Theoretic Consequentialism and the Nearest and Dearest Objection. *Ethics* 10, 461-482.

KEARNS, S. – STAR, D. (2009): Reasons as Evidence. In: Shafer-Landau, R. (ed.): *Oxford Studies in Metaethics. Volume 4.* Oxford: Oxford University Press, 215-224.

KIESEWETTER, B. (2011): "Ought" and the Perspective of the Agent. *Journal of Ethics and Social Philosophy* 5, 1-25.

KOLODNY, N. (2005): Why Be Rational? *Mind* 114, 509-563.

KOLODNY, N. (2007): How Does Coherence Matter? *Proceedings of the Aristotelian Society* 107, 229-263.

KOLODNY, N. (2008a): Why be Disposed to be Coherent? *Ethics* 118, 437-463.

KOLODNY, N. (2008b): The Myth of Practical Consistency. *European Journal of Philosophy* 16, 366-402.

KOLODNY, N. (2009): Reply to Bridges. *Mind* 118, 369-376.

KOLODNY, N. (2011): Aims as Reasons. In: Freeman, S. – Kumar, R. – Wallace, R. J. (eds.): *Reasons and Recognition: Essays on the Philosophy of T. M. Scanlon*. New York: Oxford University Press, 43-78.

KOLODNY, N. (forthcoming). Instrumental Reasons. In: Star, D (ed.): *Oxford Handbook of Reasons and Normativity*. Oxford: Oxford University Press.

KOLODNY, N. – MACFARLANE, J. (2010): Ifs and Oughts. *Journal of Philosophy* 107, 115-143.

LITTLEJOHN, C. (2010): Moore's Paradox and Epistemic Norms. *Australasian Journal of Philosophy* 88, 79-100.

LORD, E. (2013): *The Importance of Being Rational*. PhD. Dissertation. Princeton University.

PARFIT, D. (2011): *On What Matters*. Oxford: Oxford University Press.

RAZ, J. (1999): *Engaging Reason*. Oxford: Oxford University Press.

RAZ, J. (2005): The Myth of Instrumental Rationality. *Journal of Ethics and Social Philosophy* 1. www.jesp.org.

ROSS, J. (2012): Normativity, Rationality, and Commitment. In: Shafer-Landau, R. (ed.): *Oxford Studies in Metaethics, Volume 7*. Oxford: Oxford University Press, 138-181.

SCHROEDER, M. (2004): The Scope of Instrumental Reason. *Philosophical Perspectives* 18, 337-364.

SCHROEDER, M. (2009): Means-End Coherence, Stringency, and Subjective Reasons. *Philosophical Studies* 143, 223-248.

SHAH, N. (2003): How Truth Governs Belief. *The Philosophical Review* 112, 447-482.

SKORUPSKI, J. (2010): *The Domain of Reasons*. Oxford University Press.

SMITHIES, D. (2012): The Normative Role of Knowledge. *Noûs* 46, 265-288.

THOMSON, J.J. (2008): *Normativity*. Chicago – La Salle: Open Court.

WAY, J. (2010): Defending the Wide-Scope Approach to Instrumental Reason. *Philosophical Studies* 147, 213-233.

WAY, J. (2012): Explaining the Instrumental Principle. *Australasian Journal of Philosophy* 90, 487-506.

WAY, J. (forthcoming): Reasons and Rationality. In: Star, D. (ed.): *Oxford Handbook of Reasons and Normativity*. Oxford: Oxford University Press.

WEDGWOOD, R. (2002): The Aim of Belief. *Philosophical Perspectives* 16, 267-297.

WEDGWOOD, R. (2003): Choosing Rationally and Choosing Correctly. In: Stroud, S. – Tappolet, C. (eds.): *Weakness of Will and Varieties of Practical Irrationality*. Oxford: Oxford University Press, 201-229.

WEDGWOOD, R. (2011): Instrumental Rationality. In: Shafer-Landau, R. (ed.): *Oxford Studies in Metaethics. Volume 6*. Oxford: Oxford University Press, 280-309.

WHITING, D. (2010): Should I Believe the Truth? *Dialectica* 61, 213-224.

WHITING, D. (forthcoming): Nothing but the Truth: On the Aim and Norms of Belief. In: Chan, T. (ed.): *The Aim of Belief*. Oxford: Oxford University Press.

WILLIAMSON, T. (2000): *Knowledge and its Limits*. Oxford: Oxford University Press.

WILLIAMSON, T. (2005): Knowledge, Context and Agent's Point of View. In: Preyer, G. – Peter, G. (eds.): *Contextualism in Philosophy*. Oxford: Oxford University Press.

The Bootstrapping Objection

CHRISTIAN PILLER

Department of Philosophy. University of York
Heslington. York YO10 5DD. United Kingdom
christian.piller@york.ac.uk

ABSTRACT: If our mental attitudes were reasons, we could bootstrap anything into rationality simply by acquiring these mental attitudes. This, it has been argued, shows that mental attitudes cannot be reasons. In this paper, I focus on John Broome's development of the bootstrapping objection. I distinguish various versions of this objection and I argue that the bootstrapping objection to mind-based accounts of reasons fails in all its versions.

KEYWORDS: Bratman – Broome – intention – reasons – reasoning.

Before I explain the setting in which the discussion to follow will take place, I would like to start with a simple story. Its morale is something, I hope, we all agree upon. Its relevance will become apparent later on.

Suppose I believe that I am a better than average philosopher. I admit that I hardly get any invitations to conferences and it is also true that when I submit an abstract I often get a message telling me that, unfortunately, many high-quality submissions had to be rejected. I have some publications. They are, by no means, in top-journals; and even if it's only the *North Yorkshire Philosophical Gazette* I had to revise my article several times – no idea why. And so the evidence is mounting. I don't hold a conspiracy theory according to which my failures are unproblematically compatible or would even support my high opinion of myself. I do regard these things as evidence against my positive view of my abilities but I manage to put them

aside. I do believe that I am better than average – actually, when considering the work of some of my colleagues, quite a bit better. This belief of mine – I am better than average – gives rise to other beliefs, for example to the belief that my luck will change.

If you assess the normative status of this latter belief – my luck will change – it won't help that I can provide the following argument for it: 'Better than average philosophers will, by the end of their career, have a better than average publication record. In order to have such a record my luck will have to change. And it will, as I am a better than average philosopher.' The corner stone of this argument – my being a better than average philosopher – is, to say the least, unsupported. The evidence points in a different direction: I really should know better. We understand how this can happen. It is easier to delude yourself than to face uncomfortable facts. The positive view I have of myself, I said, is not justified. This, in turn, affects the normative status of my optimistic outlook. It is unreasonable to expect a positive change because the view on which this expectation is based is unreasonable. So this is the claim, I hope, we all agree upon: If you put irrationality in, you will not get rationality out, even with the best of arguments.

1. Background

The bootstrapping objection arises in debates about the foundations of normative thought. Let me explain. Many people are convinced that one ought not to drink poison. They think the fact that something is poisonous is a reason not to drink it and, if there are no reasons to the contrary, this reason alone can explain why one ought not to drink it. In this way, we introduce world-based reasons and world-based oughts. The fact that some liquid is poisonous is a reason not to drink it and this fact explains why, if nothing else needs to be considered, one ought not to drink it, all things considered.

You like anyone else do not know everything. What if you believe – on the basis of good reasons – that the stuff you intend to drink is not poisonous? You think it is completely harmless. It looks like wine; it smells like wine; it tastes like wine. If this is what you think, then it would be irrational for you not to drink it. In this way, we introduce mind-based rules of rationality. These rules tell us how to move in a rational way from one

mental state to another. In the case at hand, they tell you to move from your belief that the liquid is harmless and tasty to the intention and – if nothing interferes – to the action of drinking it. It is a natural thought that rationality is normative. If, all things considered, it is rational to do something, then, all things considered, you ought to do it.

Wait a moment. If the facts determine that you ought not to drink the liquid and if your mental states determine that you ought to drink it, then it would be the case that you ought to drink it and that you ought not to drink it. This cannot be right.[1]

I distinguish between three kinds of reactions to this problem. According to the first reaction, what I said cannot be right, can be right after all.

(1) The problem we are facing is one of normative inconsistency. If we allow for both world-based and mind-based reasons and oughts, they may pull us in opposite directions. Accepting that normativity cannot that easily collapse into inconsistency, we try to show that the inconsistency is only apparent. It would be only apparent if the opposing recommendations – one ought to drink the liquid and one ought not to drink it, relied on different sense of 'ought'. Ewing (1947) thought that in one (subjective) sense of 'ought' you ought to drink the liquid, yet in another (objective) sense of 'ought' you ought not to drink it. This move, in my view, would deny one of the presuppositions of practical thinking. The question that characterizes practical deliberation is 'What should I do?' It is not about what I should do in this sense or in that sense. Such qualified questions – should I, just

---

[1]   The distinction between facts and attitudes and the corresponding distinction between world-based and mind-based reasons is more complex than I made it out to be. This has to do with the fact that some facts (or parts of the world) consist in people having certain attitudes. This complicates the application of the distinction. Suppose I believe that everyone here hates me. Should I leave town or should I seek psychological help? Both answers find their place (though on different grounds) in a world-based as well as in a mind-based normative framework. Let's start with the world-based framework. The fact that everyone around here hates me is a good reason to leave town. The fact that I believe so – taking my believing in this case as a fact about the world – is a world-based reason to seek help. Within a mind-based framework, my believing that everyone hates me is a reason to leave town, whereas my belief that I have this belief (plus other beliefs about paranoia) are reasons to seek help. I will put this complication aside as it won't play a role in my discussion. Others have tried to draw the same distinction in similar ways. Prichard (1932, 18) describes the issue as follows: 'If a man has an obligation, i.e., a duty, to do some action, does the obligation depend on certain characteristics of the situation, or on certain characteristics of his thought about the situation?'

considering this or that aspect, do it? – can be steps towards answering what seems to be the real question of practical deliberation, which uses 'should' unambiguously.

The two remaining reactions agree on the ideal of normative consistency as well as on the threat posed to this ideal if we accepted both world-based and mind-based reasons. They differ in their views about which of these two accounts of normativity ought to be abandoned. (2a) According to Kolodny (2005), there is no mind-based normativity; all normativity is world-based. Means-end reasoning, consistency and other apparent mind-based principles of rationality are just that – they appear to tell us what we ought to do. Their appearance is only veridical if they are underwritten by world-based reasons. But if they are, such world-based reasons do all the work. (2b) Like Kolodny, John Broome accepts world-based normativity, e.g. a liquid's being poisonous is a reason not to drink it. Traditional principles of rationality, for Broome, are neither world-based nor mind-based. They are wide-scope requirements. Thus, Broome and Kolodny differ as Broome accepts a conception of rationality that goes beyond the domain of world-based reasons. Broome's wide-scope move is supposed to alleviate the worry about normative inconsistency. However, having allowed world-based reasons, Broome thinks that the normativity of wide-scope requirements of rationality would need to be anchored in such reasons. As this is no easy task, Broome remains agnostic about the normativity of principles of rationality.

There is a third option, namely (3) to think of all normativity as mind-based. I find this third option independently attractive. However, here I will not pursue the project of discussing its merits. I will be engaged in more limited project. The Bootstrapping Objection is an argument against the idea of mind-based reasons. The basic idea is the following. Reasons, if mind-based, would be too easy to come by. If thinking that a liquid tastes nice was a reason to drink it, and the liquid's taste was the only relevant consideration, all one needed was this thought in order to render drinking the liquid rational. This is taken to be implausible. The bootstrapping objection has been applied in the theoretical as well as in the practical case. Neither beliefs nor intentions are reasons. In what follows I will discuss John Broome's development of this objection.[2]

---

[2]   'Bootstrapping', in the sense in which it is relevant here, has been introduced into the philosophical debate by Michael Bratman. In Bratman (1981), he regards bootstrap-

2. Bootstrapping 1: Against attitudinal reasons. Why beliefs are not reasons. You cannot, by means of your beliefs, bootstrap a new reason into existence, to add to your evidence.

John Broome writes:

First, there are no attitudinal reasons. Attitudes are not reasons in the way I have described. Here is why. Take R3: If you believe p and you believe if p then q, your two beliefs are together a reason for you to be-lieve q. For 'p' substitute 'Carbon dioxide is poisonous' and for 'q' 'Emissions of carbon dioxide are harmful'. There are various pieces of evidence for the proposition q, and others against it. Each of these piec-es of evidence constitutes a *pro tanto* reason either for or against believ-ing q. Just for the sake of argument, let us assume that the evidence falls short of being conclusive, so, by a small margin, these evidential reasons do not require you to believe emissions of carbon dioxide are harmful. Now suppose there is no evidence for the proposition p, that carbon dioxide is poisonous, but nevertheless you believe it. Suppose you also believe that if p then q – that, if carbon dioxide is poisonous, emissions of it are harmful. According to R3, these beliefs of yours con-stitute a reason to believe q. Since the evidential reasons fall short of re-quiring you to believe q, we may assume this attitudinal reason tips the balance. Therefore all your reasons together require you to believe emissions of carbon dioxide are harmful. That is not credible. Your evi-dential reasons do not require you to believe emissions of carbon dio-xide are harmful, and it is not credible that your beliefs could add to these reasons. You cannot, by means of your beliefs, bootstrap a new reason into existence, to add to the evidence. (Broome 2009, 91)

---

ping as a problem for the view that intentions are reasons. For him it is a problem that needs to be contained – not a problem that would refute the idea that intentions are reasons. In epistemology we find discussions of a related problem under 'bootstrapping' or 'the easy knowledge problem' (see Weisberg 2012 for a general overview of the de-bate). Note that in epistemology the issue concerns the problematic character of self-supporting reasoning – your given and not further justified starting point is validated by reasoning that presupposes the legitimacy of the starting point – whereas we are dealing here with an issue regarding the nature of reasons. As Broome (2013) has not been pub-lished at the time of writing this article, I will occasionally refer to the manuscript Broome (2009). None of the quoted passages has been changed. Much of the relevant material can also be found in Broome (2001) and in Broome (2007).

I do not find this example transparent. I am asked to imagine that I believe that carbon dioxide is poisonous. Although I am convinced that emitting poisonous gases is harmful, I see myself collecting evidence for and against the harmfulness of carbon dioxide and I find the evidence inconclusive. Do I have to imagine myself as being irrational? I will come back to Broome's example later on. For the moment I will use my own example which, I hope, avoids the problem I have with Broome's example.

Suppose my evidence is such that I am not required to believe that team A will win the next game by a big margin. There is evidence for it and against it. Suppose I believe that their next opponent, team B, is lacking in confidence. I also believe that against opponents who lack in confidence, team A will win with a big margin. Do I, as I believe that their next opponent lacks in confidence, have an additional reason to believe that team A will win with a big margin?

This example seems to fit Broome's argumentative purposes. He would point out the following. 'Your evidential reasons, you said, do not require you to believe that team A will win with a big margin and it is not credible that your beliefs, for example your belief that team B lacks in confidence, could add to these reasons. You cannot, by means of your beliefs, bootstrap a new reason into existence, to add to the evidence.'

The crucial question is the following: What do we mean by evidence?

(a) Evidence for S is what S knows to be true.

Suppose that in my example, though I think that team B has low confidence, I do not know this. Given that only knowledge is evidence, I cannot add to the evidence simply by adding this belief.

(b) Evidence for S is all the facts S is aware of.

Suppose it is a fact that team B has low confidence and that I believe s. Then this belief of mine is part of my evidence. On this conception of evidence, I can add a new reason to the evidence simply by believing it, as long as it is true.

(c) Evidence for S is all S's beliefs which are such that the objective probability of the conclusion given what S believes is higher than otherwise.

The fact that B has low confidence raises the objective probability of a big win by team A. Again, you add to the evidence simply by believing something (as long as what you believe does indeed raise the conclusion's probability).

(d) Evidence for S is all those of S's beliefs which raise the value of the agent's probability function for the conclusion.

Again, under this conception of evidence, simply believing that B has low confidence will add to your reason.

Only one of these conceptions of evidence, namely evidence is what one knows, supports Broome's argument. However, on this conception of evidence we do not need any argument that would show us that we cannot add to our evidence by simply believing something. Simply believing, i.e. believing without knowing, cannot add to the evidence if we assume, via our understanding of evidence, that simple believing is not part of the evidence and does not add to it. If only known facts are, by assumption, evidence, then beliefs which do not amount to knowledge do not add to our evidence. There is no need here to refer to 'implausible bootstrapping'. Broome's objection is accurate but, I think, question-begging.

If we allow for beliefs as reasons, then by coming to believe that team B has low confidence, I have acquired a view which, given my other beliefs, is relevant to whether team A will win with a big margin. I will feel pressure to adjust my target belief. In this sense, captured by alternatives (c) and (d), I have added to my evidence simply by believing something.

3. Bootstrapping 2: If beliefs were reasons, any belief you have gives you a reason to have it. That cannot be so; it would be absurd bootstrapping.

Broome continues:

I can reinforce the example. R3 entails that, if you believe p and you believe that if p then p, these beliefs constitute a reason for you to believe p. That cannot be so. We can take it for granted that you believe the tautology that, if p then p. Given that, R3 entails that believing a proposition gives you a reason to believe it. Any belief you have gives you a reason to have it. That cannot be so; it would be absurd bootstrapping.

I have argued that Bootstrapping 1 is question-begging. This move is not available in this case. Whatever your conception of evidence, believing p should not be a reason for itself. This is a problem for a proponent of mind-based reasons, as only for him or her will what we need reasons for and what is a reason belong to the same category; thus only for him or her does this problem arise. For Broome, reasons are explanations of normative facts. He is right that pointing out that one does believe p would on no account be (by itself) a successful explanation of why one ought to believe p.

In order to assess Broome's claim that accepting beliefs as reasons commits one to accepting self-support, I look at Bayesianism for guidance. How does the Bayesian understand the notion of a reason? The starting point of a Bayesian analysis is the following. The thing that is a reason makes what it is a reason for more likely. Evidence E is a reason for hypothesis H if the probability of H given E is higher than the absolute probability of H. If there is smoke, the probability that there is fire increases: $Prob(F/S) > Prob(F)$. For any agent whose probability function contains this inequality, smoke and fire are related. The former is a reason for the latter within the agent's epistemic system.[3]

How does this fact – smoke is a reason for thinking that there is fire – influence the agent's belief that there is a fire? It depends. If the agent has no idea that there is smoke, then the fact that he regards smoke as evidence for fire, will not have any effect on the agent's belief-system. If, however, the agent notices smoke, then, by the Rule of Strict Conditionalization, the new probability of there being a fire will increase to the point at which the old conditional probability, of there being a fire given that there is smoke, has put it. Conditional probabilities tells us what, for a particular agent, is a reason for what. This fact will only influence the agent, if evidence becomes available, for example by an agent's noticing that something that is evidence for something else has actually occurred. We can conceptualize this difference between regarding something as a reason and being influenced by it in different ways. For example, we could say that smoke *is* a rea-

---

[3] I said that the idea that evidence for H raises the conditional probability of H is the starting point of a Bayesian analysis. Suppose you believe that some event E has happened because you've been there when it happened, you've seen it happening. Is the fact that someone tells you that it happened for you a reason to believe it? It does not raise your probability in its happening, though it would raise it had you not been there. The fact that you already are in possession of stronger evidence should not mean that weaker evidence is no evidence at all.

son for this agent to believe that there is fire but that this agent only *has* a reason to believe that there is fire if he notices smoke.[4]

(a) For this agent, smoke *is* a reason to believe that there is fire. It is a reason the agent *has*, if he believes that there is smoke. Then, having this reason, the probability of fire increases for him.

A different, equally legitimate way to conceptualize the difference is the following:

(b) For this agent, given his conditional probabilities, smoke *would be* a reason to believe that there is fire. It becomes or *is* a reason when the agent notices smoke.

Alternatively, we can claim that

(c) *noticing smoke* is a reason to believe that there is fire.

This reason weakens when we move from noticing smoke to weaker forms of epistemic access. Believing that there is smoke and even weaker notions captured by the degree of belief in there being smoke would be reasons for believing that there is fire. Jeffrey has generalized Strict Conditionalization. Applied to our example Jeffrey Conditionalization gives us the following formula:

Jeffrey Conditionalization:
$\text{Prob}_{new}(F) = \text{Prob}_{old}(F/S) \cdot \text{Prob}_{new}(S) + \text{Prob}_{old}(F/\text{not-}S) \cdot \text{Prob}_{new}(\text{not-}S)$

According to Jeffrey, any chance of degrees of belief can have an epistemic impact. We are interested in cases in which the degree of belief that there is smoke increases.

(d) *Acquiring the belief that S* (in any degree) or increasing one's confidence in S is a reason to believe F if and only if the new probability

---

[4]   Note that although I rely on the basic Bayesian idea of connecting reasons with increases in probability, I depart from orthodox Bayesianism by explaining different reason concepts via differences in epistemic access conditions. I have argued that this is necessary to capture the difference between 'reasons' that influence and those that do not influence an agent's epistemic system.

of F (as determined by the appropriate probabilities via Jeffrey Conditionalization) is higher than the old one.

The alternatives (a) – (d) are different ways to conceptualize a Bayesian concept of a reason. With this in hand we can now turn to Broome's example. You acquire the belief that p to some degree – it need not be one. What happens when we conditionalize? Using Jeffrey Conditionalization on the formula

$$\text{Prob}_{new}(p)=\text{Prob}(p/p).\text{Prob}_{new}(p)+\text{Prob}(p/\text{not-}p).\text{Prob}_{new}(p)$$

the conditional probabilities are 1 and 0 respectively so that we get the result (which is undoubtedly correct) that the new belief in p is exactly the same as the new belief in p: $\text{Prob}_{new}(p)=\text{Prob}_{new}(p)$. Is coming to believe that p a reason to believe p? Acquiring any new degree of belief in p will never raise the probability of p. Thus, on the Bayesian understanding of what it is to be a reason, self-support is impossible.

Think of the conditional probability as a measure of the strength by which one proposition supports another. Each proposition supports itself perfectly. There is no loss in strength of support when one moves along the entailment relation. The concept of a reason and the related concept of justification, however, demand more. What I take to be a reason needs to be able to increase my belief in what it is a reason for. This, as we have shown, is impossible on the Bayesian view. Broome is wrong to think that the acceptance of mind-based reasons would commit one to accepting implausible forms of self-justification.

Above I have given alternative conceptualization of reason concepts in a probabilistic framework. According to (a) and (b), we would have to say that p is or would be a reason to believe that p, though it is not a reason an agent can have because it cannot increase an agent's confidence in p. According to (c), we would say that noticing that p can never be a reason for believing that p as it cannot increase its own probability. The same holds for (d). The alternatives (a) – (d) all try to capture the same epistemic situation (with weakening epistemic accessibility restrictions). The question whether it is the smoke or the noticing of the smoke or a belief change regarding the presence of smoke is a reason is just a question regarding the convenience of the adoption of one of these conceptual frameworks. If (a) – (d) do not yield normative differences, it does not matter whether we talk in world-based terms (when we say that smoke is a reason) or in mind-

based terms (when we focus on belief changes regarding smoke) about reasonable changes in beliefs.

If we demand of reasons an effect on an agent's beliefs, in particular on those beliefs that are supported by the reason, then we realize that, in the probabilistic framework adopted, access to p will not increase an agent's confidence in itself. Thus, contrary to Broome, we may accept mind-based reasons without having to accept implausible cases of self-justification.

4. Bootstrapping 3: Intentions are not reasons – The Metaphysical Reading

Holton explains the Bootstrapping Objection as follows:

> Forming an intention to do something surely cannot give one a reason to do it that one would not otherwise have. If it did, we could give ourselves a reason to do something just by intending to do it; and that cannot be right. (Holton 2004, 513)

Note that we can do things which exactly have the feature that, according to the Bootstrapping Objection, would be implausible. We can do things like promising which is such that it provides us with reasons simply by having done it. Compare how implausible an analogous bootstrapping objection would sound when applied to promising:

> Promising to do something surely cannot give one a reason to do it that one would not otherwise have. If it did, we could give ourselves a reason to do something just by promising to do it; and that cannot be right.

If we can simply do things that provide us with reasons, it does seem but a small step to regard mental attitudes as reason. Having formed a plan, or having developed in interest, would then signal a change in one's normative landscape. Had I never become interested in philosophy, for example, my reasons for reading philosophical books would have been minimal. The Bootstrapping Objection needs more motivation in order to have argumentative force.

In Broome (2001), Broome explains the Bootstrapping Objection as follows: 'The objection is that you cannot bootstrap a reason into existence from nowhere, just by forming an intention.' The bootstrapping objection

to the idea that intentions are reasons is presented as an application of the general metaphysical claim that something cannot come from nothing. 'You cannot bootstrap a reason into existence from nowhere.' A proponent of mind-based normativity, however, does not create reasons out of nothing.

To see why not, we should distinguish between two things: First, there are, what we might call 'reason facts', i.e. facts about what is a reason for what. The fact that a liquid is poisonous is a reason against drinking it. The fact that it would make him understand something important is a reason for explaining it to him. You do not create reason facts; you do not create facts about what is a reason for what. What one can create, however, are, secondly, the things which, according to the reason facts, are reasons. Whatever I do, I create many things. Some of these things are reasons for me and for others to do or not to do certain things. Most people agree that promising to F is a reason to F. One does not create this reason-fact that promising is a reason, but one can create promises. The promise has not been created out of nothing. If desires or intentions are reasons, then by coming to want something a reason has come into existence. There is nothing metaphysically puzzling about this. As we are able to create things, we are able to create facts which are reasons. This does not entail that we create facts regarding what is a reason for what.

When a defender of mind-based normativity talks of mental states as reasons, Broome accepts analogous wide-scope requirements. According to the mind-based account, one creates a reason by forming an intention, whereas Broome thinks that there is a requirement with the content that, if one intends, one does it. These views would merge if one could detach the consequent of the conditional obligation. Broome (2009, 130) endorses necessary detachment. 'Nec' stands for necessity, 'O' for obligation: if Nec(a) and O(if a, the b), then O(b). He applies this rule to promises:

> Let us suppose this is a requirement of morality: Morality requires of you that, if you have promised to *F*, you *F*. Now supposed you have promised to *F*. That is a fact you can do nothing about; you cannot change the past. Let us treat it as necessary. Then Necessary Detachment allows us to conclude that morality require of you that you *F*. In general, we can derive from the above principle that 'If you have promised to *F*, morality requires you to *F*'. I find this a satisfactory explanation of a feature of promising that has puzzled some philosophers. How

are you in a position, merely by saying something, to impose a moral requirement on yourself? The answer is that you are automatically and constantly under the conditional requirement that, if you have promised to *F*, you *F*. You do not bring this requirement on yourself; it is an inescapable requirement of morality. Then, when you make a promise to *F*, it simply follows that you are required by morality to *F*. Nothing surprising happens: a conclusion follows. (Broome 2009, 130)

You do not create reason-facts; they are independent of your powers. You do create facts which, according to reason-facts, are reasons. This is not a mysterious power. You are 'constantly under a conditional requirement': when you promise you make it the case that this requirement applies to you. Why should we not hold the same view about intentions? You are always under a conditional requirement and in intending you make it applicable to you. Any rule of detachment will make the wide-scope view, given the content of the requirements, which relate mental states, normatively equivalent to the narrow-scope view, according to which normativity is mind-based. My aim in this section was not to establish the mind-based view. My aim was to show that the mind-based view should not be rejected on metaphysical grounds. Creation *ex nihilo* is neither here nor there.

Rawls famously objected to utilitarianism that it would violate the separateness of persons. This is, of course, not a metaphysical objection. Utilitarianism does not deny that I am different from the person next to me and that I am, in this sense, a separate person. Utilitarianism violates the separateness of persons because it treats the good and bad things that befall different people as if we were all one person, society. This view is normatively implausible. It tells us that we can sacrifice the one if thereby many smaller goods come to the many. The same, I think, applies to the bootstrapping objection. It does not strengthen the bootstrapping objection if it is dressed up in metaphysical clothing.[5] The real force of the bootstrapping objection must be that the idea it tries to reject, namely that beliefs, intentions, or

---

[5]   It seems to me that the attraction of metaphysical formulations of the Bootstrapping Objection is their rhetorical force. Here is an example: 'Broome draws from Michael Bratman a powerful point [...] which he calls the bootstrapping objection... why should a rational man worry whether he has any reason to psi if he has only to conceive an intention to fi, to which psi-ing is a means, and a reason to fi will spring into existence like Athene from the head of Zeus?' Price (2008, 79).

desires are reasons, has implausible normative implications. It is to this interpretation that I turn next.

5. Bootstrapping 3 – The normative reading: First Version

Broome writes:

> The view that intentions are reasons is implausible. If you have no reason to do something, it is implausible that you can give yourself a reason just by forming the intention of doing it. How could you create a reason for yourself out of nothing? Suppose, say, that you have no reason either for or against doing some act, and you happen to decide to do it. Now you intend to do it. So now, if intentions are reasons, you have a reason to do it. Since you have no contrary reason not to do it, the balance of reasons is in favour of your doing it. You now actually ought to do it, therefore. But this is implausible. It is implausible that just deciding to do something can make it the case that you ought to do it, when previously that was not the case. (Broome 2001, 98)

Suppose I can either go to the left or go to the right and, suppose, I have no reason to prefer going one way to going the other way. I have to make a decision. Whatever I decide will explain why I did what I ended up doing. Does such a decision make it rational to go one way rather than another? What else could make it rational? Any hesitation in saying 'You have decided to turn left, so turn left' stems from the possibility of reconsideration. At some point, however, reconsideration itself will look irrational.

Here is another example (see Verbeek 2007): When you can bestow a benefit on either one of two equally deserving persons but not on both, it is fair to hold an equal-chance lottery. You have to decide whom to benefit depending on the outcome of the chance event: if heads comes up the first person gets it, if tails, the second person will benefit. There is no reason to prefer this assignment of benefits to outcomes of the coin toss to the opposite one. Once you have made this decision, however, and the coin has come up heads you have a very strong reason to provide the benefit in accordance with your antecedent assignment. This reason is created by nothing but your decision and, antecedently, you had no reason to prefer one alternative over the other. We would need a more detailed example in order to assess the idea that allowing intentions to be reasons would have im-

plausible normative consequences. Thinking about cases in which one has no reason to do this or that looks like appealing to cases of indifference in which simply forming the intention will create a reason to go one way rather than the other.

6. Bootstrapping 3 – The Normative Reading: Second Version

Broome illustrates the bootstrapping objection with the following example:[6] Should you go to Paris? Suppose you have decided; you intend to go. Broome writes:

> If the balance of antecedent reasons was in favour of your going to Paris, you ought to go there. You have made the right decision and you ought to carry it out. If the balance of antecedent reasons was against your going to Paris, you ought not to go there. You have made the wrong decision and you ought not to carry it out. Your intention itself does not count one whit in favour of going to Paris. It makes no difference to what you should do. What you should do depends only on your antecedent reasons. Suppose there is a slight balance of antecedent reasons against going, but you made a mistake in your calculations and wrongly decided to go. A short time later, having invested nothing in the decision, you discover your mistake. Should you change your mind? If intentions were reasons, there would automatically be a reason not to, and if the balance of antecedent reasons would be slight enough, you should stick to your decision. But actually you should change your mind. Since you have invested nothing in your wrong decision you should change it. (Broome 2001, 99)

One could dispute Broome's conclusion if one added to his example. Broome's example presupposes a thesis about the priority of action over intention. According to this thesis, the reasons for an action determine the reasons for the corresponding intention. If your reasons favour going to Paris, you ought to go and nothing more need or can be said about the reasons for intending to go. When we think about reasons for and against

---

[6]   This example structurally mirrors Bratman's case about meeting Susan or meeting Kathy from Bratman (1981), which in Bratman (1987) was changed to the Mondale example about what to do in a political debate.

going to Paris, we can summarize their effect in a judgment of the form that going is better for you than not going. Intending to go might have autonomous effects. Intending to go to Paris might be good for you, independently of the value of going there. Even if you will not enjoy Paris much, you might like to be the kind of person who can honestly say 'See you in Paris!' or 'Well, can't make it tonight, I'm off to Paris.'[7]

Broome, I said, presupposes a contentious priority view. However, the truth of the thesis, when applied to the case at hand, might just be built into the example – and this is, I admit, a natural way of reading it. Broome argues that it would lead to normative implausible consequences if we assumed that intentions are reasons. We would, he says, have to accept that we ought to carry out our irrational plans and this, I agree, does not sound right.

On a more general level, we might want to say something in its favor. Some states do not lose their normal reason-giving force, simply because they are the result of our own irrationality. Suppose you decided to make yourself very thirsty by not drinking anything for a whole day. There was no further point to it; it was simply a silly idea. Nevertheless being thirsty retains its normal reason-giving force. If you are thirsty, you should drink.

Thirst as well as other states of deprivation, are, however, a special case. In general, irrational states are normatively significant in the sense of being 'bad', i.e. one should abandon these states, one should try to get out of them. They are not normatively significant in the sense of reasons. The normal role of a belief as a reason is to support other beliefs; it is not to be eliminated. The normal normative role of an intention is to be fulfilled by our agency. This makes thirst special, because being eliminated (what is the common feature of bad mental states) is, in this case, its normal normative role. An irrational intention, however, should not be eliminated by being fulfilled – one should simply stop intending as one does. Thus I ac-

---

[7] Intentions, plans, commitments can have autonomous benefits, i.e. benefits which do not arise from benefits of what is intended. It is good for us to be able to form plans and commitments. For example, they allow us to overcome temptations. The acceptance of what I have called attitude-related reasons for intentions (see Piller 2006), becomes relevant for what we ought to do if we accept the priority of intention view. One should do what one should intend to do. What one should intend to so is determined by 'reasons for actions', i.e. by the benefits of what it is that we intend to do, and by the autonomous reasons for intending to do it.

cept the following lesson from Broome's example: Irrational intentions do not make doing what one intends to do rational.[8]

Let me come back to the beginning of my discussion. I complained about Broome's first example. If your evidence does not decide whether emissions of carbon dioxide are harmful, how can you simply believe that it is poisonous? That, I said, would be irrational. Now we have come to a point in unraveling the bootstrapping objection at which this irrationality has become the central feature of the objection. Irrationality does not generate rationality. Even if one's reasoning process is impeccable, if one starts from something silly one will end up with something equally silly. Someone's belief that France is still a monarchy is not well supported by his belief that he is the King of France.

Silly premise beliefs (or silly intentions) don't make it the case that we ought to have the respective conclusion beliefs (or that we ought to do what we intend to). What does this show about whether mental states can be reasons? Broome thinks it shows that beliefs and intentions cannot be reasons. We have to replace attitudinal reasons with wide-scope rational requirements. To make this view immune to the problem of normative inconsistency, the normativity of rational requirements is put in doubt. I offer a different reaction. We could keep the category of attitudinal reasons as long as we restrict these reasons to attitudes which are not normatively objectionable, i.e. we simply exclude irrational premise beliefs and irrational intentions from being reasons.

Would this be an ad hoc defense of attitudinal reasons? I don't think it would because the principle that irrationality does not generate rationality is independently plausible. Consider the probabilistic view of epistemic rea-

---

[8]   Bratman (1981) disagrees. He actually accepts that an irrationally formed intention can make it rational to do what one intends to do. Even irrational intentions can tip the balance of reasons. We end up with a case in which it is rational to take the means to an intended end that one should not intend. Bratman concludes, 'The slogan that rationality in intention and action just is rationality relative to the totality of relevant considerations, is a dogma that cases of bootstrap rationality force us to reject' (Bratman 1981, 265). If Bratman is right, this would open up a different line of resisting Broome's Bootstrapping Objection – Broome would not have found a case in which accepting intentions as reasons would have implausible normative implications. However, I will not take this line. On this matter, I agree with Broome: an irrational intention will not render rational what absent the intention would be irrational. This strikes me as true as long as autonomous benefits of having intentions are being excluded.

sons explained earlier. We have learnt that it is not enough that the conditional probability of the conclusion C is higher given the premise P than it would otherwise be. We have already met one other condition in the discussion of self-justification. There have to be circumstances in which an increase in my belief of P increases my belief in C. Believing P never increases (by itself) its own probability. Now we meet another condition. Only things that are themselves not normatively objectionable can do the work of reasons which is to show us what one ought to do or ought to believe. Compare the force of reasons to the illuminating effect a source of light has on the objects of its surroundings. In order to illuminate an object the object (a) has to be in a proper place where it can be reached by the light. (This is the conditional probability or the standing-in-the-being-a-reason-for relation.) Furthermore, (b) if there is black tape around the light source, it won't illuminate either: normative force can be undermined. It is working alright internally, and it would emit light if it were not for the black tape. Furthermore, the light source has to be in normal working order. If it is internally broken, it will not emit any light. This is the condition I emphasize here. The conclusion belief is in the right place. The premise stands in the being-a-reason-for relation to the conclusion belief but it is not a reason. Being broken, it does not emit any light.

I said that the silly premise belief is not a reason for the conclusion belief although it would have been a reason had it been normatively okay. No assumption has been made about what it means to be normatively acceptable. Acceptance of the irrationality-in/no-rationality-out principle is independent of the question whether normative acceptability should be understood in mind-based or in world-based terms. Thus our discussion of the bootstrapping objection has not brought to light an argument against a mind-based account of normativity. This was the point of the story with which I started. In this example my belief that I am a better than average philosopher is irrational on the basis of all the other things I believe. If everyone agrees that irrationality does not generate rationality and if, as I have tried to show, the bootstrapping objection boils down to this very claim, it cannot dislodge any view about the nature of reasons.[9]

---

[9] One might want to object to the idea that intentions are reasons on different grounds. If intentions would be reasons, in addition to the reason there are for what one intends, we end up double-counting. This worry, however, affects mind-based views in the same way as it affects world-based views. If we counted the goodness of an

There is another bootstrapping worry which has a different and more limited target. It does not object to mind-based normativity as such. It rather objects to a specific mind-based principle. Kolodny presents this worry as follows:

> Suppose I believe that I have conclusive reason to have some attitude. In some sense, I ought to have that attitude; it would be irrational of me not to have it. Now suppose that 'ought' here means 'have reason'. Then we get the bootstrapping result that if I believe that I have conclusive reason to have some attitude, then I in fact have reason to have it. This is absurd. (Kolodny 2005, 512)

This worry affects a principle which assigns the following role to normative beliefs: If you are convinced that you ought to do something, then you ought to do it. If this principle holds, believing that one ought to do something makes it true that one ought to do it. Accepting such a principle comes at the price of rejecting our fallibility in normative matters. This is, I agree, a high price. Only revisionism about the content of our obligations – they would have to be determined by the same features as our beliefs about our obligations – could render such an infallibility principle palatable. Nevertheless beliefs about what one ought to do have to play a central role in a normative theory. Having concluded one's deliberation, one has to be committed to doing what one has thus concluded. Otherwise, one would not have concluded one's deliberation. I have argued that the perception that the bootstrapping objection would refute any mind-based account of normativity is mistaken. However the question how a mind-based account explains the normative role of normative beliefs will here remain unanswered.[10]

---

option as well as the things that make it good, we encounter the same danger of double-counting. In both cases, we can attribute normative force to goodness or to the intention, as long as we ensure that double-counting will be avoided.

References

BRATMAN, M. (1981): Intention and Means-End Reasoning. *The Philosophical Review* 90, 252-265.

BRATMAN, M. (1987): *Intentions, Plans, and Practical Reason*. Cambridge, Mass.: Harvard University Press.

BROOME, J. (2001): Are Intentions Reasons? And How Should We Cope with Incommensurable Values. In: Morris, Ch. W. – Ripstein, A. (eds.): *Practical Rationality and Preference: Essays for David Gauthier*. Cambridge: Cambridge University Press, 98-120.

BROOME, J. (2007): Does Rationality Consist in Responding Correctly to Reasons? *Journal of Moral Philosophy* 4, No. 3, 349-374.

BROOME, J (2009): *How to Be Rational?* Manuscript, version June 2009.

BROOME, J (2013): *Rationality through Reasoning*. Oxford: Wiley-Blackwell.

EWING, A.C. (1947): *The Definition of Good*. New York: Macmillan.

HOLTON, R. (2004): Rational Resolve. *Philosophical Review* 113, 507-535.

KOLODNY, N. (2005): Why Be Rational? *Mind* 114, 509-563.

PILLER, C. (2006): Content-Related and Attitude-Related Reasons for Preferences. *Royal Institute of Philosophy Supplement* 59, 155-182.

PRICE, A. (2008): *Contextuality in Practical Reason*. Oxford: Oxford University Press.

PRICHARD, H. A. (1932): Duty and Ignorance of Fact. In: Prichard, H. A: *Moral Obligations*. Oxford: Oxford University Press, 18-39.

VERBEEK, B. (2007): Rational Commitment. In: Peter, F. – Schmid, H. B. (eds.): *Rationality and Commitment*. Oxford: Oxford University Press.

Contents · Obsah

DISCUSSIONS ▪ DISKUSIE

BOOK REVIEWS ▪ RECENZIE

REPORTS · SPRÁVY