

# Reasoning with preferences?

JOHN BROOME

## 1. Reasoning and requirements of rationality

Rationality requires certain things of you. It requires you not to have contradictory beliefs or intentions, not to intend something you believe to be impossible, to believe what obviously follows from something you believe, and so on. Its requirements can be expressed using schemata such as:

*Modus ponens.* Rationality requires of  $N$  that, if  $N$  believes  $p$  and  $N$  believe that if  $p$  then  $q$ , then  $N$  believes  $q$ .

*Necessary means.* Rationality requires of  $N$  that, if  $N$  intends that  $e$ , and if  $N$  believes that  $e$  will be so only if  $m$  is so, and if  $N$  believes  $m$  will be so only if she intends that  $m$ , then  $N$  intends that  $m$ .

*Krasia.* Rationality requires of  $N$  that, if  $N$  believes she ought to  $F$ , and if  $N$  believes she will  $F$  only if she intends to  $F$ , then  $N$  intends to  $F$ .

(‘She’ is to be read as a reflexive pronoun.) It may be questioned whether any of these formulae express genuine requirements of rationality. Their precise formulation may be inaccurate, at least. But these formulae are not the subject of this paper, and for the sake of argument I shall assume they are correct. In any case, they are only examples of requirements of rationality (or ‘rational requirements’, as I shall often say); rationality requires many things of you besides these. Notice that all of these particular requirements govern conditional statements. They have a ‘wide scope’, as I shall say. None governs a single belief or intention of yours.

Many people think that rationality makes requirements on your preferences, too. In order to have an example to work with, I shall concentrate on this familiar one:

*Transitivity.* Rationality requires of  $N$  that if  $N$  prefers  $a$  to  $b$  and  $N$  prefers  $b$  to  $c$ , then  $N$  prefers  $a$  to  $c$ .

This too has a wide scope. It is particularly controversial whether or not this is a genuine requirement of rationality. But in this paper

## John Broome

I shall not engage directly in controversy about it; I shall assume that *Transitivity* expresses a genuine requirement. I shall ask how, given that it is a rational requirement, you may come to satisfy it.

By what process can you come to satisfy a particular requirement of rationality? Often, you simply find yourself satisfying it. You intend to visit Venice; you believe the only way to do so is to buy a ticket (and that you will not do so unless you intend to); and you find yourself intending to buy a ticket. You satisfy *Necessary means* in this instance. You come to do so as a result of some automatic, unconscious causal process that you do not control; it just happens. Many of your preferences satisfy *Transitivity* in a similar way. Presumably there is some evolutionary explanation of why this sort of thing happens.

Possibly an ideally rational creature would find itself satisfying all the requirements of rationality this way. But mortals fail to satisfy very many of them. However, we mortals do have a way of improving our score. We can bring ourselves to satisfy some requirements by our own activity of reasoning. Reasoning is an activity—something we do—through which we can satisfy some requirements in particular instances. For example, we can come to believe a particular consequence of what we believe by thinking the matter through.

Some unconscious processes could be called unconscious reasoning. But in this paper I am interested only in conscious processes, and I shall give the name ‘reasoning’ to those ones only. Unconscious processes are not activities, and I am interested in reasoning as an activity.

I am assuming rationality imposes requirements on your preferences, such as *Transitivity*. No doubt you find yourself satisfying some of those requirements through unconscious processes. But when you do not, can you bring yourself to satisfy them through reasoning? Briefly: can you reason with preferences? That is the topic of this paper.

I am interested in correct reasoning only. Various mental activities of yours might accidentally lead you to satisfy a rational requirement, and various of those activities might qualify as reasoning. But a reasoning activity that systematically leads you to satisfy a rational requirement would have to be *correct* reasoning.

Why does it matter whether you can reason with preferences? It is important in itself to understand the process of reasoning, but there is another reason too. In ‘Why be rational?’, Niko Kolodny argues that, for any rational requirement on you, there must be a process of reasoning through which you can bring yourself to

## Reasoning with preferences?

satisfy that requirement. If he is right, and if it turned out that you cannot reason with preferences, it would follow that there are no rational requirements on preferences.

As it happens, I am not convinced by Kolodny's arguments, for reasons I cannot set out in this paper.<sup>1</sup> I remain agnostic about his conclusion. For all I know, there may be requirements of rationality that you can come to satisfy only by unconscious processes that you do not control. But even so, if it should turn out that no process of reasoning could bring you to have, say, transitive preferences, that would cast some doubt on the claim that rationality requires you to have transitive preferences. We would certainly want an explanation of how there could be this requirement on you without your being able to bring yourself to satisfy it. In this way, the question of reasoning reflects back on to the question of what rationality requires.

You certainly cannot rely on unconscious processes to get all your preferences into rational order; anyone's system of preferences is too big and complex for that. This is particularly true of preferences among uncertain prospects. The axioms of expected utility theory are supposed to express requirements of rationality for these preferences, and no one satisfies those axioms automatically.

Reasoning with preferences, and indeed reasoning in general, has not been much discussed. Many authors write about what rationality requires of your preferences and other mental states. Having stated some requirements, they leave it at that. They do not consider by what process you may come to satisfy their requirements. Why not? I think they must take it for granted that, once you know what the requirements of rationality are, you can bring yourself to satisfy them by reasoning. I think they must implicitly rely on a particular model of reasoning. They must think you can reason your way to satisfying a requirement by starting from the requirement itself as a premise. More exactly, their model starts from your believing some proposition such as the ones I have labelled *Modus ponens*, *Necessary means*, or *Transitivity*, and you reason from there. These are propositions about your mental states, so your reasoning starts from a belief about your mental states. I shall call this a 'second-order belief', and I shall call this model of reasoning the 'second-order model'. It is an all-purpose model. It can be applied to reasoning with mental states of all kinds—beliefs, intentions, preferences and so on.

<sup>1</sup> See my 'Wide or narrow scope?'

## John Broome

But for some mental states, reasoning cannot work as the second-order model supposes. The model does not work for beliefs, for one thing. Section 2 explains why not. Section 3 describes an alternative, first-order model of reasoning, which is more successful for beliefs. It does not depend on any second-order belief about your mental states. But it is not such an all-purpose model; it is not straightforward to extend it beyond beliefs to other mental states. I shall next consider how successfully the two models can apply to preferences. Section 4 distinguishes a broad concept of preference from our ordinary one, as I need to do. Section 5 applies the second-order model to broad preferences with moderate success. Section 6 applies the first-order model to ordinary preferences, again with moderate success. The central issue that arises in section 6 is how far ordinary preferences, can be distinguished from beliefs about betterness. It may turn out that what appears to be reasoning with ordinary preferences is really nothing other than theoretical reasoning about which alternatives are better than which. Section 7 considers whether that is so.

My main conclusion is that the second-order model of reasoning is unsuccessful for ordinary preferences, as it is for beliefs. Possibly this model may work for broad preferences. Nevertheless, we may indeed be able to reason with ordinary preferences, because the first-order model is more successful. However, I remain unsure that first-order reasoning with preferences is really distinct from theoretical reasoning about betterness.

## 2. Second-order theoretical reasoning

I start with theoretical reasoning—reasoning with beliefs. I shall use an example in which you come to satisfy the requirement *Modus ponens*. It is a case of simple deductive reasoning, which should be paradigmatic of theoretical reasoning.

You wake up and hear rain, so you believe it is raining. Your long experience with snow has taught you that, if it is raining, the snow will melt. However, because you are still sleepy and have not yet thought about the snow, you do not yet believe the snow will melt. So you do not satisfy *Modus ponens* in this instance. You believe it is raining; you believe that if it is raining the snow will melt, but you do not believe the snow will melt. By reasoning, you can surely bring yourself to satisfy the requirement in this instance. How will your reasoning go?

## Reasoning with preferences?

This section investigates the second-order model. I shall take a generally sceptical stance towards it. I shall argue it does not work for theoretical reasoning, nor for reasoning with ordinary preferences. Given that, I shall be generous towards this model, and make concessions to help it on its way. I shall make assumptions that support it, even when I cannot fully justify them.

The second-order model supposes that your reasoning sets out from a belief in the requirement itself. So let us suppose you do actually believe the requirement *Modus ponens* in this instance. You believe rationality requires of you that: you believe the snow will melt if you believe it is raining and you believe that if it is raining the snow will melt. Can you get by reasoning from this belief to satisfying the requirement itself, as the second-order model supposes?

One plausible pattern of reasoning offers a clue as to how you might do so. Suppose you believe you ought to do something—buy cherries, say. You might say to yourself:

*I ought to buy cherries,  
So I shall buy cherries.*

I mean the second of these sentences to express an intention of yours, rather than a belief that you will buy cherries. I shall say more about the idea of saying to yourself in section 3. This is plausibly a little piece of reasoning, through which your normative belief that you ought to buy cherries brings you to form the intention of buying cherries. Normally, when you intend to do something, your intention causes you in due course to do it. So in due course you are likely to buy cherries, as a final result of your normative belief that you ought to do so.

I think that what you say to yourself here is indeed reasoning, and moreover correct reasoning. By means of reasoning on this pattern, you can bring yourself to satisfy the rational requirement *Krasia*: to intend to do what you believe you ought to do. I shall call it 'kratic reasoning'. In this paper I shall not argue that kratic reasoning is genuine, correct reasoning; I shall simply assume it is. I do so to smooth the way for the second-order model; it is one of my concessions to the model. In a moment, I shall show how the second-order model can make use of it.

As a second concession, I shall assume you can derive a strictly normative belief from your belief in the rational requirement. I have already assumed you believe rationality requires you to satisfy the condition that you believe the snow will melt if you believe it is raining and you believe that if it is raining the snow will melt. Now,

## John Broome

I assume you go further and derive the belief that you ought to satisfy this condition. Questions might be asked about this step.<sup>2</sup> First, even though rationality requires you to satisfy this condition, does it follow that you ought to satisfy it? Suppose, for instance, very bad consequences would result from your satisfying it; ought you to satisfy it then? Second, even if it does actually follow, how can we assume you make this inference, so it is reflected in your own beliefs?

To give the second-order model a chance, I cannot avoid making this questionable assumption. If correct second-order reasoning is to bring you to satisfy some condition, you need to believe you ought to satisfy it. It is not good enough for you to believe merely that rationality requires you to satisfy it. Suppose, say, you believed rationality requires you to satisfy a condition but also believed you ought not to satisfy it. In that case, correct reasoning could not possibly lead you to satisfy it. So correct reasoning needs an ought belief, not merely a belief about a rational requirement.

I give the model an ought belief, therefore. I assume you believe you ought to believe the snow will melt if you believe it is raining and you believe that if it is raining the snow will melt. That should put you in a position to go through this piece of kratic reasoning, modelled on the cherries example:

*I ought to believe the snow will melt if I believe it is raining and I believe that if it is raining the snow will melt*  
*So I shall believe the snow will melt if I believe it is raining and I believe that if it is raining the snow will melt*

The second sentence is supposed to express an intention. Because the content of your premise-belief has a wide scope, you end with an intention that has a wide scope. What you intend is the conditional proposition that you believe the snow will melt if you believe it is raining and you believe that if it is raining the snow will melt.

Suppose you get as far as this. What happens next? If you are to follow the precedent of cherries, this intention would normally cause you to fulfil it. But there are two difficulties standing in the way of that result.

The first is the wide scope of your intention. Kratic reasoning could take you to a more specific intention only if you started with a more specific normative belief. To get by kratic reasoning to an intention to believe the snow will melt, you would have to start

<sup>2</sup> See my 'Does rationality give us reasons?'

## Reasoning with preferences?

from a belief that you ought to believe the snow will melt. But you cannot acquire this specific normative belief by correct reasoning from your initial belief in the broad-scope rational requirement you are under.

To see why not, notice it may not be true that you ought to believe the snow will melt. Perhaps you ought not to believe it is raining; perhaps the rain you hear is on a recording that you set as your alarm call. If you ought not to believe it is raining, it may well not be the case that you ought to believe the snow will melt. On the other hand, we are assuming it is true that rationality requires you to believe the snow will melt if you believe it is raining and you believe that if it is raining the snow will melt. You cannot by correct reasoning derive a belief that may not be true from one that is true.

So by correct kratic reasoning you cannot arrive at an intention to believe specifically that the snow will melt. But it is that specific belief the reasoning is supposed to lead you to. That is the first difficulty.

It may not be a serious one. All your intentions are indefinite to some degree, and yet you manage to fulfil many of them. If you intend to buy cherries, you could fulfil your intention by going to the greengrocer or the supermarket, in the morning or the afternoon. Somehow your intention gets narrowed to a more specific one, say to buy cherries at the supermarket, leaving home at 12.30. This narrowing can happen without your having a normative belief that you ought to buy cherries at the supermarket, leaving home at 12.30. It certainly can happen; we do not have to worry about how. I shall assume the same thing could happen in the present case. I shall assume your wide-scope intention could be narrowed to an intention to believe the snow will melt. This is rather plausible, since you do in fact believe it is raining and that if it is raining the snow will melt. I treat it as another concession to the second-order model.

But now you meet the second difficulty. This is the fatal one. Intending to believe a particular proposition is normally ineffective; it normally does not get you to believe the proposition. (Because you probably know that, you probably cannot even form an intention to believe a particular proposition. You cannot intend something and at the same time believe the intention will be ineffective.)

There are exceptions. You may be able to acquire a belief in a particular proposition by using some external means—going regularly to church or taking a belief pill, for example. If an

## John Broome

external means is available to you of coming to believe a particular proposition, then you may be able to intend to believe this proposition, and this intention may cause you to believe it, using the means. However, the last step—using an external means such as going regularly to church or taking a belief pill—is not a mental process. It therefore cannot form part of a process of reasoning. So the second-order model of reasoning cannot work through your using an external means.

On the other hand, you cannot come to believe a proposition by intending to believe that proposition, without using an external means. You can do some things without using an external means; raising your hand is one example. Intending to raise your hand can bring you to raise your hand without using an external means. But intending to believe a proposition cannot bring you to believe that proposition without using an external means. In his 'Deciding to believe', Bernard Williams argued this a necessary feature of belief; I have been persuaded by an argument of Jonathan Bennett's that it is a contingent feature of our psychology.<sup>3</sup> But whether necessary or contingent, it is a truth. It prevents the second-order model of theoretical reasoning from working in the way I have been investigating.

That way was through *kratic* reasoning, by which a normative belief leads to an intention. Could the second-order model work more directly, without involving any intention? Could it be that believing rationality requires you to be in a particular mental state, or believing that you ought to be in a particular mental state, simply causes you to enter that state, without your forming an intention of doing so? Could this happen in a way that is sufficiently regular to count as reasoning?

T.M. Scanlon thinks it can happen for some states: those he calls 'judgement-sensitive attitudes'. These are 'attitudes that an ideally rational person would come to have whenever that person judged there to be sufficient reason for them ...'.<sup>4</sup> So, for instance, if you were ideally rational, you would come to have a belief whenever you judged there to be sufficient reason for you to have it or, as I prefer to say, whenever you judged you ought to have it.

I find Scanlon's view implausible. Your beliefs are not normally caused by any normative beliefs you might have about what you ought to believe. If you believe you ought to have some belief, that would not normally cause you to have the belief. Suppose you

<sup>3</sup> 'Why is belief involuntary?'

<sup>4</sup> *What We Owe to Each Other*, p. 20.



## Reasoning with preferences?

believe you ought to believe you are attractive, because believing you are attractive will relax you, make you more approachable and improve your life. This would not normally cause you to believe you are attractive. Normally, our beliefs are caused by evidence, not by normative beliefs about what we ought to believe.

I agree that beliefs are judgement-sensitive in a different sense. If you were ideally rational, you would come to have a belief whenever you judged there was sufficient evidence for the content of the belief. You would come to believe you are attractive when you judge there is sufficient evidence that you are attractive. Beliefs are genuinely judgement-sensitive in this sense, but it is not Scanlon's sense. Your judgement in this case is about the content of the belief, not about the belief itself. It is a first-order belief, not a second-order one.

Judgement-sensitivity in Scanlon's sense is sensitivity to a second-order normative judgement about the belief itself. A second-order judgement of this sort often accompanies a first-order one. When you judge there is sufficient evidence for some proposition, you may well also judge you have sufficient reason to believe the proposition. But what causes you to believe the proposition, if you do, is the first-order judgement, not the second-order one. A way to test this is to look at cases where you make the second-order judgement but not the first-order one. My example of believing you are attractive is one of those. Examples like that show a second-order judgement does not normally cause you to have the belief.

In any case, even if beliefs were judgement-sensitive in Scanlon's sense, that would not directly help the second-order model of reasoning. In my example, your second-order judgement is not that you ought to have a particular belief. Instead, it has a wide scope. It is the judgement that you ought to satisfy the conditional: that you believe the snow will melt if you believe it is raining and you believe that if it is raining the snow will melt. It is particularly implausible that this judgement could cause you to enter the complex mental state described by the conditional, without *kratic* reasoning and without your forming an intention.

I conclude that the second-order model of reasoning fails for theoretical reasoning. It requires a sort of control over your beliefs that actually you do not have. So I come to the first-order model.

### 3. First-order theoretical reasoning

I shall stick to the same paradigmatic example of theoretical reasoning. You believe it is raining, and you believe that if it is raining the snow will melt, but you do not believe the snow will melt. So you do not satisfy the requirement *Modus ponens* in this instance. But you can bring yourself to satisfy it by saying to yourself that:

It is raining  
If it is raining the snow will melt.  
So the snow will melt.

Here, I have written down a sequence of sentences, which designate propositions. You do not necessarily say the sentences to yourself; you might reason in Swedish, say. But you do say to yourself the propositions that these sentences designate. You say to yourself that it is raining, and that if it is raining the snow will melt, and then you say that the snow will melt. I shall mention the point of the word 'so' at the end of this section.

You initially believe the first two of these propositions; in saying them to yourself you are expressing your beliefs. You do not initially believe the third. But when you say it to yourself, you express a belief in it. By the time you come to say it, your reasoning has brought you to believe it. By this time, you satisfy *Modus ponens*. That is how the first-order model of reasoning works.

The propositions you say to yourself constitute the contents of your beliefs. You can reason with beliefs only because they are states that have contents. Their content gives you something to reason about.

Saying something to yourself is an act. Sometimes no doubt, you say things to yourself out loud, but more often you do it silently. In that case, I could alternatively have said you call the proposition to mind; 'saying to yourself' is just a more graphic way of describing what you do. One thing it does is to bring the beliefs together, if you have not previously done that in your mind.

Your acts of saying to yourself are part of your reasoning but not the whole. Your reasoning is the causal process whereby some of your mental states cause you to acquire a new mental state. It includes a sequence of acts, and it is itself a complex activity. To be reasoning, the process must involve acts of saying to yourself. Some of your beliefs cause you to acquire a new belief, through some acts of this sort. The process ends when you acquire your new belief.

## Reasoning with preferences?

The acquisition of this belief is an act. Described one way, the acquisition is something you intend. When you embark on your reasoning, you intend to come to believe whatever is the conclusion that emerges from the reasoning. You intend that, if  $p$  is the proposition that emerges from the reasoning, you believe  $p$ . However, you do not intend to believe the specific proposition that emerges. In the example, you do not intend to believe the snow will melt. Coming to believe the snow will melt is an act like finding your glasses under the bed, after looking for them. You intend to find your glasses, and this makes it the case that your finding them under the bed is an act. But you do not intend to find them under the bed. I said in section 2 that you cannot come to believe a particular proposition by intending to believe that proposition. But you can acquire a belief by means of a procedure you intend.

Since reasoning is a process that takes place among mental states, acts of saying to yourself can only form a part of it when they express mental states. In the example, in saying to yourself that it is raining, you must express a belief of yours that it is raining. When you say to yourself that the snow will melt, you must express a belief of yours that the snow will melt, and so on. In the context of belief, saying to yourself is asserting to yourself. True, you could say to yourself the sequence of sentences

*It is raining*  
*If it is raining the snow will melt*  
*So the snow will melt*

even if you did not have the corresponding beliefs. (In this paper, I use italics in place of quotation marks.) But in doing that you would not be reasoning because you would not be going through a process that takes place among your beliefs.

In the course of your reasoning, you do not say to yourself any propositions about your mental states; you say to yourself the propositions that constitute the contents of your mental states. In the example, you do not say to yourself that you believe it is raining, nor that you ought to believe the snow will melt. No second-order beliefs about your mental states are involved. We may say you reason *with* your beliefs. You reason *about* the content of your beliefs.

The second-order model of reasoning was supposed to set out from a belief about your beliefs. But it was blocked because there is no route of reasoning from there to actually modifying your first-order beliefs. On the other hand, the process I am now describing directly modifies your first-order beliefs, because it

## John Broome

works on their contents. When you conclude that the snow will melt, in doing that you are directly acquiring a new belief.

This needs emphasis. There are two aspects to theoretical reasoning. One is identifying a particular conclusion-proposition on the basis of the premise-propositions. The other is your coming to believe the conclusion-proposition. It is tempting to try and divide reasoning into two stages according to these two aspects: first picking out a new proposition, then coming to believe it. But if there were these two stages, at the end of the first stage the new proposition would be parked somewhere in your consciousness, without your having any particular attitude towards it. We would have to explain how you then come to believe it. The explanation could not go through your believing you ought to believe it, nor through your intending to believe it, because, as I said earlier, neither of these attitudes will succeed in getting you to believe it. At least, they cannot have this effect through any process that can be reasoning. In any case, this explanation would leave us with the equally difficult task of explaining how you come to have one of these attitudes.

The truth is that you believe the proposition as you identify it. We cannot split reasoning into the two stages. Theoretical reasoning is imbued with belief all the way through. As I put it just now: you are reasoning with beliefs. You do not reason and then acquire a belief.

To summarize what we have learned so far from this paradigmatic example: reasoning is a process whereby some of your mental states give rise to another mental state; the mental states involved must be ones that have contents; in reasoning you say to yourself the propositions that constitute these contents, and you reason about these contents.

This cannot be a full characterization of reasoning. Not just any mental process that has these features is reasoning. For example, suppose you believe that it is raining and that if it is raining the snow will melt. Suppose you say to yourself that it is raining and that if it is raining the snow will melt, and suppose this causes you to believe you hear trumpets. That bizarre process is probably not reasoning.

You might think that true reasoning can only be separated from bizarre processes like this by the presence of a second-order belief. In my example of genuine reasoning, you moved from believing it is raining and believing that if it is raining the snow will melt to believing the snow will melt. You might think this process is reasoning only if you have the second-order belief that rationality

## Reasoning with preferences?

requires you to believe the snow will melt if you believe it is raining and you believe that if it is raining the snow will melt.

Even if this was so, it would not restore the second-order model of reasoning. The reasoning is still conducted at the first order, even if a second-order belief needs to be present in the background. But actually I think it is not so. A sophisticated reasoner may have this second-order belief, but I do not see why you need so much sophistication in order to reason. I do not see why you need to have the concept of a rational requirement, or even the concept of a belief.

It is more plausible that a different sort of background belief is needed to separate your reasoning process from others such as the bizarre one. You might need to believe that, from the proposition that it is raining and the proposition that if it is raining the snow will melt, it follows that the snow will melt. That is to say, you might need in the background, not a second-order belief about what rationality requires of your beliefs, but a belief about the inferential relations that hold among the propositions that constitute the contents of your beliefs. I do not deny that a belief such as this may be a necessary conditions for you to reason. But even if it is necessary in the background, it is not itself a part of the reasoning; it does not constitute an extra premise. That is the lesson taught us by Lewis Carroll in 'What the tortoise said to Achilles'. So the first-order model of reasoning is not affected, even if this belief is necessary in the background.

My own view is that reasoning processes are computational. This is what characterizes them as reasoning and distinguishes them from bizarre ones such as the one I described. If I am right, it adds to the ways in which reasoning is an activity, since computation is something you do. You operate on the contents of your beliefs computationally. I think that, when you say to yourself the word 'so' or its equivalent in another language, it marks your computation. Computation is too big and difficult a topic to broach in this paper. I shall allow myself the assumption that theoretical reasoning is an operation on the contents of beliefs.

My snow example is paradigmatic of theoretical reasoning, in that it is an example of deductive reasoning by *Modus ponens*. But it represents only a small fraction of theoretical reasoning, and it leaves a great deal to be explained. For one thing, reasoning often does not proceed in the linear fashion illustrated in the example. In the example, your reasoning sets out from some initial beliefs and concludes with a new belief. But theoretical reasoning often leads you to drop one or more of your initial beliefs, rather than acquire a

new one.<sup>5</sup> Dropping a premise-belief will bring you to satisfy the requirement *Modus ponens* just as well as acquiring a conclusion-belief will. A fuller account of theoretical reasoning will need to explain how it can turn around and have this backwards effect. Besides that, there are many other patterns of theoretical reasoning to be accounted for too. But none of that is for this paper. I described theoretical reasoning only in order to illustrate the two different models of reasoning. Now I turn to preferences.

#### **4. Concepts of preference**

I need first to distinguish two concepts of preference. This conventional definition defines a broad concept:

*Broad preference.* *N* prefers *a* to *b* if and only if *N* is in a mental state that would typically cause *N* to choose *a* were *N* to have a choice between *a* and *b* only.

We call the mental state a *preference* for *a* over *b*.

This definition is broad because it allows mental states of various sorts to count as preferences. For one thing, it allows an intention to be a preference. Suppose you intend to choose biking if ever you have a choice between biking and driving only. This is a state that would typically cause you to choose biking, were you to have a choice between biking and driving only. So you prefer biking to driving according to the definition.

This definition is too broad to capture accurately our ordinary concept of a preference. Ordinarily, we make a difference between preferring one thing to another and intending to choose one thing rather than another. You might intend to choose biking—perhaps on grounds of health—though actually you prefer driving. You can intend to choose something you do not prefer, and you can prefer something you do not intend to choose. The definition does not allow for that possibility.

According to our ordinary concept, a preference is like a desire rather than like an intention. It is a sort of comparative desire. The notion of preference may even be reducible to the notion of desire: to prefer *A* to *B* may simply be to desire *A* more than *B*. What is the difference between a desire and an intention? To specify the difference analytically is a difficult and contentious matter. Both

<sup>5</sup> Gilbert Harman particularly emphasizes this point in *Change in View*.

## Reasoning with preferences?

desires and intentions are mental states that can be identified by their functional roles; the difficulty is to spell out what their different roles are. They are similar in that a desire to do something and an intention to do something are both dispositions to do that thing. But they are dispositions of different sorts. In so far as they cause you to do the thing, they do so in characteristically different ways. It is difficult to spell out their different roles in detail. For my purposes I do not need to. We naively have a good understanding of the difference between a desire and an intention, and I only need to remind you of it. The next two paragraphs do so.

Desires are more remote from action than intentions are. When you intend to do something, you are committed to doing it, but that is not necessarily so when you desire it. To a large extent, your intentions control your actions. Often they do so through processes of reasoning, specifically through instrumental reasoning in which you figure out appropriate means to ends that you intend.<sup>6</sup> On the other hand, in so far as your desires influence your actions, they generally do so through your intentions. To desire to do something is to be disposed to intend to do it. Since to intend to do it is itself to be disposed to do it, to desire to do something is also to be disposed to do that thing, but more remotely. A desire of yours is only one influence on your intentions. Other influences include other desires that may conflict with it, your beliefs about what you ought to do, whims that strike you, confusions that afflict you, and so on. Consequently, if you desire to do something, you may not intend to do it, and you may intend to do something without desiring to do it.

You can acquire an intention by making a decision. For example, you may one day decide to go to Venice, and you will then intend to go to Venice. But deciding to go to Venice does not make you desire to go to Venice. You cannot acquire a desire by making a decision, without using an external means. You may have an external means available of acquiring the desire to go to Venice; you might spend hours poring over glossy picture books, for example. If so, you can decide to acquire the desire, and then acquire the desire using the means. But you cannot acquire the desire by deciding to, without using an external means. In this respect a desire is like a belief. I said it is a contingent fact of our psychology that you cannot acquire a belief by deciding to acquire it, without using an external means. I think the same is true of a desire.

<sup>6</sup> Michael Bratman's *Intention, Plans and Practical Reason* is a full account of the characteristic role of intentions in controlling actions.

## John Broome

According to our ordinary concept, a preference is like a desire in this respect. You cannot acquire an ordinary preference by deciding to, without using an external means. In his paper in this volume, Christian Piller claims that you can decide to have a particular preference, but I disagree with him about that if he is thinking of an ordinary preference.<sup>7</sup> His example is this:

What if we got two pots of gold, if we preferred this saucer of mud to a pot of gold? I would certainly say 'Yes, please, can I have the saucer of mud'... . If I honestly and instantaneously say 'I want the mud, not the gold. Please!' then I do prefer the saucer of mud to the pot of gold.

If the prize of two pots of gold is awarded for having a broad preference for the saucer of mud over a pot of gold, Piller wins it fair and square. A broad preference can be acquired by decision. In this case, Piller acquires by decision the disposition to choose the saucer of mud rather than a pot of gold. This disposition is the prize-winning broad preference.

However, if the prize is awarded for having an ordinary preference for the saucer of mud rather than a pot of gold, Piller is not entitled to it. He may say 'I want the mud, not the gold. Please!', but that utterance has to be understood as a pressing request to be given the mud. I do not suggest he is dishonest in making it. However, if he really meant to assert that he wants the saucer of mud more than a pot of gold, I am sorry to say I would not believe him. His sorry tale makes it plain that gold is all he wants; he has no desire for the mud. His decision to choose the saucer of mud does not give him an ordinary preference for the mud over a pot of gold.

You can acquire some broad preferences by making a decision, because those broad preferences are intentions. Those broad preferences are not ordinary preferences. On the other hand, all ordinary preferences are broad preferences. They satisfy the definition: an ordinary preference for *a* over *b* is a mental state that typically causes you to choose *a* over *b*. But not just any mental state with this property is an ordinary preference. Evidently more conditions need to be added to the definition of a broad preference if we are to arrive at a correct definition of an ordinary preference.

<sup>7</sup> I have no quarrel with Piller's conclusion that there can be attitude-based reasons for a preference, even an ordinary preference. Just because you cannot choose to have an ordinary preference, it does not follow there are no attitude-based reasons for you to have it.



## Reasoning with preferences?

In his paper in this volume, Philip Pettit argues like me that the concept of broad preference is broader than our ordinary concept. He also thinks that more conditions must be added to the definition. He mentions conditions on the mental state's collateral connections with other mental states. The axioms of decision theory illustrate the sort of conditions he has in mind. But Pettit's objection to broad preference is different from mine. If a creature's behaviour is very chaotic, we might not be able to recognize the creature as having preferences at all. So even if it was in one particular state that met the definition of a broad preference, we might not count that state as truly a preference. That is Pettit's concern, and it is a real one. But only minimal further conditions are required for this reason. If a pigeon nearly always circles to the left, we have no difficulty in attributing to it a preference for circling to the left rather than the right, even if the rest of its behaviour is fairly chaotic. Certainly, we may have preferences that are very far from satisfying the axioms of decision theory.

To define a preference in the ordinary sense, we need to add conditions of a different sort from Pettit's. They need to distinguish a preference from an intention, and they will have to do so by specifying its functional role. As I say, this is difficult to do, and I shall not try to do it here. I hope I have said enough to separate the ordinary concept of preference from the broad one, by recalling our ordinary understanding of the difference between a preference and an intention.

### 5. Second-order reasoning for broad preferences

The central question of this paper is whether there is an activity of reasoning by means of which you can bring yourself to satisfy requirements of rationality on preferences. Now we have two concepts of preference, this question divides into two. Can you reason with broad preferences? Can you reason with ordinary preferences? I shall start with broad ones.

The broad concept of preference is an artificial, theoretical one. Nevertheless, it seems to be the one most authors have had in mind when they consider rational requirements on preferences. The most popular defence of the requirement *Transitivity* is the money-pump argument, which is directed at broad preferences. Here is the

## John Broome

argument, put briefly.<sup>8</sup> Suppose you prefer *a* to *b* and you prefer *b* to *c*, but you do not prefer *a* to *c*. For simplicity, assume that your preferences are complete, so that, since you do not prefer *a* to *c*, either you prefer *c* to *a* or you are indifferent between *a* and *c*. Suppose you initially possess *c*. Now a dealer offers to swap *b* for your *c*, provided you pay her some small fee for making the transaction. Since you prefer *b* to *c*, you agree if the fee is small enough. Now you possess *b*. Next, this dealer offers to swap *a* for your *b*, again for a small fee. If the fee is small enough, you again agree. Finally, she offers to swap *c* for your *a*, this time without a fee. Since you either prefer *c* to *a* or are indifferent between the two, you are willing to make this transaction too. If you do make it, you end up possessing *c*, having handed over two small fees. You are back where you started, but poorer. It seems irrational to have preferences that allow you to be exploited in this way. That is the money-pump argument.

In this story, it is your dispositions to choose that allow you to be exploited. These dispositions constitute your broad preferences. Your ordinary preferences do not come into the argument. So the money-pump argument applies to broad preferences and not ordinary ones. It is an example of a class of arguments known as 'pragmatic arguments', which are supposed to demonstrate that rationality imposes various requirements on your preferences. All of them are aimed at broad preferences.

Because a broad preference can be an intention, you may be able to acquire a broad preference by making a decision. This opens the possibility that the second-order model of reasoning can work for broad preferences. That is, you may be able to reason your way from a belief in the requirement itself to satisfying the requirement. Since I have already set out the steps of the second-order model in the context of theoretical reasoning, I need only retrace them very quickly here. Suppose that, in the broad sense, you prefer biking to walking, and you prefer walking to driving, but you do not prefer biking to driving. You do not satisfy *Transitivity*. But suppose you believe in the requirement of transitivity itself in this instance: you believe rationality requires you to prefer biking to driving if you prefer biking to walking and walking to driving. (Perhaps you have been convinced by the money-pump argument.) Suppose indeed you have the normative belief that you ought to prefer biking to driving if you prefer biking

<sup>8</sup> Details of the argument are debated. The most convincing version of it appears in Wlodek Rabinowicz's 'Money pump with foresight'.

## Reasoning with preferences?

to walking and walking to driving. By kratic reasoning, you might be able to form the intention of preferring biking to driving if you prefer biking to walking and walking to driving. The content of this intention is a conditional proposition, but since you actually satisfy the antecedent of the conditional—you prefer biking to walking and walking to driving—you may be able to narrow the intention down to a simple intention to prefer biking to driving. If so, you now intend to have a particular preference.

At the corresponding point in my discussion of theoretical reasoning, you had arrived at the intention to believe the snow will melt. There, I said this intention is ineffective, because intending to believe something cannot normally bring you to believe it, except by using an external means. But it seems that your intention to prefer biking to driving may be effective; it may cause you to have this preference, without your using an external means.

It is an intention to have a broad preference: to be in a mental state that would typically cause you to choose biking were you to have a choice between biking and driving only. You will have this broad preference if you intend to choose biking if ever you have a choice between biking and driving only. And that state of intention seems to be one you can put yourself into simply by deciding to choose biking if ever you have a choice between biking and driving only. So it seems your intention to prefer biking to driving may cause you to prefer biking to driving, without your using an external means. The only means you require is to make a decision. This is a mental act, and it may therefore form part of a reasoning process.

That was quick. I have apparently mapped out a complete route whereby *second-order reasoning* could bring you to satisfy the requirement *Transitivity*, by acquiring the preference you need in order to satisfy it. However, there are several questionable steps along the route. In section 2, where I developed the second-order model of reasoning, I made questionable assumptions as concessions to the model. So I do not insist that the second-order model works for broad preferences; I simply cannot rule it out. Since broad preferences are not preferences as we ordinarily understand them, I pass quickly on to those that are.

### 6. First-order reasoning with ordinary preferences

For ordinary preferences, the second-order model can quickly be ruled out. You cannot acquire an ordinary preference by making a

## John Broome

decision, without using an external means. This is one of the characteristics that distinguish an ordinary preference from other broad preferences. It follows that second-order reasoning will not work for ordinary preferences. The argument is the same as the one I gave for second-order theoretical reasoning.

What about first-order reasoning? First-order reasoning for preferences would be reasoning *with* preferences, about the contents of preferences, rather than reasoning *about* preferences. Is there such a thing? The account I gave of first-order reasoning for beliefs was special to beliefs. If we are to extend it to states other than beliefs, we shall need a separate account for each state. We need one for preferences.

There is a general difficulty in the way of understanding how you can reason with states other than beliefs, operating on their contents in the way first-order reasoning requires. Beliefs have a special feature that allows you to do this sort of reasoning. When you say to yourself that it is raining, you express your mental state of belief. You also, in a different sense, express the content of that belief. You say that it is raining, which is to express the proposition that it is raining, which is the content of your belief. So you express the belief and its content together.

First-order reasoning requires this sort of double expression. It is reasoning with mental states, and you have to express those states in order to reason with them. But as well as that, reasoning is about the contents of the mental states. You need those contents before your mind, which means you have to present them to yourself, or express them to yourself. So your expression of your states also has to express the contents of those states.

But at first sight, few mental states share with beliefs the property that you can express them and their content together. Consider a desire, for example. We normally take a desire to have a content, and most philosophers take its content to be a proposition. Suppose you want to be loved. Then according to the common view, the content of your desire is the proposition that you are loved. But suppose you expressed this content by saying 'I am loved'. Then you would not be expressing the desire. If you are expressing any mental state of yours, it would have to be a belief that you are loved. You can only express this belief if you have it, and you may or may not have it, but at any rate you are not expressing a desire to be loved. So you are not putting yourself in a position to reason with your desire to be loved.

A preference is a more complicated example. We can take a preference to be a relation between two propositions, and we can

## Reasoning with preferences?

take that pair of propositions to be its content. Suppose you prefer walking to driving. We can take this as a preference for the proposition that you walk over the proposition that you drive. What could you say to yourself to express this preference? Evidently neither of the propositions that constitute its content. And to say that you prefer walking to driving does not express the preference either. At best it would be expressing the belief that you have the preference, if you happen to have that belief. Consequently, it seems you cannot reason with preferences. That is the difficulty.

The difficulty arises over reasoning with all mental states apart from beliefs. But there is a way to overcome it. We can revise our notion of the content of a mental state. Philosophers commonly assume that mental states of different types can have the same content, which they take to be a proposition. So you might have a belief that you are loved, or a desire to be loved, and either state would have as its content the proposition that you are loved. Either state has the same content, but in the two different cases you stand in a different relation to the content—a believing relation in one and a desiring relation in the other. In the complicated case of a preference, you stand in a preferring relation to a pair of propositions. That is the common view.

The alternative is to take the content of a mental state to be a proposition together with a mark of some sort, which marks the type of state it is.<sup>9</sup> In this way the differences in mental states can be absorbed into the contents of the states. For instance, if you believe you are loved, the content of your belief is the proposition that you are loved together with a belief mark. If you desire to be loved, the content of your desire is this proposition together with a desire mark.

How do we refer to these contents? I shall explain in a moment how we do so in English. But it will be clearer if I start with an artificial language. The language must have the resources to designate marks; I shall give the name ‘markers’ to the linguistic items that do this job. Let the marker for belief be ‘yes’ and the marker for desire be ‘nice’. If you believe you are loved, you might designate the content of your belief by the artificial sentence ‘I am loved—yes’. If I also believe you are loved, I have a belief with the same content as yours, but I would designate it using the second person sentence ‘You are loved—yes’. If you want to be loved, you

<sup>9</sup> Examples of this idea appear in Richard Hare’s *The Language of Morals* and Paul Grice’s *Aspects of Reason*.

## John Broome

might say 'I am loved—nice'. If I want you to be loved, I have a desire with the same content as yours. I might say 'You are loved—nice'.

A preference is again more complicated. If you prefer walking to driving, the content of your state is the pair of propositions that you walk and that you drive, together with a preference mark. You might designate it by the artificial sentence 'I walk—rather—I drive'.

If you say this sentence to yourself, you are expressing the preference, and you are also expressing the content of the preference. In this way, a mark gives a preference the special feature that a belief has: expressing the content of the preference is also expressing the preference itself. So, when you express the preference, you make its content available to be reasoned about. Preferences become available for reasoning with.

The purpose of marks is to distinguish between different sort of mental state. One sort of state can be distinguished by the absence of a mark, provided all the others have marks. It is convenient to give beliefs this special status. So from here on, I shall drop the 'yes' marker, and take the content of a belief to be a proposition without a mark.

Marks give us the beginning of an account of first-order reasoning with mental states other than beliefs. Your reasoning will be a process in which you express your mental states to yourself using marked sentences, operate on their contents, and emerge with a new mental state. But this is only the very beginning of an account. The next thing that needs to be done is to make the account realistic. If we are really to use marked sentences in our reasoning, we must have actual marked sentences in our language. Do we?

We do. Natural languages can express beliefs and their contents. They also contain devices that allow them to express many other mental states and their contents. If their contents are indeed propositions with marks, as I am assuming, some of these devices are what I called markers. English uses special constructions or special moods of verbs to serve as markers.

For example, a desire is marked by an optative construction. Robert Browning said 'Oh, to be in England now that April's there!'. This optative sentence designates the proposition that Browning is in England now in April, together with the mark for desire. When Browning said to himself 'Oh, to be in England now that April's there', he expressed his desire to be in England, and also the content of his desire, understood as a proposition with a

## Reasoning with preferences?

mark. Translated into my artificial language, he said ‘I am in England now that April’s there—nice’.

As Jonathan Dancy pointed out to me, English has a marker for preference too. The sentence ‘Rather walk than drive’ is the English equivalent of my artificial ‘I walk—rather—I drive’. It designates the pair of propositions that you walk and that you drive, with the mark for preference.

On the face of it, this construction puts you in a position to reason with your preferences. Suppose you prefer walking to driving and biking to walking, but you do not prefer biking to driving. You do not satisfy the requirement *Transitivity*. But you may say to yourself:

*Rather walk than drive*  
*Rather bike than walk*  
*So, rather bike than drive.*

When you say each of the first two sentences, you are expressing a preference you have. Saying these sentences to yourself causes you to have a new preference that you did not previously have. By the time you say the third sentence to yourself, you are also expressing this new preference. By causing you to have it, this process has brought you to satisfy *Transitivity*. Intuitively, this seems a plausible instance of reasoning with preferences.

The contents of your preferences are pairs of propositions, with marks attached. I can designate them using sentences in my artificial language. Since I am speaking of you, I shall put them in the second person. The contents are:

You walk—rather—you drive  
You bike—rather—you walk  
You bike—rather—you drive.

The process I have described satisfies the description of first-order reasoning that I gave in section 3. It is a process whereby some of your mental states give rise to another mental state; the mental states involved have contents; in the course of the reasoning you say to yourself the propositions that constitute these contents, and you reason about these contents. So on the face of it, this is a genuine example of first-order reasoning with preferences.

However, much more needs to be done to make that conclusion secure. For one thing, we need to generalize: are there similar processes that can bring you to satisfy other requirements on preferences? For another, can we find a criterion for correct

## John Broome

reasoning with preferences, as opposed to incorrect reasoning? Certainly, if this is to be genuine reasoning, there must be such a distinction.

### 7. Preferences and beliefs about betterness

But I think the most difficult challenge is to demonstrate that this is really reasoning with *preferences*. When you use a sentence like ‘Rather walk than drive’ you may well be expressing a belief about betterness, and not a preference—in this case, the belief that walking is better than driving. The betterness in question need not be absolute betterness from the point of view of the universe. It might be betterness for you, or betterness relative to your point of view, or something else.

If your sentences express beliefs rather than preferences, the contents of the reasoning I have described would be the sequence of propositions:

It is better that you walk than that you drive  
It is better that you bike than that you walk  
So it is better that you bike than that you drive

The process that proceeds by your expressing these propositions to yourself constitutes correct reasoning, because the betterness relation is transitive. If it is better that you walk than that you drive, and better that you bike than that you walk, it is better that you bike than that you drive. But this is theoretical reasoning with beliefs. It is not reasoning with preferences. Perhaps the pattern of reasoning I presented in section 6 is always theoretical reasoning; perhaps it is never reasoning with preferences, as I suggested.

What is the difference between a preference and a belief about betterness? Not very much, possibly. A belief about betterness may satisfy the definition of broad preference that I gave in section 4: a belief that *a* is better than *b* may be a mental state that would typically cause you to choose *a* were you to have a choice between *a* and *b* only. I explained that, to define preference in its ordinary sense, we would have to add conditions to this definition of broad preference. I explained that conditions are needed to separate a preference for *a* over *b* from an intention to choose *a* rather than *b*. It now emerges that we also need conditions to separate a preference for *a* over *b* from a belief that *a* is better than *b*. But these conditions will be hard to find. The functional role of a belief



## Reasoning with preferences?

about betterness may not be very different from the functional role of a preference; it will be hard to separate them.

A belief about betterness does differ from a preference in one respect. It is a state that has a content that is a proposition. The contents of beliefs, being propositions, stand in logical relations to each other. The logical relations among contents induce rational requirements on beliefs. An example is the requirement *Modus ponens*, which derives from the logical relation among propositions known as 'modus ponens'. Moreover, we have reasoning processes for beliefs that allow us to follow up these logical relations, and thereby bring ourselves to satisfy some of the rational requirements on beliefs. These facts are special to beliefs, and seem to separate them from preferences.

But we commonly think there are rational requirements on preferences too, and I have been assuming so in this paper. Moreover, I am now investigating the idea that we have reasoning process for preferences that allow us to bring ourselves to satisfy some of these requirements. If these things are true, it further reduces the functional difference between preferences and beliefs about betterness. Both are governed by rational requirements and, for both, these rational requirements can sometimes be satisfied by reasoning.

Furthermore, there is a case for thinking that the rational requirements on preferences, if they truly exist, derive from the logical relations among propositions about betterness. Why does rationality require your preferences to be transitive? I have mentioned the money-pump argument, but here is another possible explanation. Rationality requires you to prefer *a* to *b* if and only if you believe *a* is better than *b*. And rationality requires you to believe *a* is better than *c* if you believe *a* is better than *b* and *b* is better than *c*. And this is so in turn because, as a matter of logic, if *a* is better than *b*, and *b* is better than *c*, then *a* is better than *c*. I do not insist this is the correct explanation of the *Transitivity* requirement, but it is a plausible one.

The upshot is that it is hard to distinguish the functional roles of a preferences and a belief about goodness. This explains why many noncognitivists about value think that a belief about betterness is indeed nothing other than a preference. In so far as the two converge, I am inclined in the opposite direction: a preference may

## John Broome

be nothing other than a belief about goodness. It may turn out that reasoning with preferences is really nothing other than reasoning with beliefs.<sup>10</sup>

## References

- Bennett, Jonathan, 'Why is belief involuntary?', *Analysis*, 50 (1990), pp. 87–107.
- Bratman, Michael E., *Intention, Plans and Practical Reason*, Harvard University Press, 1987.
- Broome, John, 'Does rationality give us reasons?', *Philosophical Issues*, 15 (2005), pp. 321–37.
- Broome, John, 'Wide or narrow scope?', *Mind*, forthcoming.
- Carroll, Lewis, 'What the tortoise said to Achilles', *Mind*, 4 (1895), pp. 278–80.
- Hare, R. M., *The Language of Morals*, Oxford University Press, 1952.
- Grice, Paul, *Aspects of Reason*, edited by Richard Warner, Oxford University Press, 2001.
- Harman, Gilbert, *Change in View: Principles of Reasoning*, MIT Press, 1986.
- Kolodny, Niko, 'Why be rational?' *Mind*, 114 (2005), pp. 509–63.
- Pettit, Philip, 'Preference, deliberation and satisfaction', in this volume.
- Piller, Christian, 'Content-related and attitude-related reasons for preferences', in this volume.
- Rabinowicz, Wlodek, 'Money pump with foresight', in *Imperceptible Harms and Benefits*, edited by Michael J. Almeida, Kluwer 2000, pp. 123–154.
- Scanlon, T. M., *What We Owe to Each Other*, Harvard University Press, 1998.
- Williams, Bernard, 'Deciding to believe' in his *Problems of the Self*, Cambridge University Press, 1973, pp. 136–51.

<sup>10</sup> This paper has been greatly improved as a result of extremely helpful comments I received from Krister Bykvist and Serena Olsaretti.