

## Content-Related and Attitude-Related Reasons for Preferences

Christian Piller  
University of York  
[cjp7@york.ac.uk](mailto:cjp7@york.ac.uk)  
January 2005

ABSTRACT: In this paper I argue that we should not always prefer what is better; we rather should prefer that which is such that preferring it is better. Rational preferences depend not only on reasons grounded in features of what one prefers; they also depend on reasons relating to features of the preference itself, i.e. they depend on what I call 'attitude-related reasons'. I investigate arguments which aim to show that there are no attitude-related reasons for beliefs and I find that these arguments cannot be generalized to cover the case of preferences as well. Starting with an example which I borrow from Sen, I investigate the extent of attitude-related reasons for preferences and their importance. I criticize the strategy used by deniers of attitude-related reasons of trying to accommodate the normative force of features of preferences as reasons for trying to bring it about that one has these preferences, and I end by answering objections which have been raised by Gibbard and Parfit.

In the first section of this paper I draw, on a purely conceptual level, a distinction between two kinds of reasons: content-related and attitude-related reasons. The established view is that, in the case of the attitude of believing something, there are no attitude-related reasons. I look at some arguments intended to establish this claim in the second section with an eye on whether these arguments could be generalized to cover the case of preferences as well. In the third section I argue against such generalizations and present a case in favour of accepting attitude-related reasons for preferences. In the fourth section I present an objection to which I react in the fifth section where I try to strengthen my case for attitude-related reasons for preferences. Finally I discuss and reject criticisms raised by two opponents of the view defended here.

### I The Distinction

Think of any attitude we might have towards some proposition  $p$ , which, because we have some attitude towards it, I call *the content of this attitude*. It might be an epistemic attitude like belief or doubt, an attitude like desiring or intending, or an emotional attitude like anger or gratefulness. For all these attitudes we can have reasons that speak in favour of or against adopting them. On a conceptual level we can distinguish between two kinds of reasons. Content-related reasons are such that, if we are aware of them, they show us  $p$ , the content of an attitude, in a certain light. They point to some feature of  $p$ , which makes a certain attitude towards  $p$  the appropriate one to have. Attitude-related reasons, by contrast, are such that, if we are aware of them, they show us our having a certain attitude towards  $p$  in a certain light. They point to some feature of our having this attitude that makes it the appropriate one to have. Thus, reasons for adopting a certain attitude might be grounded in two different ways: either by their relation to some feature of the content of an attitude or by their relation to some feature of the attitude itself.

The following examples show that we can handle this distinction well enough. Take the attitude of believing first. The propositional content of my belief that  $p$  might be entailed by other things I believe with good reason. This feature of the proposition that  $p$ , namely being entailed by what I reasonably believe, makes it the case that I have good reason to believe that  $p$ . Because we are talking about a feature of the

proposition that p, the reason mentioned is a content-related reason. In contrast, believing that p might be beneficial in some way. This, if there are such things, would be an attitude-related reason for believing that p, as being beneficial is a property of believing that p. Similarly with emotions: being incapable of doing something most other people can do might well be a reason for me to be ashamed of myself. I am ashamed of myself because I have the property of being incapable of doing something others can do easily. My reason for feeling shame is content-related. If my being ashamed, in fact, prevents me from doing what others can do easily, then I have an attitude-related reason against feeling ashamed. My attitude of feeling ashamed has the property of preventing me from doing something the doing of which is important to me. Similarly for preferences: I might prefer apples to oranges on grounds of how apples and oranges taste. If so, this would be a content-related reason for this preference. Preferring home-grown products might be an aspect of being supportive of one's local community. If to support one's local community is a desirable feature, then, as I am not living in a country where oranges grow easily, I would have an attitude-related reason for preferring apples to oranges, as it is my preference for apples over oranges that has this desirable feature.

The distinction between content-related and attitude-related reasons has been explained by the distinction between features of an attitude's content and features of the attitude itself. This might not look like a solid enough foundation for the following reason. Instead of saying that the attitude A with content C has a feature F, we could almost equivalently say that C is such that the attitude A towards it is or would be F. For example: To call one's confidence in one's own abilities useful is almost the same as saying that one's abilities are such that belief in them is or would be useful. Thus, features of attitudes, which, according to our initial suggestion, provide attitude-related reasons, can, it seems, simply be re-described as features of the content of these attitudes and would then provide content-related reasons. The same argument also works the other way around. What looks like a content-related reason, namely that some content C has feature F, can be re-described as an attitude-related reason, namely as A being such that its being related to C would make it an A that relates to something that is F. According to this objection, it is simply arbitrary whether we pick out some feature as a feature of an attitude A or as a feature of its content C.<sup>1</sup>

Does the idea that we can describe the same situation – something's having some feature – with varying grammatical subjects really undermine the suggested distinction between two different kinds of reasons? On the one hand we have what looks like a purely grammatical point; on the other hand we have examples that indicate an ability to draw a plausible distinction. It certainly would come as a surprise if it turned out that there couldn't be this ability to distinguish, which we are confident to have.

The objection points to a possibility of re-describing features. We can understand what, intuitively, is a feature of an attitude as a feature of the attitude's content, but we also need not do so. We can understand what, intuitively, is a content's feature as a feature of a thus directed attitude, but, again, we need not. This is all that is needed to answer the objection. Features of the content of an attitude, I said, provide content-related reasons. What this amounts to is that we can describe the features of this content without mentioning our attitude towards it. This provides us with a real contrast to attitude-related reasons. If we cannot describe the content's feature without

referring to the attitude we have towards this very content, then we are dealing with an attitude-related reason. Consider the following example: the usefulness of one's belief in one's own abilities would, intuitively, be an attitude-related reason for this belief. One might well say that one's abilities are such that belief in them would be useful. But even if, grammatically, one describes the reason-giving feature as an aspect of the belief's content, one nevertheless cannot refer to it as a feature of the content without mentioning the attitude one has towards it. This makes it an attitude-related reason for believing in one's abilities.<sup>2</sup>

Note that facts concerning people's attitudes, including my own, can be content-related reasons for attitudes. Only if the reason for an attitude cannot be described without mentioning this very attitude, do we have an attitude-related reason for this attitude. Liking him because he likes my liking of football is for me a content-related reason for liking him: From my perspective it shows him to be appreciative of others having appropriate interests and as such specifies a feature of him, which I take to be a reason. The feature which is a reason, his liking of people interested in football, can be described with mentioning the attitude it is a reason for, which is my liking him. If, however, I would like him because he would like my liking of him, then, assuming that I have good reason to please him, I would have an attitude-related reason for liking him. The feature which gives me reason to like him, namely his liking of my liking him, cannot be described without referring to what the feature is a reason for, i.e. without referring to my liking him.

Another worry about the concept of an attitude-related reason runs as follows: Attitude-related reasons, we have learnt, make essential reference to a feature of the attitude. Think about the following feature: Believing that p might be such that it has the weight of reasons on its side. This is certainly a feature that we can only identify with reference to the very attitude of believing that p. Nevertheless it seems wrong to call it an attitude-related reason. Similarly, any principle concerning reasons for belief will have to mention the belief or the kind of belief for which reasons are specified. It would, however, completely undermine our distinction if all principles of reasons could only ever specify attitude-related reasons.

Let me answer this worry: The fact that the evidence at hand best supports a belief is not an attitude-related reason because it is not a reason at all. It rather is a statement about reasons and about where their force lies. Similarly, a statement that specifies that something is a reason for something else is not itself a reason but a statement about a reason. What is the reason for believing that there is a fire? It is the smoke or the fact that I notice the smoke. The fact that noticing the smoke is a reason to believe there is a fire – this whole epistemological fact – is not itself a reason to believe that there is a fire. As we can obviously refer to what is the reason, namely my noticing the smoke, without having to refer to what it is a reason for, noticing the smoke is a content-related reason for believing that there is a fire.<sup>3</sup>

One last remark on the distinction between attitude-related and content-related reasons: A reason is always a reason for something. If we vary what the reason is a reason for, then what was an attitude-related reason for something can become a content-related reason for something else. The usefulness of my believing that p is, if it is a reason for believing at all, an attitude-related reason for believing that p. Suppose we can intend to believe that p. If so, the usefulness of the belief that p

would be a feature of what the intention is about and, thus, it would be a content-related reason for intending to believe that p. If, however, it would be a waste of time to intend to believe that p, then this feature of this intention would be an attitude-related reason against intending to believe that p, which in turn could be a content-related reason for wanting not to intend to believe that p.

The distinction between content-related and attitude-related reason is supposed to be completely general. It applies to all attitudes for which there could be reasons and, furthermore, it is supposed not to exclude any substantial view about what is a reason for what. Having clarified the concept of an attitude-related reason, we can now turn to the more interesting question: are there such reasons?

## II Attitude-Related Reasons for Beliefs

Most people think there would be something odd about accepting attitude-related reasons for beliefs. The fact that it would be beneficial for me to believe something, in general, does not, it seems, bear on the question whether my belief is held rationally or on whether it is justifiable. Only facts that somehow bear on its truth, i.e. facts that primarily have to do with the content of the belief, can play a justificatory role. What explains this common resistance to accepting attitude-related reasons for beliefs?

A salient fact about beliefs is that often we are unable to adopt a belief by simply deciding to adopt it. 'Can we, by any effort of our will, or by any strength of wish that it were true, believe ourselves well and about when we are roaring with rheumatism in bed [...]?' asks William James and he answers, 'we can say such a thing [namely that we are well] but we are absolutely impotent to believe [it].' (James 1897, 5) No one disagrees with James about his example. But there are other cases. Can't I decide to close my investigation into a matter and settle on the view most plausible at the moment, and isn't that deciding to adopt a certain view of the matter? For example, can't I decide to believe the student who assures me that he was unwell (maybe roaring with rheumatism in bed) and thus grant him an extension to his deadline (cf. Harman 1997)? And can't I decide that some matter is so serious that my usual standards for sufficient evidence are too lax, and, therefore, decide not to believe anything (cf. Nozick 1993, 86f.)? Taking these cases at face value, we should set limits to when we cannot decide to believe. The limit is easily drawn: We cannot decide to believe something in the face of strong reasons to the contrary. In short, we cannot decide to believe against our reasons.<sup>4</sup>

How does our inability to decide to believe undermine the idea that there could be attitude-related reasons for beliefs? The answer already lies in what has been said. We agreed that we couldn't decide to believe that we are well when obviously ill. Why not? Because we cannot decide to believe something if we have strong reasons to the contrary, as belief is an attitude governed by reasons. The existence of attitude-related reasons would undermine the correctness of this explanation.

Decisions are guided by reasons that have to do with the value or desirability (however we conceive of these things) of what is brought about by these decisions. Attitude-related reasons are usually of this very kind, they mention desirable or useful features of attitudes. If there were attitude-related reasons for believing something, then our inability to decide to believe that we are well, when ill, could not be

explained by the fact that we cannot decide to believe against reason. After all, the attitude-related reasons for believing to be well could be quite strong. If there were strong enough attitude-related reason for believing oneself to be well, it would not be against reason to hold this belief. Thus, we had lost our explanation of our inability to decide to believe in this case. This inability, however, needs an explanation. Regardless of how strong we imagine these attitude-related reasons to be, it will still hold that we cannot acquire the beneficial belief simply by decision. Thus, assuming that the correct explanation of the inability to decide to believe, when present, has to do with the fact that belief is governed by reasons, there cannot be attitude-related reasons for beliefs. Their existence would undermine the correct explanation of an obvious fact. For future reference I call this 'the undermining argument'.

There might be deeper reaching explanations of why there are no attitude-related reasons for beliefs. Believing, it is often said, is an attitude that has truth as its constitutive aim. David Velleman, for example, presents such a position as follows: 'We believe a proposition when we regard it as true for the sake of thereby getting the truth right with respect to that proposition: to believe something is to accept it with the aim of doing so only if it really is true.' (Velleman 1996, 183) It is the aim of getting things right, which distinguishes believing from other attitudes like assuming. The fact that it would be nice to have a certain belief need not be indicative of its truth. Thus, when we are trying to come up with a belief about a subject matter, considerations of value, insofar as value isn't truth-related, don't count. These considerations don't count because if they did, we would not be anymore in the business of coming up with a belief.

Velleman allows that someone might not be interested in getting things right; one might just be indifferent to truth (and interested in believing that *p* regardless of whether *p* is true or not) or one might want, for some reason, to believe what is wrong, whatever it is. Thus, it need not be the agent who aims at getting things right. But what things besides agents have aims? One aspect of Velleman's view is that attitudes or mechanisms which produce such attitudes have their own aims.

Here, however, I am not really concerned with the details of Velleman's argument. I use it as a contrast to the previous argument. In both arguments a feature of beliefs, namely that we can't believe against reason in the first argument and that belief aims at truth in Velleman's argument, is employed to argue against attitude-related reasons for beliefs. The first argument, however, is set on a more general level; it mentions reasons but not truth. Thus, it seems to invite us to draw a more general lesson. We can't just believe at will, because belief is guided by reasons. Similarly, we can't just desire at will, because desire is also guided by reason, and we can't just intend to do something at will, because intentions are also guided by reasons. The first argument, it seems, is generalisable to all attitudes. If so, it would support a denial of attitude-related reasons for all attitudes. David Owens endorses such a generalization in the following passage: 'I have already established that beliefs are not subject to the will because a belief could not be justified simply by reference to the desirable consequences of having the belief. Something similar is true of practical decisions: what makes a decision rational is not the advantages of the decision itself but those of the action decided upon' (Owens 2000, 81).

Let me try to support for a moment such a generalization. Attitude-related reasons look odd because they might threaten a certain form of realism about reasons. By realism I mean that there are conditions of appropriateness and/or correctness for all attitudes, which are independent of how the fulfilment of these conditions affects us. A belief is justified if what I believe is well supported by my evidence. Whether I like believing what I believe or not is irrelevant for a belief's normative status. The same holds, according to realism thus understood, for all attitudes for which there can be reasons. We can give conditions of appropriateness and/or correctness and our liking of the resulting attitudes should, again, not change their normative status. For example, my preference for A over B is correct if and only if A is indeed better than B (however we understand betterness; we could well understand it in the subjective terms of decision theory). My intention to fi is appropriate if and only if fi-ing is the best option (again, however, we understand bestness). And similarly for emotions: 'It is irrational [or inappropriate],' Allan Gibbard writes, 'say, to be angry at the messenger who brings bad tidings, but rational to be angry at the miscreant who deliberately wrongs one.' (Gibbard 1990, 36) Whether my anger at either one of them is useful or harmful does not affect the conditions of appropriateness indicated by this example.

One more piece has to be added to fully describe what could well be called the Standard View, according to which there are no attitude-related reasons. A defender of attitude-related reasons might have argued as follows: harmfulness is often a reason against doing something. Why, if some attitude is harmful, should it not be a reason against holding this attitude? The Standard Reply accepts the normative force of harmfulness but relocates it. It is not a reason against holding the harmful attitude; it rather is a reason for trying to prevent holding the harmful attitude.

It first looked like a normative conflict, when the content-related reasons speak in favour of some attitude, but the harmfulness of holding this attitude speaks against it. According to the Standard View it is not a case of conflict after all. There are only content-related reasons for attitudes. What looked like attitude-related reasons are in fact reasons for something else, namely for trying to bring it about or to prevent holding some attitude. On this view it might well be rational to do something the expected result of which is an irrational attitude. The rationality of our attempts to bring an attitude about is simply irrelevant to the rationality of the attitude itself. Believing or wanting and bringing it about that one believes or wants something are two separate things, governed by different groups of reasons.<sup>5</sup>

In this section I considered why attitude-related reasons for beliefs look odd. My central concern, however, is whether there are attitude-related reasons for preferences, and three points have emerged that bear on this issue. First, we can try to generalize the idea that we cannot simply decide to believe because belief is governed by reasons. Preferences are governed by reason as well. If, for that reasons, we can't simply decide what to prefer, we might have an analogous argument against attitude-related reasons for preferences. Secondly, a realist understanding of the conditions of appropriateness and/or correctness of attitudes might extend to preferences. A preference, we might want to say, is fitting if and only if it is a preference for what actually is better. Whether we like having this preference might well be irrelevant for such an assessment. Thirdly, the distinction between reasons for believing something and reasons for trying to bring it about that one believes something can be applied,

with the same plausibility to preferences. Thus, the normative force of the considerations that a defender of attitude-related reasons for preferences will appeal to, for example that having a preference is useful, can be accepted by the Standard View, but its relocation to the realm of attempts to influence one's preference structure renders the acceptance of attitude-related reasons for preferences spurious.

### III: Attitude-Related Reason for Preferences

Think of the following three choice situations.<sup>6</sup> In all of them your host offers you one of two fruits: you choose first and he takes the remaining one. If the choice is between a big apple and an orange, you take the big apple. Apples and oranges are pretty much on a par for you, apples just having a small advantage. If the choice is between an orange and a small apple, you take the orange. The small loss in quality of enjoyment is outweighed by the larger quantity of enjoyment the orange offers you. When the choice is between a big and a small apple, you take the small apple. Why? You don't want to leave the host with what you can reasonably expect to be his less preferred option. Doing so would be impolite. As you don't know (and are not expected to know) whether your host prefers oranges to apples, you can in those situations choose the fruit you like best. If the choice is between apples, however, you leave the fruit you expect your host to like more, and thereby choose the fruit you like less.

What does this example teach us? When we think about what to choose, we should not only think about the features of what we choose but also of the features of our so choosing. Tasting nice is a feature of what we choose; being polite is a feature of our choosing.

If it is legitimate to apply the distinction between content-related and attitude-related reasons to choice, then the apple example shows us that there are attitude-related reasons for choices. Politeness is a feature of choices, not of apples. Even if we can say that the apple is such that choosing it is polite, we have to refer to our choosing it when we describe this property of the apple, which, in accordance with our discussion in section 1, makes it an attitude-related reason for choosing the small apple.

Accepting attitude-related reasons for choices does not encounter the following problem which we would have to face if we accepted attitude-related reasons for beliefs. In the case of belief, when attitude-related reason and content related reasons pull in different directions, we would have no clue how to weigh degrees of evidential support against something like the potential harmfulness of believing something. But in the case of choices such a weighing of different kinds of reasons poses no special difficulties. It is very much like weighing reasons of one and the same kind. In our example the attitude-related reason of politeness trumps the content-related reason which points to a difference in enjoyment. However, if we strengthened the content-related reasons we might get a different result. If the meal had been rather insubstantial and the small apple on offer was really very small, the balance might shift. It would shift, if the small apple had been treated with some chemical agent you but not your host are allergic to.

When I discussed the question why we are not inclined to accept attitude-related reasons for beliefs, three considerations emerged which one could try to use against the claim that preferences allow for attitude-related reasons. But when we try to

undermine the force of the example above we realize that none of them works. The first argument doesn't get off the ground as, in contrast to the case of belief, we can decide which apple to choose. According to the second argument, the acceptance of attitude-related reasons would threaten a realist understanding of them. The correctness condition for choice suggested above told us that a choice is correct if and only if what is chosen is best. A polite person, however, chooses what is worse, and, it seems, correctly so. Thus, this correctness condition needs to be changed in the light of our example: Choosing some G is correct if and only if the choosing of G is best. Such a principle is meant to make room for both content-related and attitude-related reasons. They are what determine whether our choosing is correct. Whether this new correctness condition is satisfied or not is sufficiently independent of our attitudes towards the correct choice as these attitudes already partially determine which choice satisfies the condition. Thus, attitude-related reasons pose no threat to realism about reasons. What about the idea to separate the domains to which content-related and attitude-related reasons apply? It would be absurd to apply this idea to the case at hand, as we would have to choose the big apple but we would also have to try to bring it about that we choose the small apple. To act politely, on this view, would only be possible if we succeeded in making ourselves choose irrationally. Not even rude people will accept this view.

So far I have talked about choices. Can we simply retell the above example in terms of preferences?<sup>7</sup> Just thinking about the fruit on offer, I would prefer the big apple. But when thinking about what to prefer I have to think not only about features of the object of my preference but also about features of my preferring. Politeness goes deeper than choice: A truly polite person doesn't have to grind her teeth when doing what is polite. The polite person, on the strength of her attitude-related reasons, does indeed prefer the small apple. She does so because preferring the small apple has an attractive feature: to prefer the small apple is to be polite.

If this picture is correct, one must wonder how attitude-related reasons for preferences could ever have been in dispute. Politeness is just one minor virtue, but what we find there seems to be an important aspect of any virtue. Kindness, for example, may demand something that, in a way, inconveniences oneself. But if one really is a kind person, one does what is kind wholeheartedly. After all, the inconvenience relates only to what one chooses or prefers and will be outweighed by the features of so choosing or preferring. In general, we care not only about what we get, we also care about who we are. Who we are is revealed in the choices we make and in the preferences, which usually are thus expressed. If it matters to us who we are, it matters to us which preferences we have. And if it matters to us which preferences to have, the door has been opened for attitude-related reasons.

What role does this defence of attitude-related reasons for preferences assign to second-order desires? Preferring the small apple has a desirable feature. It also has an undesirable feature, as it is a preference for what is worse. But as politeness is more important for one than the difference in the amount of enjoyment, one will overall want to have a preference for preferring the small apple. In a case like this, however, this second-order preference plays no important normative role. A rational agent is someone who will correctly respond to the force of her reasons. Thus, a rational agent will prefer the small apple on the strength of her attitude-related reasons for this preference. The second-order desire is simply an endorsement of a preference, which



is the right one to have for independent reasons. Consequently, conflicts between second-order desires and first-order desires will raise the suspicion that a thus conflicted agent has not fully aligned his preferences with his reasons. Suppose I prefer the big apple, and prefer to prefer the small apple. There are two ways of describing this case. First, somewhat childishly, I cannot resist the promise of greater enjoyment. Thus, I prefer the big apple. But on reflection, I think that politeness should carry more weight. Thus, I prefer to prefer the small apple. My judgment that politeness is more important is obviously a criticism of my first-order preference. I don't prefer as, in my own view, I have most reason to prefer. Secondly, I can't free myself from the influence of conventional rules, though I doubt their normative significance. Thus, I do prefer to prefer the small apple. But, somehow fortunately, this preference remains motivationally idle. I prefer the big apple in spite of my opposed second-order preference. I think that politeness in this context is of negligible importance and so my second-order preference is not as, in my view, it ought to be.

In the apple case and in cases relevantly similar, second-order desires don't play an independent normative role. Conflicts between second-order and first-order desires in such cases are indicative of a failure of having all one's preferences aligned with one's reasons. In other cases however – and I will introduce them later on – second-order preferences do play a significant normative role. This happens when first-order preferences leave matters underdetermined.

My defence of attitude-related reasons for preferences bears strong similarities to a view prominent in Amartya Sen's work. A crucial thought in this respect is the following: 'Rather than expressing moral views in terms of one ordering of outcomes, it may be necessary to express them through a ranking of the possible orderings of outcomes' (Sen 1974, 62).<sup>8</sup> In the terms of our example this would imply that the importance of politeness is captured by my preference for preferring the small apple, i.e. I rank the ranking which has the small apple on top higher than the alternative ranking. I regard this second-order preference as an endorsement of the result of having correctly weighed attitude-related and content-related reasons. Thus, I agree with Sen's claim that meta-rankings add needed explanatory power to preference theory, I only add a further explanation: these second-order preferences reflect the force of attitude-related reasons.

Here is another example, in which we meet attitude-related reasons:<sup>9</sup> One player has a dominant strategy. Playing it would secure him a small advantage over playing the other strategy. If he does so, the other player, however, will have no chance to receive a comparably good outcome. Many people play the dominated strategy. I suggest they do so because the payoff of the game is not the only relevant consideration. Interacting in a way that leaves one's partner with considerably less than he could have had had one made a small sacrifice, is being greedy or being nasty. But people commonly want to be nice. To prefer the dominated strategy is to be nice and, thus, to be the kind of person one usually wants to be. This explanation does not primarily appeal to a feeling of sympathy with the opponent: it is not that the amount the player receives is valued less by him if the other player fails to achieve a good result. Thus the effect should be preserved in conditions of increasing anonymity. If one values being nice, and thus takes attitude-related reasons into account, sympathy is not needed to influence the way one behaves.

Such an account fits well with Sen's distinction between sympathy and what he calls 'commitments': 'Sympathy is, in some ways, an easier concept to analyse than commitment. When a person's sense of well-being is psychologically dependent on someone else's welfare, it is a case of sympathy; other things given, the awareness of the increase in the welfare of the other person then makes this person directly better off [...]. On the other hand, commitment does involve, in a very real sense, counterpreferential choice, destroying the crucial assumption that a chosen alternative must be better than (or at least as good as) the others for the person choosing it [...]. Commitment is, of course, closely connected with one's morals' (Sen 1977, the three quotes come from 327, 328, 329). Sen's aim in this article is to broaden the framework of economic analysis. He wants to make room for commitments. In my view, he thereby endorses attitude-related reasons. Is there a sense in which acting on one's commitments is 'counterpreferential'? One acts against preferences one had had, if one would have only thought about content-related reasons. This is not an odd counterfactual, as thinking about issues in isolation from their social dimension might well be a natural step in the process of deliberation. Thus, the acceptance of attitude-related reasons offers both an account of the nature of commitments and a justification of including them in the analysis of rational behaviour.

#### IV An Opposed Analysis

The example by which I tried to illustrate the force of attitude-related reasons moves from polite choices to polite preferences. We choose the small apple, because we prefer it, and we prefer it because preferring it is to be polite. An opposed analysis of polite behaviour – an analysis that does not introduce attitude-related reasons – is, however, possible.<sup>10</sup>

The opposed analysis starts by inviting us to specify the objects of our preferences in more detail. In the context of our example we should distinguish between two distinct preferences involving the two apples: one preference is for taking the small apple, the other is for receiving the big apple. Politeness, let us plausibly assume, comes only into play when we consider what to do, not when we consider what should happen to us. Thus, although we do not prefer to receive the small apple, we do prefer to take it. The more detailed description of the objects of preferences renders this set of preferences unproblematic. They are simply preferences for different objects and, therefore, content-related reasons alone can account for them. We can think of these objects as bundles of goods: one bundle consists of a small apple and being polite, the other bundle consists of a big apple and being rude. Our preferences are determined by our evaluations of these bundles. What we prefer, and not any features of our preferring, fully explains our preferences.<sup>11</sup>

When I described the apple example on the level of choices, I said that the set of relevant considerations is not exhausted by features of what one chooses but that it extends to features of one's so choosing. The opposed analysis will agree that features of one's choosing are normatively significant, but in order not to admit attitude-related reasons on the level of choice, it will embrace the following argumentative move, a move familiar from expositions of consequentialism: having acted in a certain way is one of the consequences of so acting. When we choose the small apple we also choose to act politely, i.e. we also choose to make a certain (polite) choice. In general, choosing A is itself one of the things that we choose when we choose A. This

self-referentiality of choice – choosing A is always also a choosing of the choosing of A – explains the success of the opposed analysis. Content-related reasons are sufficient because one's choosing something is supposed to be a part of what one chooses.<sup>12</sup>

According to my argument of the previous section, a choice of A is explained by a preference for A. If this is correct, the agent will usually also have a preference for choosing A. This preference gives the opposed analysis its starting point, because the attitude-related reason for preferring A will usually be mirrored in an analogous feature of choosing A. Such is the case in the apple example: both preferring the apple as well as preferring to choose it manifests politeness.

The search for an undisputed case of attitude-related reason, one in which above absorption of their force is not feasible, will thus have to follow one of two routes: either we consider preferences that are not preferences for choosing things or we consider cases in which the feature of the preference which grounds the attitude-related reason will not be mirrored in a feature of the corresponding choice. I will pursue both lines in the next section.

#### V More Attitude-Related Reasons for Preferences

In section III I have argued that second-order desires are not normatively significant in the apple example. In cases of indifference, however, they are normatively significant. Here is what Harry Frankfurt says about the usefulness of having final ends:

'Without ends, there are no means. And if no activity serves as a means, then no activity is useful. Thus, having a final end is a condition of engaging in useful activity. Now the fact that an activity is useful endows it with meaning. Suppose that we never acted in order to attain or to accomplish something which we regarded as desirable. Suppose, in other words, that we never did anything that we believed to be useful. In that case, our activity would appear to us to serve no purpose. We would find it empty and vain, for it would seem to have no point. It would be, to our minds, altogether meaningless. A life constituted entirely by activity of that sort would be, in an important sense, a meaningless life. Life cannot be meaningful in this sense, then, without final ends' (Frankfurt 1992, 84f.)

The reason offered for wanting something for its own sake, for having a final end, is not connected with what it is that is our final end. Thus, wanting something for its own sake is not simply a matter of getting one's wants in line with the perceived value of the states which our activities might accomplish. The reason for wanting something for its own sake is that wanting something for its own sake is useful in a fundamental sense: without such wants life couldn't be meaningful. Note the structure of Frankfurt's claim: Some attitudes have desirable features. This is a reason, an attitude-related reason, for having them.

For reasons related to those mentioned by Frankfurt, we encourage young people to engage in various activities so that they might find things that really interest them. The case that he describes, however, a case of complete indifference, of not caring about anything, is something alien to most of us. Localized indifference is a much

more common phenomenon and gives rise to examples of attitude-related reasons structurally analogous to Frankfurt's case. I was born close to Graz and, for reasons unimportant here it was always Sturm Graz and not GAK, the other local team, who had my allegiance. Supporting a team makes football more interesting. Living in England I watch English football and want to support a team, not primarily because of the 'value' of the team in football terms but because supporting a team makes football more interesting. This attitude-related reason doesn't solve my problem just on its own. (The same is obviously true in Frankfurt's case as well). It doesn't decide which team to support. It leaves me with wanting to want some team to win. And this second-order desire is normatively significant. It sets me the task of finding a team to support. I will look for content-related reasons to prefer one team to another. Content-related reasons that by themselves, i.e. without the task set by the second-order desire would not sway me one way or the other. This project need not be successful. If allegiance to a team is very important to oneself, then one could have an attitude-related reason against forming any more or less arbitrary allegiances. This attitude-related reason could well outweigh the attitude-related reason tied to enjoyment.<sup>13</sup>

Frankfurt is certainly right in that we need to care about something. Being involved with the world is a necessary condition for leading a good life. But here as elsewhere one has to strike the right balance. There can be too much interest and too much engagement which would put a strain on our limited resources and, to mention a famous Epicurean thought, would make us vulnerable to a multitude of disappointments. Frankfurt's point and its Epicurean flipside highlight important features of preferences which are not reducible to features of their objects and, thus, introduce attitude-related reasons for preferences on which the opposed analysis is silent.

We can illustrate Frankfurt's minimal point – features of attitudes matter – in other ways. Preferences can be socially unacceptable. They can make it difficult for us to fit in and gain the respect of the people around us. One might be drawn to what others find appalling or at least sufficiently odd to withdraw from normal levels of interaction. If one is attracted by death, and the dead cat still lies in the shed so that one can study the natural processes of decomposition, one won't be invited for tea at one's neighbour's house. Not so much because of what one does, rather because of the creepy character one is. If one regrets such social exclusion, the weighing of attitude-related and content-related reasons might well make one want to change one's preferences in order to become socially more acceptable.<sup>14</sup>

At the end of the last section I said that there are two areas which escape the opposed analysis: it cannot be applied if a preference is not a preference for choosing something and it fails if the benefit one gains from preferring something is not matched by a benefit of choosing it. I now turn to the latter case.

We are looking for examples in which preferences can have desirable features that cannot be described as features of choosing what the preference is a preference for. Let me start with the attempted generalization of the 'undermining argument'. Remember, we cannot choose what to believe because belief is governed by reasons. Along the same lines it is argued that we cannot simply choose to prefer something. For example: when offered a choice between a saucer of mud and a pot of gold, one cannot simply choose to prefer the saucer of mud. The suggested explanation is the

same as in the belief case: such a preference would be crazy. In the belief case, however, we had a further datum: whatever the size of the reward offered, we simply cannot decide to believe to be well, when ill. Does the same hold in case of preferences?

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'. Would there be any reason to doubt my honesty, when I say this? I think not. In a 'normal' setting in which we choose between mud and gold it would be crazy to ask for the mud but, given the offer, it strikes me as very reasonable to do so. 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. This reaction does not oppose but arises from the view that preferences are constrained by reasons. They are constraint by content-related and by attitude-related reasons.<sup>15</sup>

We cannot apply the opposing analysis to this example. In the apple case the desirable feature of preferring the small apple, namely to be polite, could also be understood as a feature of choosing the small apple. But the desirable feature of preferring the saucer of mud (or, indeed, of preferring to choose the saucer of mud), namely to be rewarded by two pots of gold, is not a feature of choosing the saucer of mud. If it were we simply would have a different example, one in which you choose between two bundles of goods: a saucer of mud and two pots of gold on the one hand and one pot of gold on the other. The example as told here has it that preference but not choice is rewarded.

I claim that we are able to prefer what is worse because doing so is better for us. This turns the undermining argument upside down. Whereas in the belief case attitude-related reasons would undermine what looks like the correct explanation of an inability, in the case of preferences we need attitude-related reasons to explain our ability to prefer what is worse.

According to the Standard View, the view I am arguing against here, it is one thing to prefer something and quite another thing to try to bring it about that one prefers something. In the case under discussion we should prefer the pot of gold but we should also try to bring it about that we prefer the saucer of mud. We should, in other words, try to make ourselves irrational.

Because beliefs and preferences are reason-guided, making ourselves irrational is always a complex project for rational people like us. But is it really plausible to assume, as the Standard View entails, that I could only get the two pots of gold if I consulted the hypnotist who could make me prefer the saucer of mud? Think of the Toxin Puzzle where intending to act but not acting is rewarded. This puzzle essentially relies on the future directedness of the intention. How can I form an intention to do something tomorrow when I know now that tomorrow I won't have any reason for doing it? If we collapse the timely split between forming the intention and acting, the puzzle disappears. The attitude-related reason, however, keeps its force. My wanting to drink the toxin now has a desirable feature that is not a feature of my drinking or choosing to drink the toxin. I will get the reward as long as I want to drink it now, whether I drink it or not. Having the preference indicated by all my reasons, I would ask for the toxin and drink it. Nothing more could be asked for to

give me the reward. I have no need for a hypnotist; I come to want to drink the toxin by a normal process of reasoning that respects attitude-related reasons.<sup>16</sup>

Why should we follow the Standard View and separate the domains to which the two kinds of reasons apply? The Standard View owes an answer to this question. A good answer would be that separation is needed because the reasons to be separated derive their normative force from different sources. This answer is defensible in the case of beliefs, because the reasons provided by the evidence and the reasons provided by the potential benefit that believing carries with it look nonnegotiable. Trying to bring about a belief that is neither in your power to make true nor supported by the evidence can only be a reasonable project if the benefits outweigh the costs. How strong the epistemic reasons are against whose force we want to believe will only be relevant if this would make a difference to the costs involved. If we had the believe-that-p pill, the strength of epistemic reasons we want to counteract would be irrelevant. This incommensurability does suggest that epistemic and practical reasons have different sources and, as I said, this lends independent support to the separation of domains to which these reasons apply. The benefits from preferring something, however, are clearly commensurable with the benefits from what we prefer. In our example it is all about gold. Thus, why should we distinguish reasons for trying to bring it about that one prefers something from reasons for preferring it, if these reasons come from the same normative force?

We could argue for such a separation by distinguishing different addressees of norms of reason. I am one of the addressees: I ought to prevent my having a preference for the pot of gold. But who or what is the addressee of the norm that gold is preferred to mud? Is it the preference itself, so that it ought to exist as my preference? Isn't it much more plausible to claim that I am the addressee of all norms, be they about what to believe, what to prefer or how to act?

The Standard View leaves us with a seriously strained picture of practical reason. I ought to prefer the pot of gold to the saucer of mud and I ought to prevent myself from having this very preference. If I ought to prefer gold to mud, then, in following this norm, should I not try to bring about that I prefer gold to mud? Similarly, if I ought to bring it about that I prefer mud to gold, I ought to do so not because of any aspects of my trying but because of the normative significance of the result of my trying. The Standard View leads practical rationality to the edge of inconsistency. Think about how implausible the analogous claim in the theoretical domain would look like. Epistemic reasons make it rational to believe that p. But you are also asked to believe that not-p, i.e. to bring this about, on purely epistemic grounds. This does sound as if both the belief in p and the belief in its negation are epistemically required. I conclude that the Standard View can only be plausible if it is underpinned by a separation of normative force. Thus, it can only be applied in a case in which we are confronted with an apparent conflict between practical and epistemic reasons for believing something. It cannot introduce a useful distinction between preferring something and trying to bring it about that one prefers, because the reasons for both cases stem from the same normative force.

Let me summarize my position: The norms of practical reason apply to me. Practical normativity is not divided in itself. The acceptance of attitude-related reasons is

necessary to endorse the idea of a coherent and unified notion of practical normativity.

But isn't there a price to pay? Doesn't the acceptance of attitude-related reasons threaten realism, the view that there are appropriateness conditions for preferences which are unaffected by whether we like preferences that fulfil these conditions or not? There is no reason for such a suspicion, as whatever the right substantial theory of reasons is, it will simply apply to attitude-related reasons as well. Some preferences have features that are reason-giving. It might be that preferring something is being polite or it might be useful as it promises some reward. Allowing attitude-related reasons extends the domain of reasons, but it doesn't change any of the principles of reason offered by substantial theories.

There is a further consequence. We have replaced one correctness condition for preferences – a preference is correct if and only if what is preferred is best – with another – a preference is correct if and only if so preferring is best. The original correctness condition invites the following appropriateness condition for preferences: A preference is appropriate or rational if and only if one reasonably believes that what one prefers is best. Thus, we could have a thoroughly cognitivist account of rational preferences, one in which rational preferences necessarily coincide with beliefs.<sup>17</sup> The same possibility still holds after we have allowed attitude-related reasons for preferences. A preference is appropriate or rational if and only if one reasonably believes that preferring in this way is best. This, I think, is a remarkable result: even if rational preferences are thus connected with beliefs, and even if there couldn't be attitude-related reasons for beliefs, there could still be attitude-related reasons for preferences. Their force is simply absorbed in the content of the belief which determines the rational preference. Thus, there really is no price to pay for accepting attitude-related reasons for preferences.<sup>18</sup>

## VI Against the Critics

Finally let me speak directly to two prominent opponents of the view outlined here. Pursuing the generalization of the inability argument discussed above, Alan Gibbard writes: 'Now preferences and intentions are not themselves voluntary. In the case of preferences, that is clear enough. I might, for instance, be convinced that I will be happy if and only if I cease to want to be happy, but I cannot on that account stop wanting to be happy' (Gibbard 1991, 39).

I have a content-related reason to want to be happy – being happy is very good. I have an attitude-related reason not to want to be happy – wanting to be happy prevents me from being happy. Let me introduce the following distinction: The force of attitude-related reasons might be genuine or borrowed. I like apples for their taste. Suppose I like my liking of apples for the following reason: liking apples makes it more likely that I stick with my apple-eating habit, even when obstacles emerge, e.g. when eating apples becomes socially less acceptable. The force of this attitude-related reason, which is reflected in my liking to like apples, is derived or borrowed from the value of apples. I endorse the apple-eating habit because I like (for good reasons) eating apples. If the discovery of the superior taste and the higher health benefits of Kiwis extinguishes my liking of apples, my liking of my liking of apples will go at the same time. Let me contrast this case with one in which my attitude-related reasons for liking

to eat apples does not depend on the value assigned to eating apples. I might like to like eating apples because I like to like home grown products. In this case, where the normative force of the attitude-related reason does not depend on those qualities of apples that make me like them in the first place, the discovery of Kiwis will result in a conflict that needs to be resolved: Is the increase of enjoyment experienced when eating Kiwis and their higher health benefit reason enough to undermine my allegiance to the local farming community?

Having distinguished between derived and genuine force of attitude-related reasons, we can turn to Gibbard's example. I don't think it shows that we couldn't want something for attitude-related reasons, reflected in our wanting to want it. The difficulty it presents is much more specific: the reason for wanting not to want to be happy derives its normative force from happiness. Thus, one only has reason not to want to be happy anymore, as long as one wants to be happy. There is something paradoxical about engaging in a project that, if successful, would have undermined its own rationale. But this is by no means a general feature of wanting something because one wants to want it. The peculiar difficulty in the project Gibbard describes is no reason for a general denial of attitude-related reasons for preferences. How we deal with attitude-related reasons has been sufficiently illustrated by the examples I have given.

Derek Parfit writes: 'If we believe that having some desire would have good effects, what that belief makes rational is not that desire itself, but our wanting and trying to have it. Irrational desires may have good effects. Thus, if I knew that I shall be tortured tomorrow, it might be better for me if I wanted to be tortured, since I would then happily look forward to what lies ahead. But this would not make my desire rational. It is irrational to want, for its own sake, to be tortured. The good effects of such a desire might make it rational for me, if I could, to cause myself to have it. But that would be a case of rational irrationality. (Parfit, 2001, 27)

In Parfit's example he has a comparably weak attitude-related reason to want to be tortured. It would make him have a more relaxed time until tomorrow. He also has a very strong content-related reason not to want to be tortured tomorrow. Any plausible aggregation of reasons will show that Parfit's verdict is correct: It would be irrational for him to want to be tortured. This, however, by no means shows that there are no attitude-related reasons. Parfit refutes the view that there are only attitude-related reasons and no content-related reasons for preferences. This is a view, I agree, no one should hold.<sup>19</sup>

## Bibliography

Broome, John, *Weighing Goods*, Blackwell, Oxford: 1991

Broome, John, 'Normative Practical Reasoning', *Proceedings of the Aristotelian Society*, Suppl. Vol. 75, 2001, 175-193.

Diamond, Peter, 'Cardinal Welfare, Individualistic Ethics and Interpersonal Comparisons of Utility: Comment', *Journal of Political Economy* 75 (1976), 765-766.



Frankfurt, Harry 1992, 'On the Usefulness of Final Ends', repr. in his *Necessity, Volition, and Love*, Cambridge University press: Cambridge 1999, 82-94.

Gibbard, Allan, *Wise Choices, Apt Feelings*, Clarendon Press, Oxford: 1990.

Harman, Gilbert 1997, 'Pragmatism and Reasons for Belief', repr. in his *Reasoning, Meaning and Mind*, Oxford University press, Oxford: 1999, 93-116.

James, William 1897, *The Will to Believe and Other Essays in Popular Philosophy*, Dover, New York: 1956.

Jeffrey, Richard, 'Preferences Among Preferences', *Journal of Philosophy* 71, 1974, 377-391.

Kavka, Gregory, 'The Toxin Puzzle', *Analysis* 43, 1983, 33-36.

Keynes, John Maynard, *A Treatise on Probability*, Macmillan, London: 1921.

Nozick, Robert, *The Nature of Rationality*, Princeton University Press, Princeton: 1993.

Olson, Jonas, 'Buck-Passing and the Wrong Kind of reason', *Philosophical Quarterly* 54, 295-300.

Owens, David, *Reason Without Freedom*, Routledge, London: 2000.

Parfit, Derek, 'Rationality and Reasons', in: Dan Egonsson, Björn Petersson, Jonas Josefsson, Toni Ronnow-Rasmussen (Eds), *Exploring Practical Philosophy: From Action to Values*, Ashgate: 2001. 17-41.

Pettit, Philip, 'Decision Theory and Folk Psychology', in M Bacharach & S Hurley (eds), *Foundations of Decision Theory*, Blackwell, Oxford: 1991.

Pettit, Philip, 'Hope and its Place in Mind', in V Braithwaite (Ed), *Annals of the American Academy of Political and Social Science*, Vol. 592, 2004, pp.152-65.

Piller, Christian, 'Doing What is Best', *Philosophical Quarterly* 50, 2000, 208-226.

Piller, Christian, 'Normative Practical Reasoning', *Proceedings of the Aristotelian Society*, Suppl. Vol. 75, 2001, 195-216.

Piller, Christian, 'Kinds of Practical Reasons: Attitude-Related Reasons and Exclusionary Reasons', forthcoming in: S Miguens, J A Pinto and C E Mauro, eds., *Actas do Encontro Nacional de Filosofia Analítica II*, Porto: Porto University, 2005

Rabinowicz, Wlodek & Ronnow-Rasmussen, Toni, 'The Strike of the Demon: On Fitting Pro-Attitudes and Value', *Ethics*.114, 2004, 391-423.

Raz, Joseph, *Practical Reason and Norms*, 2<sup>nd</sup> ed, Princeton University Press, Princeton: 1990.

Scanlon, Thomas, *What We Owe To Each Other*, Harvard University Press, Cambridge/Mass.: 1998.

Sen, Amartya, 'Choice, Orderings and Morality', in S Körner (ed), *Practical Reason*, Blackwell, Oxford: 1974, 54-67.

Sen, Amartya, 'Rational Fools: A Critique of the Behavioral Foundations of Economic Theory', *Philosophy and Public Affairs* 6, 1977, 317-344.

Sen, Amartya, 1995, 'Maximization and the Act of Choice', in his *Rationality and Freedom*, Belknap Press, Cambridge/Mass: 2002, 158-205.

Sen, Amartya, 'Introduction: Rationality and Freedom', in his *Rationality and Freedom*, Belknap Press, Cambridge/Mass: 2002, 3-64.

Velleman, David, 1996, 'The Possibility of Practical Reason', repr. in his *The Possibility of Practical Reason*, Oxford University Press, Oxford: 2000, 170-199.

---

<sup>1</sup> This objection is presented and defended in Rabinowicz&Ronnow-Rasmusson 2004, 404-410. Their concern, however, is not the denial of attitude-reasons – in fact they are sympathetic to the idea that there are such reasons – but to exclude attitude-related reasons in a deontic analysis of value.

<sup>2</sup> My account of the distinction between content-related and attitude-related reasons is similar to the more complicated account offered by Olson (2004).

<sup>3</sup> The feature that is indicted by a content-related reason for adopting some attitude will, of course, vary depending on which attitude it is. In the case of preferences we can think of it as comparative desirability, a notion that itself will vary depending on which substantial theory of desirability one adheres to. In the case of beliefs the notion of comparative likelihood can play a similar role.

<sup>4</sup> Some authors, see for example Owens (2000, 30), have suggested that we cannot decide to believe even if the evidence is indecisive. Take a coin toss: you cannot simply decide to believe it will come up heads. I agree that we cannot do that. But I'd say that this is just an instance of the above drawn limit to decisions to believe. In case of a coin toss it would be against reason to do anything but withhold one's judgment. To withhold one's judgment need not indicate lack of evidence but might, in fact, be best supported by one's evidence. See, for example, Keynes 1921, 'The Weight of Arguments', 71-91.

<sup>5</sup> I call this the Standard View as almost everyone accepts it. To mention only those who play more than one role in this paper, see the cited work of Broome, Gibbard, Nozick, Owens, and Parfit.

<sup>6</sup> The example which follows is a popular variant of a case discussed in Sen (1977, p. 328).

<sup>7</sup> Some philosophers hold that the rationality of choices is directly determined by the rationality of the preferences that are expressed by these choices: 'Whether a voluntary action is rational, then, is a matter of the rationality of preferences and intentions' (Gibbard1990, 39); see also Scanlon 1998, 21. Once we allow for attitude-related reasons, however, the relation between rational preferences and rational choices is less straightforward. A preference for doing something might have desirable features that aren't accountable for purely in terms of what such a preference is a preference for. Thus, I won't make use of the general principle endorsed by Gibbard and Scanlon. For the case in question, though, I will argue that taking the small apple is indeed rational for the polite person because in her circumstances she correctly prefers the small apple.

<sup>8</sup> Above quote is not an isolated occurrence in Sen's work, as he has endorsed the importance of meta-rankings throughout his work; see Sen (2002, 18).

<sup>9</sup> It is discussed in detail in Daniel Hausman's contribution to this volume.

<sup>10</sup> See Broome (1991) for general support of this strategy and Pettit (1991, 163-6) for applying it to the apple example. Although in the context of preferences Pettit figures as a prominent opponent of my defence of attitude-related reasons, it is noteworthy that in Pettit (2004), he argues for the rationality of hope on the basis of the positive features of having this attitude and, thus, allows attitude-related reasons for hoping.

---

<sup>11</sup> In section IV I pointed to Sen's work as sympathetic to the view defended here. In his 1995 paper, he continues to develop his ideas in a new conceptual framework, distinguishing between two conceptions of consequences or outcomes: 'A person's preferences over comprehensive outcomes (including the choice process) have to be distinguished from the conditional preferences over culmination outcomes given the acts of choice. The responsibility associated with choice can sway our ranking of narrowly-defined outcomes (such as commodity vectors possessed), and choice functions and preference relations may be parametrically influenced by specific features of the act of choice (including the identity of the chooser, the menu over which the choice is being made, and the relation of the particular act to behavioural social norms that constrain particular social actions)' (Sen 1995, 159). Against one's initial reaction, using the outcome terminology does not imply that Sen has switched to what I have called the opposing analysis. He indeed wants to show that even polite people can be seen as maximizers. This is, however, a purely technical result. About the function  $R(s^*)$  that could describe polite people as maximizers he says: 'The as if preference  $R(s^*)$  is, of course, a devised construction and need not have any intuitive plausibility seen as preference. [...] The as if preference works well enough formally, but the sociology of the phenomenon calls for something more than formal equivalences.' (Sen 1995, 191).]

<sup>12</sup> In footnote 5 I mentioned Gibbard's and Scanlon's view that the rationality of choosing some X depends on the rationality of preferring X. According to the opposing analysis new reasons come into view once we replace X as the objects of one's preference by choosing X. Thus, whenever there are what I call attitude-related reasons, the opposing analysis is incompatible with the view endorsed by Gibbard and Scanlon.

<sup>13</sup> The most common case of indifference occurs when content-related reasons are tied. I have preferences for certain cereals but no preferences for any of the packets of the same cereal. I choose one, any one, not because of what is chosen but because of the advantages of choosing. Can we retell this case in terms of preferences? Note that I don't really act randomly in such a situation. Usually I choose the one closest, i.e. I prefer the closest packet as choosing it is easiest. This preference arises – I want to suggest – only because I prefer to prefer something and thus, because of an attitude-related reason, which sets us the task of finding content-related reason that in other circumstances wouldn't decide the matter.

<sup>14</sup> In the framework presented here, this project is rational, because it is the project of aligning one's attitudes with one's (attitude-related) reasons. If these attitude-related reasons are sufficiently strong, any resistance such a project of change might meet would be a sign of irrationality. The stronger the resistance, the more we would speak of an obsession with death, thereby lending support to my view that conflicts between second-order and first-order preferences indicate failures of having one's attitudes in line with one's reasons.

<sup>15</sup> Here we may understand 'preferring the saucer of mud' as 'preferring to choose the saucer of mud'. The important point remains: preferring to choose the saucer of mud will be rewarded but choosing the saucer of mud will not be rewarded. Thus, if we prefer to choose the saucer of mud, we do so on the strength of the attitude-related reasons in play.

<sup>16</sup> For the Toxin Puzzle see Kavka 1983. Some of the points made here can also be found in the last section of Piller 2001, where I specifically focus on attitude-related reasons for intentions.

<sup>17</sup> For more details about why David Lewis's worry about this equivalence is, in my view, misguided, see Piller 2000.

<sup>18</sup> Let me add one further thought. Joseph Raz, especially in Raz (1990), has drawn our attention to what he calls 'exclusionary' reasons, i.e. reasons not to act for certain reasons. An agent's tiredness or the fact that an agent has been ordered to do something might be, on his account, such an exclusionary reason. The excluded reasons are content-related reasons. In Piller (2005) I try to show that an analysis in terms of attitude-related reasons of the phenomena on which Raz bases his exclusionary reasons is a plausible alternative to Raz's own account. If correct this would give us further examples for the role of attitude-related reasons. A view similar to my own on this matter can already be found in Jeffrey (1966).

<sup>19</sup> I want to thank all the participants of the conference where this paper was first presented for many informative and enjoyable discussions. Drafts of this paper have been read by John Broome, Philip Percival, Philip Pettit, Anthony Price and Wlodek Rabinowicz and I thank them for their helpful comments.