

Mark S. Peacock

Sen's Apples: Commitment, Agent Relativity and Social Norms

Abstract:

This paper examines Amartya Sen's notion of 'commitment' in light of Geoffrey Brennan's recent discussion thereof. Its aim is to elucidate one type of commitment which consists in following social norms. To this end, I discuss Sen's 'apples' example from his 'Rational fools' essay (*section 2*). In *section 3*, I draw some implications of commitments in Sen's work for the concept of 'agent relativity'. *Section 4* discusses the distinctiveness of Sen's conception of human beings in their supposed ability to be able to bind themselves to following social norms at the expense of their own benefit.

1. Introduction

Amartya Sen's notion of 'commitment' poses a challenge to understandings of rationality which conceive choices to 'reveal' or otherwise straightforwardly express the preferences of the chooser ('preferences' being related to the well-being of the person who has them). Sen has made several attempts to present, revise and illustrate the notion of commitment, but the concept is still surrounded by an air of mystery which is one reason it meets with both incomprehension and opposition. The fault is partly Sen's, for he sometimes presents the concept as if it were more straightforward than it is; one manifestation of this is an example he uses in his 'Rational fools' essay, in which 'commitment' was first introduced, which purports to exemplify the distinction between 'commitment' and 'sympathy'. The example concerns two boys who are each to choose one of two apples, one larger than the other. In a recent volume dedicated to commitment and rationality (Peter and Schmid 2007), Geoffrey Brennan offers an examination of Sen's apples, to my knowledge, the first sustained scrutiny which the example has received. Brennan (2007, 118) notes the "risk of weighing down a striking example with an excess of analysis", yet analysis is the only way of ascertaining whether the example does what Sen claims of it. I will burden it with yet more analysis in what follows despite my fear that analysis will sully the beauty of an example which, one might hold, is better left free of analysis. Brennan takes issue with Sen's view that the example illustrates the distinction between sympathy and commitment. In offering an alternative interpretation of the example, Brennan exposes facets of Sen's work which I draw out in this essay. My task is

‘exegetical’, in that I do not ‘take sides’ with or against Brennan or Sen; rather I try to elucidate precisely what Sen means by commitment and how his view is distinct from those of others, including Brennan. I do not tackle every aspect of what Sen writes on commitment—something which the format of a journal article does not permit; instead I focus on only one type of commitment which, Sen holds, people are able to make, namely, the following of social norms. Sen’s conception of what it means to follow self-imposed norms is quite different to other conceptions of the phenomenon put forward by economists. The difference bespeaks an entirely different *Menschenbild*. In limiting my focus to norm following behaviour, I believe I can reveal what is distinct about Sen’s conception of human beings without becoming unduly bogged down in his terminology, over which many a commentator has stumbled, of choosing ‘counterpreferentially’ or of acting on goals which are not one’s own.

In *section 2*, I contrast Sen’s and Brennan’s interpretations of the apples example, whereafter, in *section 3*, I explore an aspect of Sen’s work which Brennan’s analysis suggests but leaves unexplored, namely, that there be a connection between committed behaviour and what Sen calls ‘agent relativity’. The *fourth section* looks at people’s commitment to follow of social norms. Here, I purport to expose the crux of the difference between Sen and Brennan which revolves around the status of norms and the decision of the agent to follow or break a norm once she has committed herself to adhering to the norm.

2. Apples, Sympathy and Commitment

First the example—“of two boys who find two apples, one large and one small”—and Sen’s remarks thereon:

“Boy *A* tells boy *B*, ‘You choose’. *B* immediately picks the larger apple. *A* is upset and permits himself the remark that this was grossly unfair. ‘Why?’ asks *B*. ‘Which one would *you* have chosen, if you were to choose rather than me?’ ‘The smaller one, of course’, *A* replies. *B* is now triumphant: ‘Then what are you complaining about? That’s the one you’ve got!’ *B* certainly wins this round of the argument, but in fact *A* would have lost nothing from *B*’s choice had his own hypothetical choice of the smaller apple been based on sympathy as opposed to commitment. *A*’s anger indicates that this was probably not the case.” (Sen 1982, 93)¹

‘Sympathy’, Sen tells us, may be interpreted egoistically, for one has sympathy with another person when one’s own welfare is affected by that of the other; one takes pleasure in their pleasure and grieves at their pain. ‘Commitment’, by

¹ References to ‘Rational fools’, originally published in 1977, are to the reprint in Sen (1982). Sen’s apples example might have been inspired by a similar one offered by Ragnar Frisch which Sen (2002, 177) quotes in (published) translation. Sen was aware of the unpublished translation of Frisch’s article when he wrote ‘Rational fools’ (Sen 1982, 97; 2002, 177 note 31).

contrast, admits of no egoistic interpretation, for it can involve actions which the actor “believes will yield a lower level of personal welfare to him than an alternative” (Sen 1982, 92). Committed actions can, but need not, be one and the same as the action from which one expects to maximize one’s own welfare, but what characterises these cases as (non-egoistic) commitment rather than (egoistic) sympathy is the following counterfactual: the actor would undertake the committed act even if she expected it to detract from her personal welfare *vis-à-vis* an alternative act which would maximize her expected personal welfare. How is the apples example supposed to illustrate the distinction between sympathy and commitment?

One problem with the apples example is, as Brennan points out, its lack of detail and context; how one interprets it will vary according to the way in which one ‘fills in’ the gaps Sen leaves in his vignette. Let us begin with the hypothetical choice with which Sen ends the indented passage quoted above. If *A* had chosen first, had chosen the smaller apple and had done so out of sympathy, *A*, Sen holds, would have ‘lost nothing’. The argumentation underlying this claim is presumably thus: had *A* acted from sympathy, he would have received personal benefit from beholding *B*’s satisfaction at getting the bigger apple (assuming, of course, that any displeasure *A* receives from having the smaller instead of the larger apple is outweighed by *A*’s sympathetic pleasure with *B*’s pleasure at getting the larger apple). By implication, if *A* were imbued with a heavy enough dose of sympathy for *B*, *A* would have no grounds for complaint if *B*, in the role of first chooser, chose the bigger apple. Now let us turn to Sen’s actual example, namely that in which *B* chooses first and chooses the larger apple. How is *A*’s umbrage supposed to indicate that he has a commitment, and to what is he committed?

Brennan (2007, 118) suggests that *A* adheres to a norm of politeness and that *B*’s failure to observe this norm is the source of *A*’s irritation. The interpretation is plausible and conforms to a type of commitment which Sen associates with norm following (2002, 216; 2007, 348; 2009, 192–3), and to which I will return in *section 4*. But Brennan and Sen part ways in their interpretation of the example, with Brennan (2007, 119) holding that Sen claims too much for it. Were *A* to choose first, Brennan argues, he would face a simple trade-off between two values he wishes to pursue: being polite and getting the largest apple; and if, in the role of first chooser, *A* were to choose the smaller apple, it would be because the value of being polite weighed heavier on him than the value of getting the larger apple. For *B*, things are otherwise: politeness may be sacrificed for receiving the larger apple. Interpreted thus, the example poses no challenge to the orthodox notion of rational choice (Brennan, 2007, 118–19): there are two ‘goods’ or ‘objects of desire’, namely, adhering to the norms of polite behaviour and getting the larger apple; *A* and *B* have different preferences for these two goods, with *A* preferring to sacrifice the larger apple in favour of polite behaviour, and *B* preferring the opposite. The extra conceptual baggage—in particular, the understanding of commitment as counterpreferential choice—which Sen believes he requires to explain the situation is, according to Brennan, otiose. Brennan’s

contention that *A* and *B* face two ‘objects of desire’ in the example is crucial to grasping the difference between him and Sen, and I ask the reader to keep it in mind until I return to it in *section 4*. I finish this section, though, with an amendment to something both Brennan and Sen suggest about the apples example.

Brennan (2007, 120) asks how we would judge *A*’s response to *B*’s taking the larger apple (i) if *A* had sympathy with *B* (such that *A* derives pleasure from *B*’s (extra) pleasure from consuming the larger rather than the smaller apple) or (ii) if *A* acted from a “commitment (in Sen’s sense) to *B*’s flourishing”. In both cases Brennan suggests *A*’s complaint towards *B* would be ‘bizarre’. Regarding (i), Sen apparently agrees, for he writes that *A*’s anger at *B* taking the larger apple indicates that *A* acts from commitment, not from sympathy, the implication being that *A* would have no cause for complaint if he had sympathy for *B*. What lies behind these claims is presumably that a sympathetic *A* who cared sufficiently about *B*’s happiness, would be delighted to see *B* take the bigger apple (and possibly maximise his own utility if *B* takes the larger apple); hence *A* would have ‘lost nothing’, as Sen writes, but rather gained everything, which is the whole point of sympathy (understood egoistically). In Brennan’s second case, in which *A* is committed to *B*’s flourishing, *A*’s offer that *B* choose first may be understood as an instantiation of *A*’s commitment which, whilst it might not redound to *A*’s personal benefit, is a sacrifice *A* is willing (‘committed’) to make for the cause of furthering *B*’s flourishing. Thus follows Brennan’s conclusion that in both cases, (i) and (ii), a complaint from *A* would be ‘bizarre’. Nevertheless, I can imagine variants of (i) and (ii) in which *A*’s disgruntlement is not inexplicable. In the sympathetic case, *A* might derive enough satisfaction from *B*’s consumption of the larger apple to outweigh any dissatisfaction *A* feels from being left with the smaller apple, and hence it is correct, as Sen writes, that *A* ‘loses nothing’. But *A* might nevertheless feel peeved at *B*’s rudeness and have preferred that *B* show some sign of politeness and gratitude after being asked by *A* to choose first; *B*’s replying: ‘no, after you, old chap’ or ‘which one would you prefer?’, might have come as a welcome response to *A* who would then nevertheless have gladly insisted (with all the more joy) that *B* ‘simply must’ choose first and (hopefully) take the larger apple. In the case of an *A* committed to *B*’s flourishing, the same applies: a thus committed *A* may still take offence at *B*’s rudeness and chide *B* because of it, whilst remaining true to his commitment to *B*’s flourishing. Consider a similar example, of a parent committed to her child’s flourishing, who asks her child whether he would like a bed-time story to be read to him. If the child replies: ‘yes’, as opposed to: ‘yes, please’, the parent might berate the child, rather as *A* chides *B* in the apples example, but the parent might do so not in spite, but because, of her commitment to the child whom she understands as somebody in whom the conventions of politeness ought to be instilled. The parent’s commitment might be of the form in which Sen expresses an example of a commitment, viz.: ‘No matter how many times my son forgets to say ‘please’, I

will see to it that he gets a bed-time story come what may'.² This commitment does not preclude the parent's becoming irate at the child's rudeness. And so it might be with boy *A* who could maintain his commitment to *B*'s flourishing whilst bemoaning his impoliteness.

3. Commitment and Agent Relativity

Brennan's further analysis of Sen's apples reveals a more substantial point to which I attend in this section. In trying to ascertain whether Sen's distinction between sympathy and commitment corresponds to that between preferential and counterpreferential choice, Brennan (2007, 120) suggests the following contrast:

"Suppose *C* has a commitment to looking after *D* and makes transfers to *D* in amount *T*. Suppose *C** has sympathy for *D** to an extent that is behaviourally indistinguishable: the transfer *T* is the same in both cases. Now consider the responses of *C* and *C** to an increase of some amount in *D*'s] (and *D**[']s) well-being. One might conclude that in the sympathy case, *C** will be made better off by the increase in *D*'s well-being; whereas in the commitment case, *C* will not be affected one way or the other. Because *C* has not brought about the increase in *D*'s well-being, the increase simply does not bear on *C*'s commitments."

Perhaps unbeknownst to him, Brennan reveals therewith an important facet of Sen's notion of commitment, namely, the term's relationship to another which Sen discusses—'agent relativity'. To explore this facet further, let us pursue a lightly embellished form of Brennan's parallel examples involving *C* and *D*, and *C** and *D**. Of particular note is Brennan's final sentence in the passage just quoted, in which he suggests that, if *D* were transferred a sum by a third party, *E*, this would have no bearing on *C*'s commitment to care for *D*; *C* would 'not be affected one way or the other'. I agree with the conclusion that *E*'s transfer would have no bearing on *C*'s commitment in the sense that *C* would not see *E*'s transfer of a sum *T* as the fulfilment of her own commitment, and since *C*'s personal well-being might be unaffected by *D* receiving the transfer, it is plausible to write that *C* will not be affected. But there is reason to dispute the claim that *C* would 'not be affected one way or the other' in the sense of *C*'s evaluation of the situation which arises when *E* makes the transfer to *D* compared to her evaluation of the situation which arises when *C* makes the same transfer herself.

If *C**, who is sympathetic toward *D**, learns of *E**'s donation of a sum of money to *D**, *C**'s well-being (as a result of his sympathy) will increase. Indeed,

² I am adapting Sen's example of a commitment which he expresses thus: "no matter what happens to me, I must help her" (2009, 190).

C* might deem D* to have adequate funds as a result of E*'s act of munificence, and so C* will be better off if E* (and not C*) makes the donation, for, in both cases, C* would derive satisfaction from seeing D*'s wealth (and, by implication, D*'s well-being) increase, but, if E* makes the donation, C* does not have to 'foot the bill' for the improvement in D*'s well-being and for the concomitant improvement in C*'s own well-being (via sympathy with D*). Now consider the case of commitment whereby third party E makes a similar contribution to the well-being of D to whose flourishing C has made a commitment. Let us assume that D is disabled and C's commitment to D consists in improving D's mobility. E beats C to it and provides a wheelchair for D; C learns of this and no longer has to provide the wheelchair himself; the goal of C's commitment has been realized. One might think that C is better off now because he does not have to channel resources to D which he, C, could use for other causes. And indeed, on Sen's understanding of commitment, C might be better off in terms of her own well-being after E (and not C) provides the wheelchair, because, relieved of the cost of providing the wheelchair, C could spend the money on a summer holiday for herself and thereby further her own well-being. But there is something about E's provision of the wheelchair to D which, from C's perspective, constitutes a morally relevant difference from the case in which C provides the wheelchair herself, and this difference lies in the very fact that the wheelchair was provided by E and not by C. If, that is, C is committed to promoting D's well-being, then it seems not to be a matter of indifference to C whether the goal of furthering D's mobility be effected through C's own agency and not through the agency of a third party.

To take stock thus far, the sympathetic C* can be just as happy that D* receives a gift from a third party as he would be were C* herself to have given that gift; indeed, if given the choice, C* might actually try to engineer the situation such that D* would receive the gift from a third party. But the committed C, whilst possibly being personally better off if a third party provides a wheelchair to D, would not choose to have the third party provide the wheelchair but would insist on doing so herself. And, if she has a genuine Senian commitment to D, C would choose to provide the wheelchair independently of any 'warm glow' or other feeling of personal advantage she feels as a result of doing something good for D.

But what sort of difference does it make to C whether she or E provide the wheelchair to D? It cannot be a difference to her personal well-being because commitments (and whether they are realized) can, by definition, leave one's personal well-being unaffected. The difference it makes to a committed C, whether she herself or a third party provides the wheelchair to D, is in C's evaluation of the two possible states of affairs. From C's perspective of commitment, a strictly consequentialist analysis of the situation, in which one looks solely at what is provided to D but not by whom, would omit something of value from the

picture, that 'something' being *who actually provides D with the wheelchair*.³ And according to whether C himself or a third party, E, provides D with the wheelchair, C would evaluate the situation differently. When C provides the wheelchair herself, the state of affairs which comes to pass has greater moral significance or value for her because she has realised something which she values, namely, her commitment, something which nobody else can realise for her. The sympathetic C* might also have reasons for bemoaning the adequacy of consequentialist analysis of the transfer to D*, but he would do so for different a reason: C*'s reason for caring about who transfers the sum to D* is that, if someone else, not C*, transfers the sum to D*, C* will thereby save himself some money, safe in the knowledge (and satisfaction) that D*'s well-being has been enhanced, something from which C*, via sympathy with D*, benefits. C, on the other hand, cares about who transfers the sum to D because the commitment to enhancing D's mobility is her own.

For C, then, we have a case of agent relativity or, in Sen's taxonomy of agent relativity, we have a case of 'viewer relativity' (Sen 1981, 19–28). That is, the two states of affairs—(i) D receiving a wheelchair from E, and (ii) D receiving a wheelchair from C—are of differential moral significance from C's perspective, even though the strictly consequentialist outcome, viz., D receiving a wheelchair, is the same in each case. From the viewpoint of a dispassionate observer, however, who is not directly involved in the doings of C, D or E, it is neither here nor there who provides D with the wheelchair. Hence, to judge these two states of affairs, it makes a difference who is making the judgement, the (committed) doer, C, or a spectator. This difference in moral evaluation between C and a non-participating observer can arise even if both hold similar moral views according to which it is good that D receive the wheelchair. Therein lies the agent relative nature of the evaluation.

May we also conclude conversely that the sympathetic C*'s evaluation of the two states of affairs in which D* receives a transfer (in the one case from C*, in the other from a third party, E*) is agent neutral rather than agent relative? To answer this question, let us ask whether a person who has sympathy with another should be indifferent to whether the other's well-being be improved by the acts of a third party rather than by acts undertaken by herself. If, that is, my well-being is positively correlated with yours (through sympathetic sentiments from me to you), are there circumstances under which it becomes important to me through *whose* actions—mine or those of a third party—your well-being increases? Although an affirmative answer is possible and thus agent relativity conceivable in the case of sympathy, it is, I hold, unlikely to be the case. To see why, let us look, at the way in which Sen relates commitment and sympathy to action.

³ Sen (1981, 29) refers to what I have called 'strictly consequential analysis' as "especially narrow formulations" of consequentialism in which "consequences are defined *excluding* the actions that bring them about".

Sen (1981) relates his discussion of agent relativity to neither sympathy nor commitment, but his description of sympathy and commitment give as a clue to his position. Sympathy, he writes:

“corresponds to the case in which the concern for others directly affects one’s welfare. If the knowledge of torture of others makes you sick, it is a case of sympathy. If its [sic] does not make you feel personally worse off, but you think it is wrong and you are ready to do something to stop it, it is a case of commitment.” (1982, 91–2)

Of note here is that Sen immediately relates a person’s commitment to her own action (... ‘you are ready to *do* something to stop it’ ...), but there is no corresponding relation between sympathy and action; sympathy is related to a ‘concern for others’ but not necessarily to doing something oneself to increase their well-being. More generally, Sen casts sympathy as a more passive sentiment than commitment (commitment probably being ill-described as a ‘sentiment’ at all): with sympathy, “the awareness of the increase in the welfare of the other person [...] makes this [sympathetic] person better off” (1982, 92). Mere ‘awareness’ that the other person’s well-being has been increased is sufficient to increase the well-being of the sympathetic person; there is apparently no relevance as to whether the sympathetic person or someone else brings about this increase. In the introduction to *Choice, Welfare and Measurement* of which ‘Rational fools’ is a chapter, Sen also distinguishes the way in which sympathy and commitment relate to action: sympathy refers to:

“one person’s welfare being affected by the position of others (e.g., feeling depressed at the sight of misery). Commitment, on the other hand, is concerned with breaking the tight link between individual welfare (with or without sympathy) and the choice of action (e.g., acting to help remove some misery even though one personally does not suffer from it).” (Sen 1982, 7–8)

In these explications (which Sen quotes in his later works (2002, 214; 2007, 19; 2009, 188–9), commitment, but not sympathy, is clearly aligned with the actions of the person who has the commitment; the passage quoted connects sympathy to a ‘feeling’ (see the first parenthetical clause of the quotation) but it connects commitment to ‘acting’ (in the last parenthetical clause). It might, nevertheless, appear possible to construct cases for which sympathy goes together with agent relativity. Consider, for instance, a person, *P*, who acts altruistically towards another person, *S*, because *P* wishes to derive a good feeling (and hence personal well-being) from doing so. If *P* were to derive less satisfaction were someone else instead of *P* herself to act altruistically towards *S*, it is obviously important to *P* and that she herself (not that other person) performs the act. But this does not lead us to a situation of agent relativity because *P* is only evaluating two states of affairs (one in which she is altruistic to *S*, another in which a third party is altruistic to *S*) from the viewpoint of her own well-being. This is too partial

an evaluation to be classed as the sort of 'moral' evaluation with which agent relativity is normally associated. To evaluate the two states of affairs morally, *P* would have to abstract from her personal well-being whilst evaluating the alternative situations. But by being sympathetic (rather than committed) to *S*, *P* cannot make this abstraction; the whole motivation of her action is self-interested. Hence, I conclude that commitment but not sympathy is aligned with agent relativity.

I now turn to the crux of the disagreement between Brennan and Sen, namely, their interpretation of norm following behaviour.

4. Binding Oneself

One matter which fuels controversy over the notion of commitment is the apparent facility with which Sen conceives the possibility of committed actions. The notion that choices be 'counterpreferential' or that an agent pursue goals which are not exclusively her own is perplexing, if not, utterly counterintuitive to the ears of many. Furthermore, that Sen's examples of these phenomena are amenable to less controversial interpretations has led some to let parsimony (and plausibility) in explanation take precedent and to eschew the notion of commitment in explaining these examples. Brennan, as we have seen, holds that one can conceptualise Sen's apples example without making theoretical claims of such a controversial or apparently far-fetched nature. Daniel Hausman (2007), too, urges an understanding of commitment which does not rely on the idea of counterpreferential choices. He argues that, if one broadens the concept of preference to encompass *all* considerations, including those under the heading of which Sen places commitments, one can account for Sen's examples without introducing the idea of counterpreferential choices. When Hausman writes of 'all things considered' preferences, a 'consideration' may be understood as any factor which can motivate a particular choice and would include narrow self-interest, sympathy, commitments or anything else which could give one a reason to choose a course of action. By grouping all such motivations under the heading 'preference', counterpreferential choice would be impossible, for if something is to motivate a person to act, it must be a preference (in the broad, all things considered, sense), and if it is not a preference (all things considered), it cannot motivate action. But whilst banishing counterpreferential choice, Hausman sees the need to (re)introduce the sort of distinctions which Sen makes to get a richer understanding of human behaviour. The difference between the two lies in the fact that, whereas Sen distinguishes *between* choices motivated by preference (narrow self-interest and self-interested sympathy) and those motivated by (non-self-interested) commitment, Hausman makes such distinctions under the mantle of preferences all things considered. For Hausman, then, and in contrast to Sen, not only are narrowly self-interested and sympathetically motivated choices both preference-based (in a capacious understanding of the term), but so are choices based on commitment; commitment, though a different

type of preference to self-interested and sympathetic preferences, is nevertheless the manifestation of a preference. It is tempting to hold that the disagreement between Hausman and Sen is ‘merely’ a terminological matter, for both distinguish types of behaviour similarly, though with different vocabulary. Sen (2007, 353) confirms this when he registers no hostility to Hausman’s conception of ‘preference’ in the all things considered sense (though he does doubt whether Hausman will be successful in getting theorists of rationality to adhere to this notion of preference). What Sen believes distinguishes him from Hausman is Sen’s view that commitments stem from an agent’s ‘voluntary’, or *self-imposed*, (rather than ‘external’) restrictions on her own behaviour (Sen 2007, 353–4). His remarks on this difference, though, are both lapidary and somewhat cryptic. I hold that the differences between Sen and his critics can indeed be exemplified if one looks at Sen’s remarks on self-imposed restrictions which make an agent adhere to social norms. Because this difference is crucial to the exchange between Brennan and Sen, I return to the work of the former to elucidate the disagreement.

Consider two alternative presentations of norm following behaviour which we may call:

- (a) the preference abiding view,
- (b) the counterpreferential view.

The preference abiding view is manifest in Brennan’s interpretation of Sen’s apple example and is encapsulated in the phrase with which Brennan (2007, 122) concludes his argument against the view that committed behaviour be counterpreferential: “it is perfectly proper to treat agents as having preferences for norm-fulfilment”. As we saw in *section 2*, Sen’s boy *A*, according to Brennan, has a preference for adhering to the norm of politeness and prefers to sacrifice the bigger apple in favour of fulfilling the norm of politeness. There is, according to this view, nothing ‘counterpreferential’ going on here; both getting the bigger apple and adhering to the norm of politeness are ‘objects of desire’, to recall Brennan’s phrase, between which *A* chooses the one he prefers more which, as it happens, is politeness. Sen, on the other hand, adheres to (b), the counterpreferential view of norm following, according to which, if I commit myself to a norm, e.g. being polite, not stealing, not defecting in prisoners’ dilemma situations, then breaking the norm (being impolite, stealing opportunistically, defecting in a prisoners’ dilemma if I think I can gain personally from perpetrating such breaches) is no longer an object of choice for me; I have put it off limits to my rational calculation by committing myself to adhering to the norm. Consider boy *A* again. For Brennan, one may say, if boy *A* were to have first choice of apple, he would be ‘free to choose’ between the larger apple and being polite, these being the objects of desire between which he is to choose. For Sen, by contrast, *A*, having committed himself to adhering to the norms of politeness, is no longer free to choose in the manner just depicted; *A* might still consider taking the larger apple and adhering to norms of politeness as objects of *desire* in Brennan’s sense, but, if he has a commitment to politeness as Sen understands it, *A*

has no choice but to adhere to the norms of politeness; taking the larger apple would therefore not be an object of *choice* to *A*, however much he desires it (and desires it more than adhering to the norm of politeness). And if *A* had had first choice of apple and have permitted himself to take the larger apple, this would show that he did not really have a commitment to the norms of politeness at all, at least not in Sen's sense. What is more (and to reiterate what I wrote at the end of the previous paragraph), Sen conceives of *A* (or of anyone who can make a commitment to norm following) as somebody who can impose such constraints on his own contingent choices voluntarily and who can abide by his self-imposed constraints without the need of external sanction or fillip. External sanction or fillip might help *A* abide by his commitments, but Sen nowhere states that they are necessary to make *A* abide by his own commitments. By sticking to one's commitment to fulfil a norm, Sen claims that one's commitment can be strong enough to make one refrain from opportunistically acting to one's personal benefit by breaking the norm. A commitment is therefore a device with which an agent can ensure that her choices are time consistent; an agent, that is, who, at time t_1 , commits herself to adhering to a particular norm will adhere to this norm in acting at time t_2 even if, at t_2 , she does not expect adhering to the norm to redound to her personal advantage; breaking the norm would be preferable to her but nevertheless she adheres to the norm. This is what Sen calls the counterpreferential nature commitments. Sen, to my knowledge, has never expressed his notion of commitment in quite this way.

5. Conclusion: Sen and Behavioural Economics

To conclude, I elucidate the difference between Sen and many economists by contrasting his notion of commitment with the theories of human beings which have issued from behavioural and experimental economists. To see this, let us create a variation on Brennan's norm-following interpretation of *A*'s anger that *B* takes the larger apple by bringing to the fore that it is not only *B*'s failure to be polite which disturbs *A* but the fact that *A* expects *B* to reciprocate the politeness or generosity that *A* would manifest toward *B*, were *A* to choose first. *A*'s anger is to be explained not only by *B*'s failure to follow the content of the norm ('be polite!'), but also by the lack of reciprocation from *B*. The expectation of reciprocity is attached to many, perhaps all, norms unless there is good reason for relieving one party to a situation of the obligation to follow the norm. *A* sees no such reason in the case of *B*. *A* sees *B* and himself as part of a 'community' in which both are bound by restrictions imposed by norms of politeness; *B* sees the relationship between himself and *A* differently because, for *B*, apparently no such norm presents itself as binding on him (and presumably, unless he has an outrageous cheek, *B* would not hold it against *A* if, in the position of first-chooser, *A* would take the bigger apple). *A*'s claim (made on his behalf by Sen) that, had he chosen first, he would have taken the smaller apple, may be seen as an offer of 'cooperation' towards *B* (analogous to cooperation in a prisoners'

dilemma or in making contributions to the provision of public goods). Experimental economists have produced results which suggest that cooperators like *A* expect reciprocal cooperation from those with whom they interact, and that, if reciprocal cooperation is not forthcoming, they are prepared to ‘punish’ non-reciprocators at a cost to themselves (see, e.g. Fehr and Fischbacher 2003; Gintis et al. 2003). Although Sen does not conceive the interaction between *A* and *B* as ‘iterated’, one may nevertheless conceive *A*’s complaint towards *B* to stem from a willingness to ‘punish’ *B* for his transgression, and thus *A*’s complaint is a type of ‘altruistic punishment’ (Fehr and Gächter 2002). Experimental economists rightly see the theoretical innovations they have made in light of experimental results as a challenge to mainstream economists’ notion of rationality. In particular, they challenge one assumption which Sen challenges in ‘Rational fools’, viz., that human beings are motivated solely to pursue their own self-interest (narrowly conceived) (Fehr and Fischbacher 2003, 788; Sen 1982, 105). But experimental and behavioural economists conceive cooperation and altruistic punishment to be based on *preferences* for fairness or for distributions which answer to the name of ‘equality’ (Fehr and Schmidt 1999). This is at odds with Sen’s conception of human beings who can act not only on preferences which are not ‘self-interested’ but can bind themselves to act on principles for which they have no preference. If I have a preference for fairness or equality, I look at the attainment of fair or egalitarian outcomes as objects of choice which I must compare to other objects of choice between which I can choose. But if one has a commitment to social justice (one of Sen’s examples), then attaining socially just outcomes are no longer objects of choice but obligations. Commitments therefore serve as ‘side constraints’ on one’s own behaviour; if the things to which we commit ourselves remain objects of choice, then they are no more than objects which we may choose or decline to choose according to our whims; to capture such a situation, a single (but encompassing) notion of preference would suffice, and Brennan and Hausman would be correct to reject the notion of counterpreferential choice.

Sen’s *Menschenbild*, though, deviates further from that of mainstream theorists of rationality than that of behavioural economists or of theorists like Brennan and Hausman (cf. Sen 2002, 26–33). I have not attempted here to defend Sen’s *Menschenbild* or to investigate examples in which it is valid. I hope, though, to have elucidated the peculiarity of his position in a new and perspicuous way.

References

- Brennan, G. (2007), “The Grammar of Rationality”, in: Peter, F. and H. B. Schmid (eds.), *Rationality and Commitment*, Oxford: Oxford University Press, 105–23.
- Fehr, E. and U. Fischbacher (2003), “The Nature of Human Altruism”, *Nature* 425, 785–91.
- and S. Gächter (2002), “Altruistic Punishment in Humans”, *Nature* 415, 137–40.

-
- and K. Schmidt (1999), “A Theory of Fairness, Competition, and Cooperation”, *Quarterly Journal of Economics* 114, 817–68.
- Gintis, H., S. Bowles, R. Boyd and E. Fehr (2003), “Explaining Altruistic Behaviour in Humans”, *Evolution and Human Behavior* 24, 153–72.
- Hausman, D. (2007), “Sympathy, Commitment, and Preference”, in: Peter, F. and H. B. Schmid (eds.), *Rationality and Commitment*, Oxford: Oxford University Press, 49–69.
- Peter, F. and H. B. Schmid (2007) (eds.), *Rationality and Commitment*, Oxford: Oxford University Press.
- Sen, A. (1981), “Rights and Agency”, *Philosophy and Public Affairs* 11, 3–39.
- (1982), *Choice, Measurement and Welfare*, Cambridge, Mass.: Harvard University Press.
- (2002), *Rationality and Freedom*, Cambridge, Mass.: Harvard University Press.
- (2007), “Rational Choice: Discipline, Brand Name, and Substance”, in: Peter, F. and H. B. Schmid (eds.), *Rationality and Commitment*, Oxford: Oxford University Press, 339–61.
- (2009), *The Idea of Justice*, Cambridge, Mass.: Harvard University Press.