Economics and Philosophy, **17** (2001) 181–205 Copyright © Cambridge University Press

CONSEQUENTIALISM, RATIONALITY AND THE RELEVANT DESCRIPTION OF OUTCOMES

Bruno Verbeek
Research School of Social Sciences, Australian National University

1. INTRODUCTION

Instrumental rationality requires that an agent selects those actions that give her the best outcomes. This is the principle of *consequentialism*.[1] It may be that it is not the only requirement of this form of rationality.

I would like to thank Gijs van Donselaar, Govert den Hartogh, Edward McClennen, Hans Rott, the editors of Economics and Philosophy, three anonymous referees, as well as audiences at the University of Toronto, Bowling Green State University and Carnegie Mellon University. Special thanks are due to Wlodek Rabinowicz for his meticulous and helpful comments.

[1] The traditional conception of consequentialism characterizes consequentialist theories as theories that have two components. First, they have a principle that ranks states of affairs from best to worst from an impersonal point of view. Secondly, they require the agent to produce the highest-ranking state of affairs he is in a position to produce (Scheffler, 1988, p. 1).

My definition of consequentialism is weaker for it does not assume the first part. It does not invoke the idea that the standard of evaluation of the consequentialist needs to be impersonal. Even an egoist is a consequentialist on my definition. All that I assume, therefore, is the notion that a consequentialist is dedicated to maximizing with regard to his values. Recently this assumption has been under attack. For example, the idea that all considerations of value can be compared has been challenged (e.g., Chang, 1997). Clearly, if it is not the case that A is at least as good as B, nor that B is at least as good as A, we cannot maximize since it is impossible to determine which of these two options is the best.

Another attack on the ideal of maximization comes from authors who defend *satisficing* as the correct response to values (e.g., Slote, 1989). In so far as the defense of satisficing appeals to notions of maximization (e.g., the claim that satisficing is the best way to deal with the epistemic limitations of agents) one can legitimately wonder whether it really is something different than maximization.

https://doi.org/10.1017/S0266267101000220 Published online by Cambridge University Press

Considerations other than the outcomes may enter the picture as well. However, the outcome(s) of an action always play a role in determining its rationality. Seen in this light consequentialism is a minimum requirement of instrumental rationality. Therefore, any theory that tries to spell out the implications of instrumental rationality, in particular expected utility theory, should subscribe to the principle of consequentialism.[2] Or so it seems.

Peter Hammond (1988) has claimed that the connection between consequentialism and rational choice is even stronger. He claims that one can derive expected utility theory from the principle of consequentialism.[3] Though I will not discuss expected utility theory, nor its derivation, explicitly in this essay, this is sufficient reason to investigate whether consequentialism is indeed an unproblematic and minimal requirement of rationality.

Conventional wisdom has it that consequentialism is more or less restrictive depending on what can count as the outcomes of an act. The following example illustrates the point. Suppose Mom is to determine which of her two children, Jane and Peter, will get a single indivisible treat. She could either give it to Jane or to Peter. However, Mom believes it is better to flip a coin first and give it to one of her children, depending on the outcome, since this is fair. Is there a relevant difference in outcome if she gives it to Jane right away, or if she flips a coin first and then gives it to Jane? In other words, should the result of Mom's actions be interpreted as 'Jane gets a treat' or 'Jane gets a treat in a fair manner'? This is the problem of relevant description of the outcomes of an action.

Many economists and philosophers believe this problem is independent from the question whether consequentialism is a plausible requirement of rational choice. They believe that consequentialism is a purely formal requirement that does not constrain any choice until one has identified the appropriate domain of outcomes. In this paper I want to

[2] 'An almost unquestioned hypothesis of modern normative decision theory is that acts are valued by their consequences'. (Hammond, 1988, p. 25).

[3] To be precise, Hammond (1988) has shown that the assumption of consequentialism together with an assumption about the domain of decision trees to which consequentialism is applied, is necessary and sufficient to deduce the weak-ordering axiom and the independence axiom. If one also assumes continuity of the preference ordering, one can deduce the central result of expected utility theory, i.e., the existence of a representing function that is unique up to affine linear translations. Hammond's result is important because it gives a strong normative foundation to the assumptions of expected utility theory. Weak ordering and independence, the two most contested assumptions of expected utility theory, are theorems in his construction rather than independent assumptions. This seems to settle that debate. However, his critics have been quick to point out that Hammond's notion of consequentialism is far from unproblematic itself. See especially McClennen, 1990; Levi, 1991.

https://doi.org/10.1017/S0266267101000220 Published online by Cambridge University Press

show that stating the problem in this way obscures an important issue. I will argue that describing outcomes in terms of a particular conception of fairness is incompatible with consequentialism as it is usually interpreted in decision theory. I believe this conception of fairness to be a paradigm example of a whole class of values, dispositions and emotions, which cannot be reconciled with the decision-theoretical sense of consequentialism. As a result one should either reject the idea that there is a rationally relevant difference between the situation where Jane gets her treat without tossing a coin, and the situation where Jane gets her treat after Mom has flipped a coin. That is, one should reject the idea that the fairness, as well as certain other values, can make a difference, thus saving consequentialism. Or – and this is the solution I favor – one should reject consequentialism and replace it with a principle of choice that does allow for such a difference.

The remainder of this paper is organized as follows. In Section 2 I introduce two fundamental consequentialist assumptions. First, the notion that considerations that are irrelevant from the perspective of the agent's values should not determine her choices. Secondly, the assumption that consequentialism is forward-looking. I argue that Mom's preferences violate either the former or the latter. Section 3 discusses three possible responses to these violations; in particular the idea that one could save both expected utility theory and consequentialism, by re-describing the relevant outcomes. Section 4 discusses the worry of some theorists that such a move might rob both consequentialism (and expected utility theory) of its normative content. It focuses mainly on Broome's idea of *rational justifiers*. Section 5 discusses Broome's special response to cases like the one under discussion, thus clearing the way for the argument of Section 6, that consequentialism is incompatible with outcome descriptions that invoke fairness. Sections 7 and 8 discuss the question of whether the consequentialist principles introduced in Section 2 are acceptable for a consequentialist. I argue that they are at the very heart of what consequentialism entails. However, since they exclude from deliberation notions of fairness, guilt, disappointment, and regret, they are unacceptable as requirements of rationality. Section 9 summarizes the main conclusions and speculates on the implications for expected utility theory as a theory of rational choice.

2. A VIOLATION OF CONSEQUENTIALISM

Like most mothers, Mom loves her two children, Peter and Jane.[4] She would like to give them both a treat. Unfortunately, she can give only

[4] The example is based on the discussion in Diamond, 1967.

https://doi.org/10.1017/S0266267101000220 Published online by Cambridge University Press

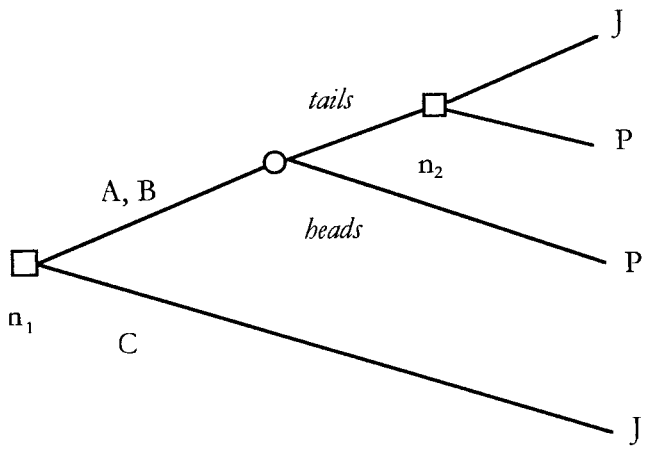


FIGURE 1

one of them a treat. Since both children are equally deserving, she is indifferent which of the two should get it.[5] Therefore, she is indifferent between the outcome in which Peter receives the treat (P), and the outcome in which Jane receives the treat (J). However, she prefers to flip a fair coin and let it decide who gets the treat. Therefore we have the following information about her preference ordering: [P, 0.5; J, 0.5]>P~J. (The expression [P, 0.5; J, 0.5] should be interpreted as the lottery which has as prizes P, with associated probability 0.5, and J, with associated probability 0.5.)

On the face of it, Mom's preferences seem consistent with consequentialism. Clearly, if the lottery is the best outcome, consequentialism requires Mom to choose it. However, Mom's preferences pose a dilemma from a consequentialist perspective. She violates one of two fundamental consequentialist assumptions. This becomes apparent when we look at Mom's situation as a dynamic choice, that is, as a situation where the agent has to make a series of choices before reaching the preferred outcome (see Figure 1).

First some conventions. A square node represents a point in the tree where a choice has to be made, whereas a circular node represents a chance event. A *plan* is a specification of the agent's moves at all choice nodes that the agent can reach, given the relevant chance events, by

[5] In this paper I often use terminology that is related to the ordering of preferences, whereas consequentialism is defined in terms of 'good'. Though I use the two interchangeably, I do not mean to imply that preferences constitute what is good. I try to be agnostic about such meta-ethical issues here. Some consequentialists are preferentialists; others are not (see also Section 5). I address both varieties of the species in this paper.

https://doi.org/10.1017/S0266267101000220 Published online by Cambridge University Press

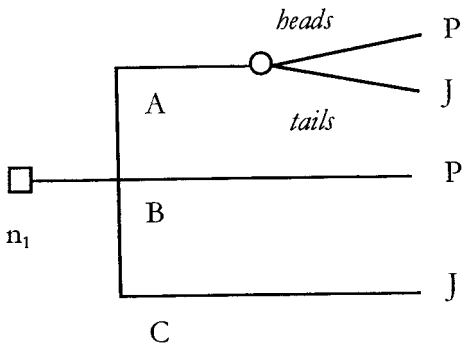


FIGURE 2

making the earlier moves determined by the plan.[6] In this tree there are three plans: plan A (go up at n_1, if tails, go up at n_2), plan B (go up at n_1, if tails, go down at n_2), and plan C (go down at n_1). The outcomes that could be reached by the execution of a plan are the set of associated outcomes of that plan. (For example, plan A has the set of associated outcomes {J, P}.) If this set of associated outcomes is not a singleton, the resulting outcome of the implementation of the plan will depend on the intervening chance events. The *prospect* of a plan is a lottery that has the elements of the set of associated outcomes of that plan as prizes with their respective probabilities. Thus, in this tree there are three prospects: that of plan A, [P, prob(*heads*), J, prob(*tails*)] (the preferred prospect); that of plan B, [P, prob(*tails*); P, prob(*heads*)] (which reduces to [P]), and that of plan C, [J]. A plan is admissible for a consequentialist if and only if there is no plan available with a better prospect.

Having fixed our terminology, we are now in a position to explain why Mom violates consequentialism. If Mom were presented with the three prospects of Figure 1 in such a way that she only had to make one choice before realizing the desired prospect, she would regard plan A uniquely admissible. That is, if Mom is faced with the normal form reduction of the tree in Figure 1 (see Figure 2), she chooses for the lottery that gives both her children an even chance of getting the treat. This follows straightforwardly from her preference ordering: [P, prob(*heads*); J, prob(*tails*)]>P~J.

However, when the very same prospects are offered to her in the tree of Figure 1, Mom should not only consider plan A admissible but plan B as well given her own preference ordering. This perhaps surprising result is inevitable if we look closer at the choices Mom faces in the tree of Figure 1. At n_1, Mom has three plans available, A, B and C. Given her

[6] This is the same informal definition as that of Rabinowicz, 1995, p. 590.

https://doi.org/10.1017/S0266267101000220 Published online by Cambridge University Press

preference ordering she will rule out plan C, so she will go 'up'. Let us assume 'heads' comes up and Mom finds herself at n_2. Now Mom faces a second choice between P and J. Assuming her preferences have not changed along the way, she is indifferent between the two. However, that means that both going 'up' and going 'down' are admissible at n_2. Since going 'up' at n_2 is exactly what plan A prescribes at that choice node and going 'down' is part of plan B, it now appears that both A and B are acceptable![7]

What has gone wrong here? Why does Mom judge plan B unacceptable in Figure 1 but acceptable in Figure 2? Since her preferences have not changed between n_1 and n_2, consequentialism allows such differences in acceptability only if the (value of the) prospects in Figures 1 and 2 differ. Since the prospects in Figures 1 and 2 are identical, it must be the case that the shape of the tree affects their value. In this case, this is only possible if the very process of making two subsequent choices somehow affected the value of the outcomes, in such a way that plan A and plan B both become acceptable. However, it is hard to imagine that making a choice the second time is costly. Moreover, if it is, why does that affect the value of J, the outcome under plan A, more than that of P, the resulting outcome of plan B? Therefore, the conclusion is warranted that Mom allows irrelevant considerations to play a determining role in her choices. As a result we should conclude that Mom is not a good consequentialist. Consequentialism requires that she does not differentiate between the two trees.

We can state the consequentialist intuition under consideration here in a bit more detail. A consequentialist chooses her plans in accordance with the following principle:

> *Normal-form/extensive-form coincidence* (NEC): Let T be any decision tree with associated set of plans S and let T^n be the normal-form representation of T, then for any $s \in S$, s is acceptable in T if and only if it is acceptable in T^n.[8]

[7] This is the very reason why Seidenfeld (1988) argues that violations of independence lead the agent to 'sequential inconsistency'. Given Mom's preferences, in Figure 1 the admissible plan in n_1 is ('up' and if 'heads' choose J). But this plan implies that at n_2 the only admissible plan-continuation is ('. . ., up'). However, since Mom is supposed to be indifferent between P and J, ('. . ., down') should be admissible as well. Seidenfeld maintains that something has got to give here. Either one rejects that ('. . ., up') is an admissible plan, which would contradict Mom's preferences, or it cannot be the case that ('up' and if 'heads' J) is uniquely admissible at n1, which would contradict Mom's preferences as well. Both options, therefore, would imply Mom should change her preferences such that they no longer violate independence. However, see (McClennen, 1988; Rabinowicz, 1995, 1997) for a criticism of this argument. See note 15 as to why Mom violates independence.

[8] McClennen (1990 p. 115) presents NEC in the context of his reconstruction of the notion of consequentialism that is employed by Hammond (1988, pp. 28, 37).

https://doi.org/10.1017/S0266267101000220 Published online by Cambridge University Press

Mom violates NEC, because in the situation in Figure 2 there is only one acceptable plan (i.e., plan A), whereas there are at least two acceptable plans (i.e., A and B) in Figure 1.

What exactly causes Mom to violate NEC? It is the result of another, equally fundamental, consequentialist assumption, to wit, that deliberation should be forward-looking. It is assumed that a consequentialist is concerned with what lies ahead of her and nothing else. She deliberates about the outcomes she can realize, not about what she could have realized. In other words a consequentialist ignores sunk costs. Having reached n_2 in Figure 1, Mom deliberates about what lies ahead and judges that she should be indifferent between [J] and [P]. Therefore, both moves are admissible from a consequentialist point of view. As a result Mom is bound to violate NEC.

We can make the forward-looking nature of consequentialism more precise in the following way.

> *Separability* (SEP): Let T be a tree and $T(n_i)$ be a separate tree, identical to the tree continuation of T from n_i onward. Let $s(n_i)$ stand for the plan continuation in T from n_i on. Every plan available in T from n_i on is also available in $T(n_i)$. Consequentialism requires that $s(n_i)$ is an acceptable plan continuation if and only if its corresponding plan in $T(n_i)$ is also acceptable.[9]

Suppose Mom would not violate NEC and treat the trees in Figures 1 and 2 the same. Then she is bound to violate SEP. If she were to face the *de novo* choice between [J] and [P] she would be indifferent.[10] Therefore, SEP requires that she is indifferent at n_2 between both plan continuations. However, in the context of the tree in Figure 1, compliance to NEC requires that she is not indifferent because in Figure 2 she prefers the lottery. Therefore, Mom cannot avoid violating either NEC or SEP. As a result she is not a good consequentialist.

[9] McClennen, 1990, pp. 120–22. Hammond regards SEP as an unproblematic consistency requirement. E.g., Hammond, 1988, p. 34. Note that SEP restricts the acceptability of plans if and only if one assumes that acceptable plans consist of acceptable plan continuations. There is reason to doubt the plausibility of this assumption if it is possible that the agent receives new, unforeseen information during the execution of his plan. However, in the case of Mom this is not the case.

McClennen (1997) weakens the separability condition. Now he allows $T(n_i)$ to contain '. . . reports as to how *as a matter of fact* you resolved earlier choice options' [original italics]. He goes on to argue that this formulation of SEP is still too restrictive. However, he does not explain if, and if so, how, such reports can make a difference for choice in comparison to a *de novo* tree. So while this seems an important concession, it is unclear if this difference allows for a difference in planning.

[10] As one referee pointed out to me, this assumes that Mom does not change her preferences in such a *de novo* situation. I take this to be a plausible assumption in the present example.

https://doi.org/10.1017/S0266267101000220 Published online by Cambridge University Press

3. THREE RESPONSES

As far as I can see there are three possible responses for a choice theorist confronted with Mom's choices. First, the theorist might argue that, apparently, consequentialism is not an adequate principle to describe and predict the actual choices that mothers make. Moreover, since consequentialism is an implicit assumption in expected utility theory, this theory is inadequate insofar as a theory of rational choice is supposed to have empirical value. However, so the rejoinder proceeds, consequentialism is required by rationality understood as a normative enterprise. So, precisely because the agent in the example violates consequentialism, one should condemn her behavior as irrational.

Secondly, one might take a hard-nosed empiricist stand. If the theory is empirically inadequate – so much the worse for the theory. Consequentialism loses its normative appeal when we discover that actual agents violate it in systematic ways.

Both responses are too radical. The first response is too strong in the sense that it is willing to accuse the majority of subjects of blatantly irrational behavior. It reminds one of the scientist who blames her observations for producing the wrong data without ever questioning the correctness of her theory. The second response is too strong in that it seems to be unwilling ever to accuse people of irrationality. It is to be expected that situations will arise which are so complex that agents will not be able to figure out the best course of action even with the help of technologies such as pencil, paper, books, mathematical approximations and computers.

Therefore, a more plausible response takes a middle position between the two extremes. Rather than condemning Mom, we might interpret her choice as an initial indication that consequentialism has insecure normative standing. If otherwise seemingly reasonable agents display Mom-like behavior, consequentialism becomes less secure as a core requirement of rationality. Therefore, we need additional, independent grounds for accepting it as a requirement for rational choice.

However, there is a third possible response, which avoids this difficult task altogether. The theorist could argue that Mom's choices are completely consistent with consequentialism. For, so the disputant argues, there is no structural resemblance between the three prospects in Figure 1. Plan A, the plan that leads to the lottery over P and J, results in a *fair* outcome, whereas the outcomes of plans B and C are unfair. If we substitute this fairness in Figures 1, we see that there is no violation of either NEC or SEP at all (see Figure 3). The reason Mom supposedly violated NEC was the fact that in Figure 1 her choices at n_2 reduce to a choice between the indifferent options P and J. Her preference for the lottery between P and J in the normal form representation of the

https://doi.org/10.1017/S0266267101000220 Published online by Cambridge University Press

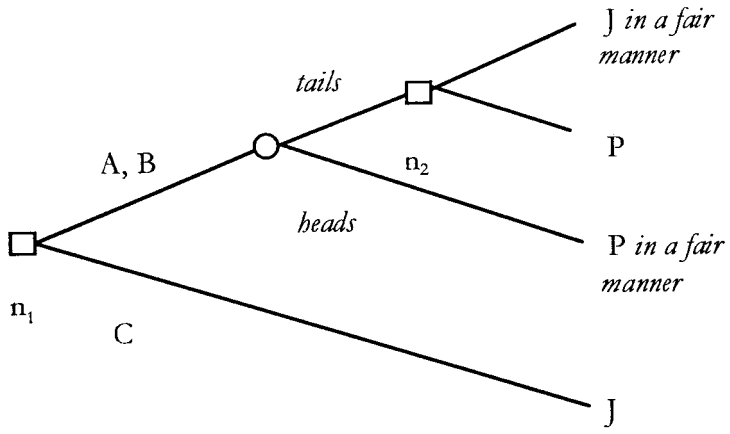


FIGURE 3

available options is not consistent with the admissible plans in the dynamic form representation. However, if at n2 she faces the choice between P and ('J, *in a fair manner*') as is the case in Figure 3, one can appreciate that Mom has a strict preference for the latter, even though she is indifferent between P and J.

In addition, Mom's strict preference for the plan continuation ('. . ., up') at n_2 is compatible with SEP, since in the *de novo* tree corresponding to this continuation ('up') is the uniquely acceptable plan.

This third response is the most common among rational choice theorists. Thus Amos Tversky, otherwise quite critical of expected utility theory, says:

> The question of whether utility theory is compatible with the data or not, therefore, depends critically on the interpretation of the consequences. (Tversky, 1975)

Making the same point in his discussion of the Allais paradox, John Broome argues:

> All the rationalisations of Allais' preferences point out a good or bad feeling you may experience. This feeling . . . can rationally be taken into account, as well as the money prizes, in determining your preferences. Let us write this feeling into our table of lotteries . . . [11] (Broome, 1991)

Similarly, Peter Hammond argues that:

> It would be false if missed opportunities, regrets, sunk costs, etc., affected behaviour and yet were excluded from the domain of consequences. As a

[11] However, Broome adds important qualifications, which I discuss below.

https://doi.org/10.1017/S0266267101000220 Published online by Cambridge University Press

> normative principle, however, consequentialism requires everything which should be allowed to affect decisions to count as a relevant consequence – behaviour is evaluated by its consequences, and nothing else. If regrets, sunk costs, even the structure of the decision tree itself are relevant to normative behaviour, they are therefore already in the consequence domain. (Hammond, 1988)

What these and other authors ask us to do is to re-describe the outcomes in the face of apparent violations of the axioms of rational choice.

4. IS CONSEQUENTIALISM TRIVIAL?

There are some serious worries about the force of this response. The most widely shared concern is that this move threatens to rob the theory of rational choice of its normative content. Suppose we can find re-descriptions for every apparent violation of consequentialism. It seems that this would save consequentialism as a principle of rational choice. However, we will be left with a principle that cannot be violated. Consequentialism will be a tautology on this proposal. What good is a requirement of a normative theory of rationality if it does not constrain choice? Consequentialism becomes an empty constraint because every choice will satisfy it after re-description of the relevant outcomes.[12]

Some welcome this tautological status of consequentialism. For example, Peter Hammond argues in several places in his work, that consequentialism (as well as expected utility theory) is empty as long as we have not settled on what the proper domain of outcomes is. In his view, the theory of rational choice has no normative content until one has determined what is to count as a relevant outcome. That is, until one has determined the proper values. It is, therefore, the value theory that does all the normative work, not the abstract axioms of standard expected utility theory.

Others do take seriously the worry that consequentialism is normatively vacuous.[13] For example, John Broome (1991) is concerned about

[12] In addition, it becomes unclear whether Hammond is entitled to deduce weak ordering and independence from consequentialism. Hammond suggests that if the structure of the decision tree itself is relevant to normative behavior it is therefore already in the consequence domain. In his insightful comment on Hammond's proof (Munier, 1996) has argued that if we take this suggestion seriously, the domain of logically possible trees with end nodes in the given domain can be very restricted. It might be restricted to just one tree. As a result, one cannot deduce either weak ordering or independence.

This result is not just important for those seeking to establish the normative plausibility of standard expected utility theory, but also for those moral philosophers who subscribe to consequentialism. The implication of Munier's comments is that one cannot *deduce* general duties to maximize values that by their very nature are limited to one tree, i.e., one choice situation, only.

[13] E.g., Machina, 1989; Tversky, 1975.

https://doi.org/10.1017/S0266267101000220 Published online by Cambridge University Press

the complaint that the theory of rational choice is empty. His solution is orthodox in that he tries to answer the complaint with a theory of value. In *Weighing Goods* he argues that what is needed is a theory that tells us how we should individuate the relevant outcomes of our choices. Such a theory would help us in determining how far we are allowed to go with re-describing the relevant outcomes. Broome proposes the following principle for this individuation:

> *Principle of individuation by justifiers*: outcomes should be distinguished as [preferentially] different if and only if they differ in a way that makes it rational to have a preference between them. (1991, p. 103)

If we apply Broome's principle to the re-description of Mom's predicament in Figure 3, we get the following result. It is rational to prefer the plan-continuation of A to that of B and to be indifferent between B and C, if and only if it is rational to prefer 'J *in a fair manner*' to 'J'. If this is the case we do not have a violation of consequentialism because we are dealing with a relevant difference in the prospects of plans A, B and C. On the other hand, if it is irrational to be anything but indifferent between 'J' and 'J *in a fair manner*', the parent who persists in judging plan A better than B, while being indifferent between B and C, is irrational.

This seems to provide a way out of the worry that consequentialism is trivially true, but it does so at a cost. In Broome's view there are requirements of rationality that are independent of the principles of rational choice. Traditionally, the axioms of expected utility theory are regarded as criteria of consistency. It is a widely held opinion that rational choice theory cannot demand more than this kind of consistency since it is a theory of instrumental rationality. Once we allow for additional criteria to determine the rationality of individual preferences, it seems we have abandoned the uncompromising 'thinness' of the ideal of instrumental rationality that made it so attractive to philosophers and social scientists alike. If we follow Broome's suggestion, we will have incorporated some form of objectivism or realism, in the underlying value theory.[14] As a result, one cannot claim that it is rational to choose A instead of B because one prefers A to B. Rather, it is the other way round: one should prefer A to B (and, therefore, choose A) because it is rational to prefer A to B.

I am not saying that this is wrong. However, one should recognize (as Broome does) that this is a radical departure from how most rational choice theorists regard the concept of rationality. In general, any attempt

[14] I use these terms very loosely here to refer to the idea that there are constraints on what can count as valuable for an agent which are in some sense independent of the agent.

https://doi.org/10.1017/S0266267101000220 Published online by Cambridge University Press

to avoid the normative vacuity of consequentialism by introducing rational constraints within the theory of value amounts to such a departure from the strictly instrumental conception of rationality.

5. INTERMEZZO: BROOME ON FAIRNESS

Broome has a special, independent response to cases like Mom's preference for a fair lottery.[15] In his view, the real worry about Mom is not the rationality of the re-description of the object of her preference. Rather, it is the worry that her preference reveals that the value of fairness emerges 'in the interactions between states of nature', which would contradict the consequentialist claim that all value is completely contained in outcomes alone. He proceeds to argue that value can be 'dispersed' – as he calls it – over states of nature if we take a sufficiently broad view about what is contained in an outcome. To all intents and purposes an outcome should be conceived of as a complete description of the state of affairs in that particular possible world. Broome continues to argue that such a complete description contains references to the causal path through which this world was realized. He then claims that the description of the causal processes that have led to the actualization of *this* world necessarily contains a reference to 'what could have been'. The value of fairness then supervenes on all of those references (just as the disposition of being inflammable supervenes on the chemical composition of an object regardless of whether it actually burns or not).[16]

We need not be too concerned about the reference to supervenience to see Broome's point. If a possible world contains meaningful references to other possible worlds, the description 'J while P could have been the

[15] He discusses his worries in the context of violations of the independence principle. However, they bear directly on the problems with Mom's preferences, since her preference order is a clear violation of the independence principle.

Independence: let g_1, g_2 and g_3 be any three alternative gambles. Then, $g_1 \geq g_2$ if and only if $g_{13} \geq g_{23}$. Where $g_{ij}=[g_i, p; g_j, 1-p]$ is a complex gamble with probability p to be exposed to g_i and probability $1-p$ to be exposed to g_j (where $0 \leq p \leq 1$).

Mom's preferences, [P, 0.5; J, 0.5]>P~J, can be rewritten as: [P, 0.5; J, 0.5]>[P, 0.5; P, 0.5]~[J, 0.5; J, 0.5]. Independence requires that the preference order over two lotteries is determined by those outcomes of each lottery that are different. Since Mom prefers [P, 0.5; J, 0.5] to [P, 0.5; P, 0.5] it must be the case that she prefers J to P since those are the only outcomes that differ. But then she ought to strictly prefer J=[J, 0.5; J, 0.5] to P=[P, 0.5; P, 0.5] instead of being indifferent. As a result, Mom's preferences violate the independence principle.

Note that this preference pattern does not just violate the strong independence condition above, it also violates a whole family of weaker conditions related to independence. E.g., Mom's preferences violate the axiom of *substitution* which stipulates that:

$\mathbf{o} \sim \mathbf{o}' \leftrightarrow g=[o_1, o_2, \ldots, \mathbf{o}, \ldots, o_n] \sim g'=[o_1, o_2, \ldots, \mathbf{o}', \ldots, o_n]$ (Luce and Raiffa, 1957).

[16] Broome, 1991, pp. 110–15.

https://doi.org/10.1017/S0266267101000220 Published online by Cambridge University Press

case' is a true proposition in one of the outcomes of the lottery. If fairness somehow supervenes on such descriptions, it could be argued that we have no particular reason to worry about violations of consequentialism (assuming, of course, that fairness is a 'rational justifier'). This is certainly an attractive idea. It would manage to save, within a strictly consequentialist framework, the notion that certain values depend on the process by which they are produced.

However, I am not convinced, because this way of thinking about fairness does not work in the dynamic context. In the tree of Figure 2, the normal form representation of Mom's predicament, Broome's idea makes good sense. The outcome 'P *in a fair manner*' is a correct description of that outcome because it was achieved through a fair procedure (i.e., the flip of a fair coin), and this procedure could have resulted in another, equally fair outcome (i.e., 'J *in a fair manner*'). If the outcome P contains references to both the procedure and the alternative outcomes, it is correctly identified as a fair outcome. The fairness will supervene on these references. As a result we have avoided the worry that fairness emerges in the interaction of states of nature and have reduced it to aspects of the outcome.

Whereas Broome's suggestion seems to work in the normal form, it runs into a paradox when we look at the dynamic form representation of Mom's choices. In Figure 1, there are four end nodes with the following outcomes: P, J, P and J. We identified three possible plans, A, B and C. Plan A has the prospect [P, prob(*heads*); J prob(*tails*)], plan B has the prospect [P, prob(*heads*); P prob(*tails*)] (which reduces to [P]) and plan C the prospect [J].

If Broome is correct in claiming that fairness supervenes on the individual outcomes of each prospect, we can re-describe the prospect of plan A as follows: [P *in a fair manner*, prob(*heads*); J *in a fair manner*, prob(*tails*)]. Fairness supervenes on P and J here because P and J are the result of a fair procedure that gives equal weight to the other possible outcome. As a result, so Broome would argue, we are entitled to identify P as 'P *in a fair manner*'. Similarly for J so the result is the tree of Figure 3. But now look at what happens to the prospect of plan B. Plan B (go 'up' if heads P) now has the prospect [P, prob(*tails*); P *in a fair manner*, prob(*heads*)]. Apart from the fact that this prospect does not obviously reduce to the prospect [P], we have a problem with the supervenience of fairness. Above, it was argued that fairness supervenes on an outcome if it is the result of a procedure that gives equal weight to all possible outcomes *and* if that procedure could have resulted in another outcome. Clearly, as the prospect of plan B shows, that is not the case here so fairness cannot supervene. Therefore, if fairness supervenes on the outcomes in Figure 1, we get the paradoxical conclusion that we are both entitled and are not entitled to ascribe fairness to the outcome P.

https://doi.org/10.1017/S0266267101000220 Published online by Cambridge University Press

Is there a way to avoid this paradox? One answer would be to argue that there is no fairness in Figure 1 at all.[17] The idea would be that an outcome is fair if and only if it is the *direct* result of a fair procedure without a choice of Mom standing in between the procedure and the final outcome. More precisely, what is relevant to Mom is that her children perceive her decision as a fair one. This means that she has to avoid the impression that there is some element of arbitrariness on her part. That can only be achieved if the outcome of the procedure is 'out of her hands' so to speak. I find this reply counter-intuitive at best. Surely, if Mom were to give the treat to Jane in n_2 this is fair. The fact that she has made a decision to do so does not change this fairness since Peter has had his chance. More precisely, Mom's choice in n_2 is fair if and only if the fact that Peter has had his chance is part of Mom's considerations in favor of Jane. This means that it is not the mere fact that Mom has a choice that determines the presence or absence of fairness, but the grounds for her choice that determine this. That means that there is a fair solution possible in Figure 1, and that is the solution that Mom has selected by adopting the plan that has the prospect [P, prob(*heads*); J, prob(*tails*)].

The obvious way to avoid the paradox is to argue that fairness does not supervene on individual outcomes but on lotteries. That means that fairness does emerge in the interaction between states of nature.[18] Broome's solution has a lot going for it because he limits himself to normal form situations where choices precede all intervening chance events. In such a context it is never the case that one and the same outcome, the same terminal node, can be part of two different prospects. One does not run into paradoxes of the type I described above.

6. CAN A CONSEQUENTIALIST APPEAL TO FAIRNESS?

Whatever we are to make of either one of Broome's strategies (i.e., the idea of rational justifiers and the special response in Section 5), it is

[17] Robyn Dawes and Teddy Seidenfeld in personal communication have both argued for this claim.

[18] Another suggestion to avoid the paradox is the following: the outcome 'P' should not just contain references to the procedure by which it was realized but also to the plan that selected it. Thus, depending on which plan was adopted, 'P' can be 'P, as a result of plan A' or 'P, as a result of plan B'. The argument proceeds by arguing that the choice of plan is crucial in the attribution of fairness. Since 'P' can be the result of plan A as well as plan B, it is fair if it is the result of the former, but not when it is the result of the latter.

In general, I believe we should resist the suggestion that plans themselves become part of the description of the outcomes because it leads us into a regress. The proper representation of the deliberation process would have to include a prior choice node which represents the plan that is to be chosen. But then we get 'meta-plans' as to which plan to select. If those meta-plans then also become part of the outcome descriptions, we have to represent these also as prior choice-nodes, which leads us to formulate 'meta-meta-plans', etc.

https://doi.org/10.1017/S0266267101000220 Published online by Cambridge University Press

important to note that both implicitly assume that the principles of consequentialism and expected utility theory in and of themselves have no *independent* restrictive normative content. Broome, like most modern consequentialists, clearly believes that consequentialism is neutral until a rationally acceptable domain of values has been identified.

In what follows I will ignore Broome's special response outlined in Section 5 and concentrate on the principle of rational justifiers as a way to save the principle of consequentialism from normative vacuity. In addition, I will abstract from our worries about the nature of the value theory that Broome's principle of rational justifiers implies and accept it as a reasonable constraint on re-describing outcomes. Let us accept also – which seems plausible – that fairness is a value that can make a rationally acceptable difference.

Unlike Hammond or Broome, I believe that consequentialism *by itself* puts serious substantial constraints on what values are acceptable. These restrictions are so severe that consequentialism is incompatible with plausible candidates for what can count as a rational justifier on Broome's theory. In other words, we should not only worry about the normative vacuity of consequentialism. In some cases we should also worry about the limitations that consequentialism brings us.

Consider the tree in Figure 3. What could it mean to claim that Jane received the treat '*in a fair manner*'? What makes this particular outcome fair? It is fair only in so far as Peter could have received the treat as a result from the same event.[19] That means that the re-description of an outcome in terms of fairness only makes sense in the presence of alternative outcomes, which can be reached through an alternative path through the tree. This is true for the tree in Figure 3. The outcome 'J *in a fair manner*' only makes sense in the presence of the branch leading to 'P *in a fair manner*'.

However, in so far as a commitment to consequentialism implies a commitment to separability, Mom is forbidden to take the branch 'tails' into account in her assessment of the outcomes following n_2. Therefore, she could not possibly identify J after n_2 as a fair outcome, regardless of whether the principle of rational justifiers allows her to make a distinction between 'J' and 'J *in a fair manner*'. As a result consequentialism does not allow Mom to appeal to fairness in order to justify her choice for Jane at n_2. Since NEC requires that Mom's choices in the extended tree and the normal-form reduction of her predicament are consistent, it follows that Mom cannot identify the lottery as a fair one in the normal form reduction of the situation. The application of both SEP

[19] What is more, the likelihood of him receiving the treat must be equal to the likelihood that Jane receives it. Therefore, Mom should use a *fair* coin.

https://doi.org/10.1017/S0266267101000220 Published online by Cambridge University Press

and NEC requires that Mom is strictly indifferent between all prospects. Her preference for the fair lottery is inadmissible. It appears that there is an unavoidable trilemma for the consequentialist.

On the first prong, Mom is forced to deny that fairness is a rationally acceptable value. She then violates the principle of rational justifiers and is indifferent between A, B, and C, thus saving NEC and SEP. That way her judgments about the admissibility of plans satisfy the requirements of consequentialism. However, it does so at a cost. If even a basic value such as fairness is not allowed to count as a rational difference between outcomes in some situations, we must ask the question how plausible is the doctrine that disallows this to be the case?

That brings us to the second prong, where we insist that fairness is a consideration a rational agent can take into account, but we reject SEP, thus rejecting a fundamental part of consequentialism. If we reject SEP, strictly preferring J at n_2 is an admissible plan continuation even though Mom would be indifferent in a *de novo* tree with just P and J as options.[20]

The third prong of the trilemma appears if we abandon NEC. That would lead us to the conclusion that while it is rational to prefer to flip a coin in the normal form representation, this cannot make a difference in Figure 1. This means that Mom cannot acknowledge the value of fairness in the dynamic case of Figure 1, whereas she can do so in the normal form reduction of that tree (Figure 2). Is there no real fairness in Figure 1? I submit that this is not the case. There is a fair way of going about things in both Figure 1 and Figure 2.

We can conclude that Mom's preferences leave us no way to save both the idea that fairness is a legitimate consideration and hold on to consequentialism. So either we give up on the principle of rational justifiers (first prong), or we abandon separability (second prong), or we reject the notion that fairness enters into the picture in Figure 1 (third prong). The latter two moves both lead us to the conclusion that consequentialism has to be rejected.

I have spelled out this trilemma for the value of fairness. I do not believe it is limited to just this particular value. There are implications for a whole class of values, dispositions and emotions, such as regret, disappointment, guilt and resentment, which all depend critically on judgments of 'what could have been'.[21] If my argument so far is correct,

[20] It could be argued that it is unclear whether Mom is really indifferent between P and J. Seidenfeld (1988) claims that indifference is betrayed by indifference to tiebreakers. Since Mom prefers to flip a fair coin to many other possible tiebreakers, she is not indifferent. However, it is also obvious that Mom does not prefer either P or J. I am not sure how to respond to this suggestion. Does this example show that Seidenfeld is mistaken? Or perhaps it indicates that P and J are unconnected since it is not the case that $P \geq J$ or that $J \geq P$. I hesitate to accept either conclusion and leave this matter unresolved here.

[21] It is interesting to note that Peter Hammond now has been persuaded of this point by

https://doi.org/10.1017/S0266267101000220 Published online by Cambridge University Press

the same trilemma I presented in this section for fairness can be constructed for each of these considerations. In conclusion, a commitment to consequentialism implies that one should reject these values from rational deliberation. A principle of choice that eliminates these from consideration cannot possibly be a plausible requirement of rationality.

7. CONSEQUENTIALISM RECONSIDERED: LEVI'S CRITICISM OF NEC

It is perfectly legitimate to question whether the two principles I laid out in Section 2 really capture what is implied by 'consequentialism' as it is employed in decision theory. I will not venture to give a complete consequentialist vindication of these principles. Instead I shall discuss and reject two arguments in this section and the next. The first argument aims to show that consequentialism does not imply NEC. The second argument, which I will discuss in section eight, makes the same claim with respect to SEP. Refuting arguments against a position does not amount to a knock-down argument in favor of that position. However, it is the best I can do here to avoid the charge that the result of the previous section is determined by an arbitrarily chosen definition of consequentialism.

First, let us consider NEC. A consequentialist is committed to realizing the best (feasible) outcomes. Unless the process of decision making itself intrinsically affects either the outcomes or their feasibility it is hard to see how the value of the outcomes changes in a tree in comparison with the same choice in normal form. It might be argued that there is a fundamental difference between dynamic choice and normal form choice, since in the case of the former the agent could be confronted with new, unanticipated information. Clearly, if there is such new information a consequentialist might have good reasons for choosing differently in comparison to the situation where he had chosen *ex ante* on a fixed course of action (as is the case in the normal form). However, there is no new information in the case of Mom's predicament. She foresees and deliberates about all the possible events (i.e., the outcomes of the coin flip). Under such (admittedly rare) conditions the equivalence of dynamic form and normal form is self-evident for the consequentialist. Should the order of choice and chance nodes make a difference to an agent in such a situation, we would be correct in claiming that this agent would allow considerations other than the

Munier and others. In Hammond (1996) he explicitly states that his proof only applies when '. . . the consequences themselves do not depend on the structure of the tree. *Of course such independence has been the standard assumption in classical decision theory*'. [Italics added]. We can interpret my entire argument, therefore, as an illustration of the implications of this 'independence'.

https://doi.org/10.1017/S0266267101000220 Published online by Cambridge University Press

evaluation of the prospects of each plan to determine her choice.[22] It is, therefore, only natural to suppose that a consequentialist should opt for the same prospect, whether she has to make one, two or more decisions to get there.

However, Isaac Levi (1986, 1991) has gone on record as claiming that NEC is not required by consequentialism. He has several arguments for these claims. The most fundamental one I will call the argument from availability. Levi argues that one should make a strict distinction between what is an available outcome and what is an admissible outcome.[23] Consequentialism, just like the theory of rational choice, is supposed to determine which outcome is admissible of the available ones. However, it cannot be used to determine the available outcomes. It follows, according to Levi, that if one takes this distinction seriously, one cannot maintain that in Figure 1 Mom has the same prospects available at n_1 as in the normal form reduction of that tree, Figure 2.

His argument is deceptively simple. Suppose Mom believes she has the prospect [J, tails; P, heads] available at n_1. This implies that she knows *for sure* that at reaching n_2, she would choose J.[24] What warrants this belief? Levi argues that it is unwarranted unless one can predict with certainty that this is how one will choose. However, such a prediction is incompatible with the belief that the agent also has the prospect [P] available at n_1 since this implies that one will choose P in n_2. All Mom can predict, therefore, is that she will realize an opportunity for choice should tails come up. Therefore, it is simply not true that Mom has the same prospects available in Figure 1 and the tree of Figure 2. Arguing that she does, confuses the availability of a prospect with the admissibility of that prospect. Therefore, consequentialism cannot imply NEC.[25]

[22] For a discussion of choice behavior that is sensitive to the time of dissolution of uncertainty see Kreps and Porteus, 1978. This makes sense in certain contexts, e.g., if one is uncertain with regard to one's future income, one might prefer this uncertainty to be dissolved as soon as possible, so as to make appropriate financial arrangements. However, this type of consideration does not play a role in Mom's situation. Her preference is for fairness, not for a timely dissolution of uncertainty.

[23] Levi, 1974, 1991, 1992. He uses the term 'feasible' rather than 'available'. His main reason for this preference is his claim that the possible outcomes an agent could realize do not just depend on what is 'objectively given' to the agent. They also depend on the beliefs of the agent. The intuition is that if you believe that you cannot realize A (even if in fact you could), you should not consider A as one of the feasible outcomes.

[24] I wonder if Levi sets up a straw man here? Is it really necessary for Mom to be *sure* that she will choose J at n_2? It seems to me that this is too strong. In order for Mom to believe that [J prob(*tails*); P prob(*heads*)] is available, it is enough that she believes that the choice for J at n_2 cannot be ruled out at n_1.

[25] Levi (1980, 1986) also believes NEC cannot be part of any plausible theory of rational choice. His main reason is the following. Standard expected utility theory requires that the agent can attribute determinate numerical values to the probabilities of all possible

https://doi.org/10.1017/S0266267101000220 Published online by Cambridge University Press

If Levi is correct in his conclusion, the argument from availability has far-reaching implications. The most disturbing one from the point of view of a traditional understanding of consequentialism, is the implication that consequentialism should no longer be understood as a necessarily maximizing form of deliberation. This is anathema to the traditional consequentialist. For her, consequentialism is the requirement to choose the *best* feasible prospect.

In order to appreciate why the rejection of NEC forces one to abandon maximization of value in some situations, consider the example of *the minimax regret chooser*.[26] Imagine an agent who orders her prospects so as to minimize her possible regret. She determines for each prospect what could have happened under the same conditioning event had she chosen otherwise. Suppose this person faces a decision tree with three possible prospects, A, B and C. Prospect A is a lottery which, depending on certain events (E1, E2, or E3), will either give her \$10, \$5, or \$4. B stands for a lottery that gives \$2, \$9, \$5, under the same conditioning events. C, finally, will give \$8, \$0, or \$10 under those events.

Suppose she lacks all knowledge of the likelihood of any of these events. In that case she will decide to take the course of action that will minimize her maximal regret. In order to determine this, she looks for each alternative that she could have had under the conditioning event had she chosen otherwise. The difference is the amount of regret of that alternative given that event. She does this for each alternative and each event and then she determines the maximum possible regret. Next, she opts for the alternative with the smallest maximal regret. In Figure 4, I have illustrated the process of comparison in the case where this agent compares all three alternatives. Since A has the least possible maximal regret, this is the best option for this agent, when comparing all three lotteries.

This way of comparing alternatives is highly problematic for an agent dedicated to realizing the best possible result because the minimax

states of affairs. (Regardless of whether these numerical values are objective probabilities, as in the case of von Neumann–Morgenstern utility theory, or subjective probabilities, as in the case of Savage utility theory.) Levi argues that this is often not possible or desirable. This has led him to construct an index that allows for both indeterminate probabilities as well as indeterminate value indices, the notion of *e-admissibility*. Roughly speaking, an outcome is e-admissible if its expected utility is high enough on *some* value of its related probability interval. In case there is more than one e-admissible outcome available (which is virtually always the case if we are dealing with multi-attribute decision making), a tiebreaker is called for. Depending on the kind of decision one is facing, Levi favors maximin when dealing with benefits, whereas minimax is the appropriate tiebreaker in the case of expected losses. If one measures the value of an outcome in this way, one's value judgements will result in planning that sometimes violates NEC.

[26] Savage, 1972, Chapter 9.

https://doi.org/10.1017/S0266267101000220 Published online by Cambridge University Press

	E_1	E_2	E_3	*regret*	E_1	E_2	E_3	*Maximum regret*
A	\$10	\$5	\$4	**A**	0	4	6	6
B	\$2	\$9	\$5	**B**	8	0	5	8
C	\$8	\$0	\$10	**C**	2	9	0	9

FIGURE 4. Calculating the maximum possible regret when comparing A, B, and C.

regret chooser has intransitive preferences over these three lotteries. Transitivity requires that the agent's preferences are consistent in the following manner. For any three outcomes x, y and z: if $x \geq y$ & $y \geq z$ then $x \geq z$.[27] The minimax regret chooser will judge that A is better than B and B is better than C. However, if she were to compare just A and C, C turns out to be better than A.[28] The minimax regret chooser has intransitive preferences. Clearly, if one's preferences are intransitive in this manner there is no 'best' prospect. Consequently, one cannot maximize and this, I submit, is not acceptable to a traditional understanding of consequentialism.[29]

[27] Where '$\geq$' stands for 'is at least as good as'. An alternative interpretation of the choices of the minimax regret chooser does not use the concept of preferences, but instead takes the notion of a choice function as primitive. Thus interpreted, this agent does not use minimax regret to order the available alternatives from best to worst, but to pick out the admissible, or choiceworthy element(s) from the set. This interpretation would be closer to Levi, since he does not assume that a rational agent necessarily has a complete well-ordered preference ranking of the feasible alternatives. However, in the present example, this difference in interpretation is not very important. The predicament of the minimax regret chooser can be expressed just as forcibly with the language of choice functions. Let C(•) be the function that the minimax regret chooser applies. Her predicament then consists in the following observations: C(A, B, C)=A; C(A, B)=A; C(B, C)= B, whereas C(A, C)=C. Therefore, in this instance we can plausibly claim that her choice behavior reveals her preference over the alternatives. (Such a claim would not be plausible if we were dealing with choice functions over incommensurable options.)

[28] Note, however, that this agent's preferences in this particular example are acyclical. A preference order is cyclical if $x>y>z>x$, where '>' stands for 'is strictly better than'. That is not the case in this example, since A>B>C.

[29] It could be argued that I have too strict a view of consequentialism. A slightly weaker formulation would be a conditional one. That is, *if* there is an outcome, such that there is none better than it, then the agent should act so as to realize it. This re-formulation of the requirement to maximize might seem to answer my concerns here. Since the minimax regret chooser has intransitive preferences, there is no best prospect. Consequently, there is no violation of consequentialism if the agent chooses B or C. However, I do not believe that this suggestion helps us here. There is a best prospect overall, to wit A (since A>B>C). Even though A gets defeated in one of the pair-wise comparisons it is superior to all others in terms of the values of the minimax regret chooser. In other words, a consequentialist, even a conditional consequentialist, should act so as the realize A.

https://doi.org/10.1017/S0266267101000220 Published online by Cambridge University Press

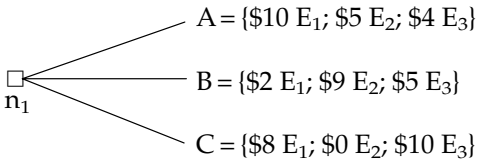


FIGURE 5

It turns out that a consequentialism, committed to NEC, should reject minimax regret because it is incompatible with NEC. Suppose the minimax regret chooser is offered the three lotteries as one up-front choice (see Figure 5). She would have no difficulties in settling on the plan leading to A.

However, if the same agent were to face the very same options in the tree of Figure 6, things seem to be different. At n_1, she would judge A the best prospect, but if she were to go 'down' she would be at n_2 facing a choice between A and C, in which case she would judge C the best prospect. Clearly, if NEC is required, such a divergence in selected plans is unacceptable. Without NEC, however, there is nothing to stop the agent from going about it in just this way. In other words, Levi would find no fault with the minimax regret chooser.

So what is the correct choice for the minimax regret chooser in this situation according to Levi? We can speculate about this. If we apply Levi's argument about availability and admissibility, one might judge prospect B admissible in this situation. Since the agent faces a choice between prospect B for sure or a further choice between A and C, she might just as well opt for B if she judges B better than such an opportunity for choice. Suppose, however, that the agent judges the further choice opportunity better than B. Then the agent will end up choosing C. As a result, C would be admissible. Similarly, if the choice opportunity were to be judged as good as B, both B and C would be admissible. Levi would argue that the minimax regret chooser will judge the choice opportunity inferior to B because it would lead to the selection of the worst outcome and that B is uniquely admissible here as a result. However, it should be noted that this claim only goes through if the agent is allowed to predict at n_1 how she will in fact choose at n_2. Since this is ruled out by the argument from availability one needs an additional assumption as to how the choice opportunity ranks in comparison to B. Without such an assumption, either B, C, or both, will be admissible.

Finally, one might question whether option A could be admissible at all in the tree in Figure 6 if we take Levi's argument seriously. At n_1 this agent does not have A available according to the argument from availability (it becomes available only after the agent has chosen to go

https://doi.org/10.1017/S0266267101000220 Published online by Cambridge University Press

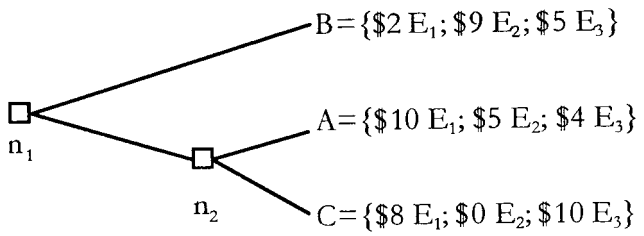


FIGURE 6

'down') and at n_2, when A is available, it is inadmissible. If this is correct, it means that the best overall prospect is inadmissible in this situation. Levi's argument leads us to conclude that B, C, or both, can be admissible in the tree of Figure 6. However, the uniquely best outcome overall, A, is not admissible here.[30]

It is important that I make myself absolutely clear here. I do not believe that the minimax regret chooser is irrational. Neither, I presume, does Levi. Her standard of evaluation is perhaps somewhat unusual since it leads to such intransitivities, but there is nothing in the notion of instrumental rationality as such that gives us grounds for accusing her of irrationality.[31] However, I do believe that the minimax regret chooser is acting in a non-consequentialist way. Her method of evaluation cannot be interpreted as maximizing with regard to her values.

8. CONSEQUENTIALISM RECONSIDERED: McCLENNEN'S CRITICISM OF SEP

This leaves us with determining the plausibility of SEP from a consequentialist point of view. The case for separability as a requirement of consequentialism is straightforward in my opinion. Consequentialism is a doctrine of forward-looking reasons. A consequentialist agent is concerned with what lies ahead of her. Given that this is the case, plan

[30] Isaac Levi, in his discussion of a similar problem involving two incommensurate values, explicitly endorses the choice for C, the worst alternative overall, and claims that the agent opting for C has stayed true to her values (Levi, 1991, pp. 111–12). In the example Levi discusses there, his claim is intuitively more plausible since his example involves a situation where the agent has incommensurable standards of evaluation. If one alternative is better considering standard X and another is better considering standard Y, both can be admissible given the appropriate choice of tiebreaker. However, in the example of the minimax regret chooser there are no incommensurabilities. There is only one standard for determining the acceptability of an outcome. Therefore, the intuitive case for the admissibility of C here is dubious at best.

[31] That is, there are circumstances imaginable where the agent cannot be required to have a (weak) ordering of her preferences.

https://doi.org/10.1017/S0266267101000220 Published online by Cambridge University Press

continuations should be evaluated in terms of their prospects, just as whole plans should be evaluated by their prospects.

Edward McClennen (1990) has argued that there are reasons to doubt that SEP is part of consequentialism. His main argument uses consequentialism, understood as the requirement to realize the best available outcomes, to argue against SEP. It can be shown that there are circumstances where consequentialists do worse by adhering to SEP than if they were to abandon it. The tree in Figure 6 illustrates this. Assuming that NEC is valid, the agent has the choice between A, B, and C at n_1. However, because she realizes, through her commitment to SEP, that she will choose C – the worst alternative – at n_2, she has to choose the second best option, to wit, B. However, suppose she did not adhere to SEP. Then she could plan to go 'down' and choose A at n_2 and at n_2 *resolutely* choose A. Because A is the best possible alternative for the minimax regret chooser, it seems that there are good consequentialist reasons to abandon SEP in situations like this.[32]

I am not convinced by this argument. McClennen's argument is an instance of what he calls 'pragmatic foundationalism'. The central claim of this form of foundationalism is that the content of a theory of rational choice is *fully* determined by what generates the best possible results. I believe that pragmatic foundationalism is incorrect. McClennen's argument is a good illustration of this. Consequentialism requires that one should choose that plan which realizes the best possible outcomes. This is the content of consequentialism. Consequentialism, therefore, is (part of) a decision procedure. It tells us how to decide. However, we cannot invoke considerations of consequentialism when we have to decide how we should decide. That decision cannot be based on consequentialist considerations since it would simply beg the question of the correctness of consequentialism. Such higher-order decisions should be based on what one believes to be correct. In general, one should believe what is true. The mere fact that it 'pays' to believe ϕ does not count as a reason for believing that ϕ is true. Similarly, the mere fact that it would sometimes 'pay' (i.e., realize desirable outcomes) to believe that one should not plan in accordance with SEP, is not a good reason for believing that consequentialism does not entail SEP. Additional arguments are needed for why SEP should be separated from the notion of consequentialism.

9. CONCLUSIONS

In this essay I have argued that there are reasonable preference orderings which violate consequentialism understood as a commitment to both

[32] This example is not quite felicitous. McClennen's own examples deal with situations where it is clear that the *ex ante* and *ex post* evaluations are in agreement as to which is the superior outcome. Arguably, in this situation that is not the case.

https://doi.org/10.1017/S0266267101000220 Published online by Cambridge University Press

NEC and SEP. I argued that one cannot avoid these violations by re-describing the outcomes, because the required re-descriptions are not compatible with consequentialism. As a result, this particular set of principles – whether properly labeled as 'consequentialism' or not – cannot be required by rationality, since it implies giving up the notion of instrumental rationality. That is, it is a move away from the notion that rationality is neutral with respect to the values of the agent. Whether that means we have to give up on consequentialist foundations altogether is but one of the many questions that the problem of the relevant description of outcomes provokes.

Finally, let me put my last card on the table without any real argument for it. If my conclusion is acceptable one could still try to recover parts of consequentialism in rational choice. One could reject NEC, like Levi, or one could reject SEP, like McClennen. Either way, we would have given up on consequentialism but at least we would have maintained a consequentialist 'flavor'. My intuition is that consequentialism is a mistake altogether. I believe that *both* NEC and SEP should be rejected as principles of rationality. NEC is too strong because it requires the agent to disregard all the considerations of the structure of the decision tree that cannot be reduced to the outcomes (hence the required equivalence between normal form and extended form). SEP is too strong because it requires the agent to disregard relevant parts of his knowledge of the past that cannot be reduced to the outcomes. I owe the reader arguments that explain, first, why such considerations and knowledge might be relevant and, secondly, why these cannot be reduced to the outcomes. In this essay I have only suggested an argument for the second claim.

REFERENCES

Broome, J. 1991. *Weighing Goods: Equality, Uncertainty and Time*. Blackwell

Chang, R. (ed). 1997. *Incommensurability, Incomparability, and Practical Reason*. Harvard University Press

Diamond, P. 1967. 'Cardinal welfare, individualistic ethics, and interpersonal comparisons: comment'. *Journal of Political Economy*, 75:765–66

Hammond, P. 1988. 'Consequentialist foundations for expected utility theory'. *Theory and Decision*, 25:25–78

Hammond, P. J. 1996. 'Consequentialism, structural rationality and game theory'. In *The Rational Foundations of Economic Behaviour: Proceedings of the IEA Conference Held in Turin, Italy*, pp. 25–42. K. J. Arrow, E. Colombatto, M. Perlman and C. Schmidt (eds.). MacMillan

Kreps, D. M. and E. L. Porteus. 1978. 'Temporal resolution of uncertainty and dynamic choice theory'. *Econometrica*, 46:185–200

Levi, I. 1974. Indeterminate probabilities. *Journal of Philosophy*, 71:391–418

Levi, I. 1980. *The Enterprise of Knowledge*. MIT Press

Levi, I. 1986. *Hard Choices*. Cambridge University Press

Levi, I. 1991. 'Consequentialism and sequential choice'. In *Foundations of Decision Theory*, pp. 92–122. M. Bacharach and S. Hurley (eds.). Blackwell

https://doi.org/10.1017/S0266267101000220 Published online by Cambridge University Press

Levi, I. 1992. 'Feasibility'. In *Knowledge, Belief and Strategic Interaction*, pp. 1–20. C. Bicchieri and M. L. Dalla Chiera (eds.). Cambridge University Press

Luce, R. D. and H. Raiffa. 1957. *Games and Decisions: Introduction and Critical Survey*. Wiley

Machina, M. 1989. Dynamic consistency and non-expected utility models of choice under uncertainty. *Journal of Economic Literature*, 27:1622–68

McClennen, E. F. 1988. 'Ordering and independence: a comment'. *Economics and Philosophy*, 4:298–308

McClennen, E. F. 1990. *Rationality and Dynamic Choice: Foundational Explorations*. Cambridge University Press

McClennen, E. F. 1997. 'Pragmatic rules and rationality'. *Philosophy and Public Affairs*, 26:210–58

Munier, B. 1996. 'Consequentialism, structural rationality and game theory: comment'. In *The Rational Foundations of Economic Behaviour: Proceedings of the IEA Conference Held in Turin, Italy*. K. J. Arrow, E. Colombatto, M. Perlman and C. Schmidt (eds.). MacMillan

Rabinowicz, W. 1995. 'To have one's cake and eat it too: sequential choice and expected-utility violations'. *Journal of Philosophy*, 92:586–620

Rabinowicz, W. 1997. 'On Seidenfeld's criticism of sophisticated violations of the independence axiom'. *Theory & Decision*, 43:279–92

Savage, L. J. 1972. *The Foundations of Statistics*. Dover

Scheffler, Samuel (ed.). 1988. *Consequentialism and Its Critics*. Oxford University Press

Seidenfeld, T. 1988. 'Decision theory without ''independence'' or without ''ordering'': what is the difference?' *Economics and Philosophy*, 4:267–90

Slote, M. 1989. *Beyond Optimizing: A Study of Rational Choice*. Harvard University Press

Tversky, A. 1975. 'A critique of expected utility theory: descriptive and normative issues'. *Erkenntnis*, 9:163–73

https://doi.org/10.1017/S0266267101000220 Published online by Cambridge University Press