

RAWLSIAN CONTRACTUALISM AND HEALTHCARE ALLOCATION: A RESPONSE TO TORBJÖRN TÄNNSJÖ

– Quinn Hiroshi Gibson –

Abstract: The consideration of the problem of healthcare allocation as a special case of distributive justice is especially alluring when we only consider consequentialist theories. I articulate here an alternative Rawlsian non-consequentialist theory which prioritizes the fairness of healthcare allocation procedures rather than directly setting distributive parameters. The theory in question stems from Rawlsian commitments that, it is argued, have a better Rawlsian pedigree than those considered as such by Tännsjö. The alternative framework is worthy of consideration on its own merits, but it also casts light on two related difficulties with Tännsjö’s approach: (i) the limits of his supposedly ecumenical methodology, which is revealed to be dialectically suspect and (ii) issues with the type of abstraction and idealization from actual judgements and preferences which the approach requires.

Keywords: Consequentialism, Rawls, Contractualism, healthcare, fairness, distributive justice

Published online: 26 May 2021

Introduction

It might seem that to ask about distributive justice when it comes to the allocation of healthcare is a relatively straightforward application of our concern with just distribution in general. My goal in this paper is to explore the limits of such an approach with reference to Rawlsian Contractualism. Torbjörn Tännsjö’s recent *Setting Health-Care Priorities* follows this approach. Although Tännsjö’s strategy is deliberately what I term *ecumenical*, I will argue that there is a key perspective missing that deserves to be heard. That is, Tännsjö is clearly trying to get to his practical conclusions with the weakest theoretical premises possible. He wants to cover all the bases, as he sees them. Accordingly, he argues that a number of striking conclusions – that we should de-prioritize marginal life-extending treatments and focus more on the care and cure of patients suffering from mental illness, for example – follow on *any* of three theories of distributive justice that he considers. Further, he claims that “the family of theories [discussed] consists of the most plausible theories about distributive justice we know of.”¹ My main contention will be that there is a view that he leaves off the table which not only deserves consideration, but also shows the limits of the approach just described. The view in question is broadly

Quinn Hiroshi Gibson
American University of Beirut
e-mail: qhgibson@berkeley.edu

¹ Tännsjö (2019): 5.

Rawlsian, one which tries to articulate “what we owe each other by way of assistance in cooperative healthcare schemes.”²

Tännsjö considers utilitarianism (with or without a prioritarian amendment), egalitarianism, and what he calls the ‘maximin/leximin’ theory. It is worth noting at the outset that, as Tännsjö understands these views, they are all variants of hedonic or welfare consequentialism, with or without the addition of weightings or constraints. Allow me to begin by briefly describing how Tännsjö understands these three views.

Tännsjö sorts the views he considers with respect to how they stand on two diagnostic issues (which he considers to be metaphysical): (a) whether they accept or reject the “separateness of persons and the integrity of lives” and (b) whether they accept that “compensation” can occur between lives, or must occur, if it is to occur, within lives.

Tännsjö reveals his preference for utilitarianism early on, saying, “I happen to believe that utilitarianism is true.” However, he also acknowledges that the strongest challenges to utilitarianism come from those who take issue with what utilitarians say about (a) and (b). He continues, “I realize that strong objections have been made against utilitarianism from the point of view of adherents of the maximin/leximin theory and more generally from adherents of egalitarian thought. The claim has been that utilitarianism cannot in a satisfactory manner handle distributive issues. It doesn’t acknowledge the moral importance of differences between persons [...]”³ The version of utilitarianism that Tännsjö prefers is classical hedonic act utilitarianism. Of course, this form of the view accepts that losses to some *can* be (indeed ought to be, if the compensation is large enough) compensated for by gains to others. Because on this version of utilitarianism there are no side constraints that place limits on when and in what manner those losses, if sufficiently compensated elsewhere, may be imposed, one might be tempted to say that such a view *rejects* the separateness and the integrity of persons *by* insisting on the possibility of compensation *between* persons.

Of course, one can (attempt to) respect, in some form, the integrity and separateness of persons and also accept that compensation can occur between persons by placing some constraints on how the losses and gains can be distributed – that is, by rejecting the value monism that is characteristic of utilitarianism. This is how Tännsjö understands egalitarianism, viz., as a form of constrained consequentialism which acknowledges not only the total hedonic value in a state of affairs but applies a moral weighting to each person’s hedonic status depending on how they fare in relation to others in the distribution. As Tännsjö puts it, “we can see the value of an outcome as the sum of what each gets (her additional happiness) weighted with a separable factor defined with reference to where the person is as compared to where others are in terms of whole-life happiness.”⁴

Tännsjö’s characterization of the maximin/leximin theory requires a more extensive commentary. As one might expect, it is supposed to be the view that accepts the separateness and integrity of persons (unlike utilitarianism) and *rejects* the idea of compensation between persons (unlike egalitarianism). He says, “I will interpret the

² Daniels (2001): 3. Much of what follows is consonant with Daniels’ elaboration of a Rawlsian conception of healthcare ethics.

³ Tännsjö (2019).

⁴ *Ibidem*: 30.

maximin/leximin theory as the view that we should give absolute priority to those who are worst off. We should maximize the happiness of the person who is worst off.”⁵ It is the *maximin* view because of what it says about the worst off; *leximin* comes in to tell us that when we compare situations in which the worst off are at the same level, we consider the level of the next least badly off and pick the distribution on which *that* person is best off. That is, leximin allows us to prefer (+6, +6, +2) to (+55, +5, +2) where maximin, on its own (allegedly – I will return to this below), will not allow this.

Rawls on the Difference Principle

Obviously, maximin/leximin is loosely Rawlsian in inspiration, yet it is unclear what exactly its relation is to Rawls’ considered view of justice. To be fair, Tännsjö is clear that he doesn’t intend maximin/leximin to be an accurate reflection of Rawls’ conception of justice as fairness. He says:

Here a caveat about Rawls is in place. I described his rationale behind the maximin/leximin theory but glossed over a problem with Rawls himself. He does not seem to conceive his view of justice as a basic moral theory. In particular, he is not prepared to apply it to individual cases of the sort discussed here.⁶

But he also claims an important kind of Rawlsian provenance for maximin/leximin, viz., what it has to say about his diagnostic concerns, (a) and (b):

However, if we want to be true to the rationale behind Rawls’ theory of justice, we *ought* to apply it to individual persons in the manner I do in this book. Unless we do so, we are not true to the rationale behind the theory, to wit, the idea of the integrity and separateness of persons, followed up by the idea that we cannot compensate the loss suffered by one person with a gain made by some other person.⁷

So, Tännsjö takes the rationale for some key Rawlsian ideas about justice to involve the separateness of persons and a prohibition on compensation for one person’s loss by another’s gain. These are indeed important Rawlsian ideas but they are not the rationale for anything like maximin or leximin. I wish to take this as a starting point for articulating a genuine non-consequentialist alternative to all three of the theories that Tännsjö considers.

For one thing, what Tännsjö refers to as “maximin” is somewhat imprecisely named. When Rawls speaks of “maximin,” he is concerned with a certain risk-averse decision rule that it is rational to apply under certain conditions of uncertainty. It is distinct from the Difference Principle (DP), which partly composes one of the principles of justice that, Rawls argues, the parties in the Original Position (OP) would select. The maximin rule simply says that when confronted with a decision under uncertainty, one

⁵ Ibidem: 22.

⁶ Ibidem: 28.

⁷ Ibidem.

should choose the option whose worst possible outcome is the least bad of any of the worst outcomes that might eventuate. Maximin says, in other words, to be very risk averse, viz., to guarantee that the worst outcome of a particular decision be the least bad that it can be. This is not a rule of practical rationality that Rawls thinks we should normally live by, but he does think that it is the rule that the parties in the OP should adopt because of the unique (or nearly unique) features of that situation.⁸ So, Rawls thinks that the reasoning that the parties in the OP would go through to select his two principles of justice is reasoning that proceeds through maximin. But it is not, even in situations where it is supposedly rational to apply it as a rule of decision, meant as a principle of justice or a principle of distribution.

Of course, the DP is (part of) a principle of justice, and that (or something like it) could be what Tännsjö really has in mind when he speaks of “maximin”. For example, he says, when discussing the reasons for adding leximin to maximin: “The way Rawls writes about his difference principle invites the thought that his sole interest is in those who are worst off, relatively speaking. Let us term this as the maximin theory.”⁹ I will take Tännsjö to be speaking of an ethical theory inspired by the DP when he speaks of maximin/leximin.

However, we still must (eventually) face what we merely earlier sidestepped: none of the principles of justice, in Rawls’ hands, are meant to compose a comprehensive moral theory. Rawls is clear that they are to apply to the institutions of what he calls the “basic structure”, viz., “the way in which the major social institutions distribute fundamental rights and duties and determine the division of advantages from social cooperation.”¹⁰ First, since Tännsjö has claimed that maximin/leximin rests on a Rawlsian rationale, and since I am proposing to articulate a genuine alternative to maximin/leximin which rests on (what I take to be) a rationale with a better claim to Rawlsian pedigree, it is worth investigating just what constitutes the rationale for the DP.

One major theme in Rawls is the idea that a theory of justice should “attempt to mitigate the arbitrariness of natural contingency and social fortune.”¹¹ He further says:

Once we decide to look for a conception of justice that nullifies the accidents of natural endowment and the contingencies of social circumstance as counters in quest for political and economic advantage, we are led to [the principles of justice]. They express the result of leaving aside those aspects of the social world that seem arbitrary from a moral point of view.¹²

⁸ Those features are: (i) the parties don’t know the probabilities of the various outcomes they are considering (this information is denied to them in the OP), (ii) they are not interested in gains to welfare beyond a certain point, and (iii) some of the possible alternatives (ones where the liberties referred to in the first principle of justice are not guaranteed) involve outcomes that are unacceptable (Rawls 1971: 152-155).

⁹ Tännsjö (2019): 26.

¹⁰ Rawls, (1971): 7. Rawls, of course, has been challenged on his insistence that the basic structure is the “primary subject of justice” (1971: 3), e.g. by Cohen (1997), who has argued that restricting the scope of justice in this way leads to tolerating inequalities that seem unjustifiable given the actual rationale for the principles of justice, and for the DP in particular.

¹¹ Rawls (1971): 96.

¹² Ibidem: 15.

But, as Cohen has pointed out,¹³ there are ways of cashing out this concern which are more and less egalitarian. On one reading, Rawls is here pointing out that standard causes of *inequality* have a morally arbitrary basis. On this reading, it is unfair that some people *have more than others* when that difference is due to such causes. This is an egalitarian idea: Comparisons between people matter, and equality is the moral baseline.

If this were our concern, we could argue for the DP by claiming that inequalities which otherwise would not be justified are justified because they make possible an otherwise impossible *strict* benefit. We could then be tempted to formulate the DP somehow like this:

Intuitive Formulation: "Inequalities are forbidden unless they render the worst off better off."¹⁴

As Cohen says,

the thought behind the insistence in the familiar [intuitive] form of the principle that inequality must benefit the worst off, is that inequality is (at least *prima facie*) unfair, specifically to those at the bottom of the inequality, but that it would be absurd to let a concern for those people dictate a prohibition on an inequality from which they benefit.¹⁵

But, Cohen thinks, there is another way that the DP is often stated, which is not as egalitarian, and which appears to rest on a different rationale. Consider the DP in its more familiar

Canonical Formulation: Inequalities are justified only if "they are to the greatest benefit of the least advantaged"¹⁶ members of society.

Suppose we are comparing state of affairs (5, 8) to state of affairs (5, 10). It could be thought that since the least well off cannot do better than 5 on the available options, both of the options where they are at 5 count as being "to the greatest advantage" of the least well off. That condition being satisfied, we would then be permitted to allow the more advantaged to be even better off without any violation of the DP. Here, we do not take a greater inequality as something that needs to be compensated by greater benefits to the least well off. Rather, we might think that it is "justified on the different basis that, if something is good for some and bad for none, then it is to be endorsed."¹⁷

As Cohen points out, these two formulations rest on different rationales. Which of these two interpretations more closely matches Rawls' own reasons for endorsing the DP? As we have seen, Tännsjö's addition of *leximin* to *maximin* yields a theory that allows us to select (5, 10) over (5, 8). In discussing this, Tännsjö cites a passage from Rawls where Rawls is discussing a situation where "close-knitness" fails, *viz.*, situations where benefits to the most advantaged make no difference to the least advantaged (though they make a difference to those in between). In the relevant passage, Rawls appears

¹³ Cohen (2008): Ch. 4.

¹⁴ *Ibidem*: 157.

¹⁵ *Ibidem*: 157-158.

¹⁶ Rawls (1971): 83.

¹⁷ Cohen (2008): 158.

to endorse the Pareto-optimality argument (for the canonical formulation) that Cohen attributes to him.¹⁸ On the other hand, Cohen does seem correct when he accuses Rawls of frequently slipping into speaking as though the benefit to the least advantaged is meant as a *compensation* for the concomitant increase in inequality (and thus suggesting that the benefit must be strict).

Notice that while it is Cohen's point that Rawls can't have the DP both ways – as he puts it, the rationale for the intuitive formulation “contradicts” the DP's “content”¹⁹ – either of Cohen's rationales seems compatible with what Tännsjö claims is the “meta-physical foundation” on which the DP rests. Here is Tännsjö explaining that foundation again:

One aspect of the ‘distinction’ between persons is the putative separateness of our personal lives. If it is a fact, then it renders reasonable the maximin/leximin theory, it is assumed. *The crucial assumption is that we cannot compensate for the loss suffered by one person with reference to a larger gain made by another.*²⁰

Even if *losses* to one person can't be compensated by gains to another, interpersonal comparisons could still come in to play a role justifying the position of the least well off *either* by guaranteeing that gains to the better off are accompanied by a corresponding gain to the least well off, even if smaller (Cohen's rationale for the intuitive formulation), *or* on Pareto-optimality grounds (Cohen's justification for the canonical formulation).

I think there is a more fundamental rationale for the DP, one which helps to provide a Rawlsian response to Cohen, but is also a better candidate than the one that Tännsjö offers. Indeed, I think that Tännsjö and Cohen are guilty of similar misunderstandings of Rawls on the DP, as I hope the alternative candidate for Rawls' rationale will make clear.

This more fundamental rationale also takes seriously Rawls' earlier-mentioned concern with the influence of factors that are arbitrary from the moral point of view. Consider the following passage from Rawls:

[T]he difference principle expresses a conception of reciprocity. It is a principle of mutual benefit. [...] *The social order can be justified to everyone*, and in particular to those who are least favored; and in this sense it is egalitarian. [...] Consider any two representative men A and B, and let B be the one who is less favored. [...] Now B can accept A's being better off since A's advantages have been gained in ways that improve B's prospects. [...] The difficulty is to show that A has no grounds for complaint. Perhaps he is required to have less than he might since his having more would result in some loss to B. Now what can be said to the more favored man? To begin with, it is clear that the well-being of each *depends on a scheme of social cooperation without which no one could have a satisfactory life*. Secondly, we can ask for the willing cooperation of everyone only if the terms of the scheme are reasonable.

¹⁸ Rawls (1971): 81-82.

¹⁹ Cohen (2008): 157.

²⁰ Tännsjö (2019): 23. Emphasis in original.

The difference principle, then, seems to be a fair basis on which those better endowed, or more fortunate in their social circumstances, could expect others to collaborate with them when some workable arrangement is a necessary condition of the good of all.²¹

There are a number of things to take away from this passage. First, Rawls here highlights the key contractualist idea that social arrangements need to be *justifiable to each*. To achieve that, we sometimes have to accept Pareto-suboptimal arrangements. He also brings out a point I have not yet made: that what we are concerned about when we are asking about distributive justice are the products of social cooperation, viz., what we produce together, on which we all have a *prima facie* claim. What is unfair, from Rawls' point of view, is that the claim we all have on what we could not produce except by working together should be *defeated* by factors that are arbitrary from the moral point of view. What the DP expresses is a commitment to mitigating the influence of those factors that are arbitrary from the moral point of view on one party's contingently greater efficiency at acquiring the products of social cooperation.

On this interpretation of Rawls, it is not that deviation from equality of some goods requires compensation by a strict benefit to the least well off, and this speaks against interpreting Rawls as offering the more egalitarian justification that Cohen associates with the intuitive formulation of the DP. But, because Rawls also seems to suggest that the DP will force us to accept precisely the kind of Pareto-suboptimal distributions that the rationale Cohen pairs with the canonical formulation does *not* seem to permit, that rationale also does not seem to be the correct one. Rather, it seems that, since the DP is meant as a principle of reciprocity in the sense of offering fair terms of cooperation to parties who have *prima facie* equal entitlement to the products of such cooperation, on this showing it is not, in the first instance, a principle of distribution (though of course it has distributive consequences). It is a principle expressing a commitment to the justifiability of social arrangements to each free and equal party to that arrangement.

The key idea here is that the DP cannot be reduced to a set of distributive parameters. It is one aspect of Rawls' commitment to the ideal of no one being – or being made to feel – left behind, of *being in this together*: “[L]iberty corresponds to the first principle, equality to the idea of equality in the first principle together with equality of fair opportunity, and *fraternity* to the difference principle.”²²

Cohen's challenge rests on the misconception that the DP must provide a clear, perfectly general answer to whether a move from (5, 8) to (5, 10) *and all analogous moves* are permissible, viz., whether there must be a strict benefit to the least well off in the case of any inequality at all. Tännsjö, since he does not claim to have given a definitive interpretation of the DP, does not make exactly the same mistake, but it should indicate to us that the apparent innocence of the above-quoted passage where Tännsjö proposes that the way to be most faithful to Rawls' rationale for the DP is to apply it to “individual cases” is merely apparent. Since his goal is not exegesis, this is not an interpretative sin, but it does seem to indicate that he seems mostly interested in discussing theories of the

²¹ Rawls (1971): 102-103. My emphasis.

²² Ibidem: 106. My emphasis.

sort that Cohen wrongly assumes Rawls is on the hook to give. Seeing that Rawls is not required to give such a theory helps us to see why a Rawlsian alternative is a genuine alternative.

Further, I do not *deny* Tännsjö's claim that Rawls is committed to the separateness of persons and the impermissibility of compensating losses to one by gains to another. Nor do I deny that these commitments are playing some role in Rawls' defense of the DP. But these commitments alone underdetermine the thick role that Rawls assigns to the DP. How do we get to a principle of reciprocity in the above sense simply given a commitment to the separateness of persons and the impermissibility of interpersonal compensation? The parties in the OP have the resources to get themselves there, but that is because Rawls has also built in a host of other commitments, ones which, of course, he thinks *we*, as actual members of liberal democratic societies, in fact have. If parties in the OP had only the metaphysical commitments that Tännsjö identifies, it is far from obvious that they would choose the two principles of justice. Indeed, it is possible that, given other assumptions, they would choose what Tännsjö calls maximin/leximin, a view that is designed to directly output distributive parameters. *But that's not what Rawls' parties would choose.* And that's why maximin/leximin isn't a Rawlsian theory.

A genuinely Rawlsian theory, one which at its core held the contractualist commitment to the idea that basic social policies should be *justifiable to each*, would be structurally different from any of the three views that Tännsjö considers. As mentioned above, all three of those views are forms of hedonic or welfare consequentialism. They thus seem particularly well-suited to a straightforward application to first-order distributive questions, such as those that arise in connection with healthcare allocation. But it would be a mistake to infer from this that all the plausible theories of distributive justice apply to such questions in a similar way.

Indeed, on the kind of alternative I am envisioning, we are not likely to hold that every feature of a distribution, specified at an arbitrary level of grain, or even that distributive decisions, be required to meet the bar of justifiability to each. Rawls is clear on this:

It is a mistake to focus attention on the varying relative positions of individuals and to require that every change, considered as a single transaction viewed in isolation, be in itself just. It is the arrangement of the basic structure which is to be judged, and judged from a general point of view.²³

There are two things that need to be addressed about this point. First, are the institutions responsible for the provision and allocation of healthcare within the basic structure? This, I submit, is not clear. For example, there is controversy in the literature over whether the "official" criterion for being part of the basic structure is having effects that are "profound and pervasive"²⁴ or having the coercive power of law.²⁵ This is not a controversy I can settle presently, but I will simply note that even if we include a society's main healthcare institutions within the primary ambit of justice, there is still

²³ Ibidem: 87.

²⁴ Ibidem: 96.

²⁵ See, e.g., Cohen (1997).

an observable gap between the *site* of justice, as conceived by a Rawlsian contractualist, and first-order questions of distribution that is not observable on the consequentialist views that Tännsjö considers.

Second, one might simply wonder whether the fact that a theory doesn't *immediately* yield answers to first-order distributive questions shouldn't just count as a mark against the theory, or, at least, should mean that such a theory can be safely set aside when discussion turns to applied issues. I think this would be a mistake. I devote the remainder of this paper to the defense of this claim.

Healthcare and Procedural Fairness

It is common in discussion of healthcare allocation to emphasize scarcity. The empirical facts here are undeniable: regardless of how one chooses to measure it, there is far more "demand" for healthcare than is effectively provided, and existing healthcare systems are inefficient, in some places woefully so. Tännsjö takes the present inefficiencies to be so great that they account for the entire gap between what we are currently getting for our money and the best that we could expect to get for *any* amount of money. He thinks that if we were to have an ideal level of compliance with any of the three theories he considers, then "no more resources [would have] to be invested in health care other than what is now standard practice in developed Western nations."²⁶

Though the *fact* of scarcity ought to be agreed to by all parties, it is often doing unacknowledged theoretical work for theorists who appeal to it. The risk of this is especially large when we are considering healthcare allocation as a special case of distributive justice generally because, while the scarcity of healthcare is especially salient, all the goods whose distribution is the concern of distributive justice are scarce in the relevant sense. The very subject matter of justice seems constrained by the Humean "circumstances of justice," under which

[n]atural and other resources are not so abundant that schemes of cooperation become superfluous, nor are conditions so harsh that fruitful ventures must inevitably break down. While mutually advantageous arrangements are feasible, the benefits they yield fall short of the demands men put forward.²⁷

In the discussion of healthcare allocation, theorists have emphasized scarcity to argue for any number of (inconsistent) distributive criteria. For example, Rescher argued, decades ago, in a foundational article for bioethics as a nascent field, that scarcity means that we should determine who gets scarce medical resources by appeal to a patient's prospective service to society, as well as past services rendered²⁸ (something few would defend now); Singer has appealed to scarcity in defense of rationing by use of the Quality-Adjusted Life Year (QALY).²⁹ The argument is nearly always the same.

²⁶ Tännsjö (2019): 173.

²⁷ Rawls (1971): 127.

²⁸ Rescher (1969): 178-179.

²⁹ Singer (2009).

“There is scarcity, so we need *some* way of deciding.” For Rescher, we appeal to “probable prospective ‘return’ on [our] investment”³⁰ or we ask whether we have an “obligation in recognition and reward of services rendered.”³¹ For Singer, scarcity simply recommends the elimination of inefficiencies. I happen to think that if we are going to leap from scarcity to particular distributive criteria, the latter sort of move is much more defensible, but the more general point is that we should be suspicious of an overemphasis on scarcity; it may do nothing more than remind us that we are in the circumstances of justice, but it may also be a Trojan horse.

I certainly do not deny that our healthcare systems can and should be more efficient. But again, we should ask, why is the distribution of healthcare a different matter, in respect of scarcity, from the distribution of any other goods that are within the ambit of justice? If there is no such difference, and we think scarcity recommends, above all else, the reduction of inefficiency in healthcare provision and allocation, we should think the same for those other socially important goods. This is presumably what consequentialists would say about the matter. But we could just as easily run things the other way. Indeed, we may just as well be inclined to follow Daniels,³² who has argued that the good of healthcare is intimately connected with the good of equality of opportunity, which is not only scarce in the relevant sense, but is also among Rawls’ primary goods, the distribution of which is *not* to be regulated, in the first instance, by considerations of efficiency. So, if we are serious about considering alternatives to consequentialism when it comes to healthcare allocation, we do best not to overestimate what the fact of healthcare scarcity demonstrates. In other words, scarcity seems to point wherever a given theorist is already inclined to go, and, in particular, it is difficult to see how we can get any rational traction in favor of consequentialism from a mere acknowledgement of scarcity, unless we are already prepared to accept that the correct response to scarcity is to adopt an imperfective conception of the goals of healthcare provision.

The Rawlsian contractualism that I have been advocating can provide what one might call an alternative *perfective* framework. Rather than evaluating our healthcare priorities with reference to how inefficient they are at present, and how much more efficient they would be if we revised them to more closely resemble the ideals recommended by efficiency-based theories of distributive justice, we can evaluate our healthcare priorities with respect to whether they are arrived at through procedures which are *fair* and which embody our commitments to the freedom and equality of all persons. The primary site of inquiry, on such a framework, is the set of background conditions that either enable or impede the making of distributive decisions in accordance with fair procedures.

There is an important dialectical point here that bears directly on the overall shape of Tännsjö’s project. Unless we want to begin by simply assuming that the correct theory of distributive justice is welfare consequentialism, we can’t assume that we know what the currency is in terms of which the success conditions for our healthcare policies are to be specified. It is not just that the strategy of delivering an ecumenical argument beginning from a disjunction ranging over a number of views is limited in inverse proportion

³⁰ Rescher (1969): 178.

³¹ *Ibidem*: 179.

³² Daniels (2001).

to the variety of views that it ranges over. It seems, further, that we risk begging no small number of questions if we start by helping ourselves to a conception of the currency in terms of which the success conditions are given. It thus seems that a more faithful way of cashing out the ecumenical strategy that Tännsjö is proposing would be to leave unspecified, at the outset, whether such a currency is imperfectible and insatiable (as welfare is). Because procedural fairness is formally silent on what that currency is, the peril of overlooking a view focused on procedural fairness seems especially great.

Revisionism and Idealization

Tännsjö wants to move from abstract theory to real life questions in healthcare allocation, and the views that he chooses to consider are well-suited for that task. But if we center a concern with procedural fairness, another issue with the ecumenical strategy becomes apparent. Fair procedures are devised, and yield outcomes that are justifiable to each, only by taking into account people's actual judgements and preferences, or *suitable idealizations* of those actual judgements and preferences. The idea here is supposed to be that determinations of what is justifiable to each should not be held hostage to vice or irrationality, but that what counts as fair serves no purpose in providing such a justification if it is not tethered in some way by the actual facts of human psychology. The problem for Tännsjö is that many of the attitudes we find when we are in the weeds of actual human psychology are inconsistent with the conclusions of his book, those conclusions allegedly being converged on by the three theories he considers. On the face of it, this provides an obstacle to yielding a procedure that will be justifiable to each, since we can expect resistance from those who have the recalcitrant attitudes in question. Obviously, the trick is to provide some principled way of distinguishing between acceptable and unacceptable idealizations, but I fear that the ecumenical strategy is not connected in the right way with people's actual preferences and judgements to deliver a procedure that we can expect to be justifiable to each. I will close by briefly elaborating this point.

Tännsjö seems primarily concerned with arguing that it is no mark against a theory that intellectual assent to the theory, on its own, does not guarantee cognitive or behavioral conformity with the theory. But it is important to note that this does not commit Tännsjö to the idea that justice can float free of the facts of human nature.³³ Indeed, for the present purposes we can safely set aside that kind of "radically fact insensitive kind of 'pure' ideal theory,"³⁴ since the views that Tännsjö considers are forms of welfare consequentialism which derive their plausibility from the presumption that welfare is something that is valuable for human beings as they are (perhaps as they are necessarily). I am willing to concede that it will not be a mark against a theory, as such, that human beings who believe it, as irrational as they are, can't fully think and act in accordance with it. To this extent, Tännsjö and I are both following Rawls in first seeking to identify the correct *ideal theory* (which, for Rawls at least, is partly characterized by

³³ See, e.g., Cohen (2003) for an argument that not all normative principles can be "sensitive to fact." See Miller (2013) for a reply.

³⁴ Roebyns (2008): 343.

“strict compliance”³⁵) and only subsequently seeking to apply it to things as they are.³⁶ According to this conception of the relation between ideal and nonideal theory, ideal theory “identifies the objective at which nonideal theory ought to aim and thereby gives nonideal theory its ultimate target.”^{37 38}

Accordingly, I do not object that Tännsjö is engaged in idealization when one should do no such thing. There is an extensive and complex literature on ideal vs. non-ideal theory in political philosophy. For instance, as mentioned above, some take that debate to be about whether the bounds of justice are contained in the bounds of the possible. This is what is at issue when Cohen attacks Rawls’s theory for being *insufficiently* ideal. Another issue is whether an articulation of an ideal is required to measure and regulate the nonideal.³⁹ On both issues, I take Tännsjö to be *with Rawls*, and so I take them to be orthogonal to present purposes. What *is* presently relevant is the sense in which I think that Tännsjö’s approach is “not connected in the right way” with people’s actual preferences. The problem I see is that the scope of the alleged irrationality that he seeks to identify is in tension with the requirement that I have been pressing, *viz.*, that the theory yield fair procedures that are justifiable to each.

Rawls plainly makes use of abstractions from human irrationality. The OP choice problem is essentially the problem of what rational, mutually disinterested parties motivated to find principles of justice would decide. But *is it* an example of the irrationality of the kind we need to eliminate in the OP, to prefer, for instance, a marginal life-extending therapy to one’s imminent death plus some other treatment to someone else? Perhaps. But measured by what standard? If I am a utilitarian, then perhaps I would be guilty of internal inconsistency if I were to have such preferences. That is certainly one conception of irrationality. But no party in the OP is known to him- or herself to be a utilitarian. A party in the OP must take into account the possibility that they represent someone (or some group including someone) who has a different (and perhaps equally rational) ordering of preferences. When we eliminate irrationality for the parties in the OP, we can’t do so from the perspective of some comprehensive doctrine.

Some people possess attitudes and preferences that are in tension with Tännsjö’s first-order recommendations, e.g., people who opt for marginal life-extending therapies. *Whose* attitudes and preferences of that sort, among those who possess them, are allegedly irrational? I see two options: (a) those who intellectually accept the truth of one of the three theories that Tännsjö discusses or (b) everyone. If Tännsjö opts for (a), he will have an easier time showing that the relevant preferences are indeed irrational – though

³⁵ Rawls (1971): 245.

³⁶ Of course, talking about the very need for healthcare seems to involve abandoning some idealizations that Rawls himself makes, such as stipulating that the parties in the OP represent people who are “fully functional over a normal lifespan” (Daniels 2001: 3). But that is perfectly consistent with leaving others in place.

³⁷ Stemplowska and Swift (2012): 376.

³⁸ I thus adopt a different conception (and for obvious reasons Tännsjö must do the same) from Stemplowska’s conception of ideal theory as “theory that fails to issue recommendations for how to improve our society that are applicable for us here and now” (2008: 319).

³⁹ For instance, Sen (2006) argues that we only need to compare the relative justice and injustice of available alternatives.

perhaps only by helping himself to standards of rationality that are not in play on Rawls' picture. Even then, he will be left without any resources for arguing that those who *don't* affirm one of his theories should accept his conclusions as justifiable. If he opts for (b), which is the interpretation on which his conclusions have the largest scope, it will be a much larger challenge to show why the attitudes in question are irrational in the first place. In both cases, this will leave many people with seemingly perfectly sound bases for regarding Tännsjö's conclusions as unjustified.

This is connected with the larger worry about Tännsjö's approach that I have been pressing throughout. The goal of what I have been calling his "ecumenical" approach is to cast as wide a net as possible, but if the net is only meant to cover those who already adhere to one of the theories he discusses, many will be left out. On the other hand, one may read Tännsjö as wanting to "scope out" his conclusions from under the disjunction of the three theories. The idea then could be that even those who are not adherents of the theories discussed could construe the apparent convergence of those theories as meta-inductive evidence in favor of the concrete proposals Tännsjö advocates. But the more widely we cast the net, the more difficult it is to arrive at justifiability to each, because of the greater diversity of input views constraining the agreement. The goal of being ecumenical is thus in tension (as it is quite generally) with the goal of justifiability to each.

So, if we go with option (a), it might be easier to explain why the recalcitrant attitudes are irrational. It is relatively straightforward to see how run-of-the-mill human irrationality could play out when it comes to being faithful to an abstract theory that one intellectually accepts. For example, even patients who are convinced adherents of one of Tännsjö's three theories may self-servingly deny one or more of the minor premises required to get from the abstract philosophical commitment to a concrete claim about their own cases. So, while a patient may adhere to utilitarianism and in general believe that marginal life-extending procedures cannot be justified given their costs, given that such resources could provide more happiness if spent elsewhere, she may simply deny that those resources will make it to the people who could benefit from them *in this case*. What of the danger of bending the rule if it becomes known? Surely a single exception doesn't make a difference.⁴⁰ And so on. And as Goodin has helpfully pointed out, if we are interested in accounting for why people opt for marginal life extensions, there are other psychological sources as well, perhaps most obviously: the terminally ill patients who are candidates for marginal life extensions are present in front of the physicians, in contrast to the merely "statistical" patients that are invoked by the three theories.⁴¹

But, again, if we go with option (a), we have given no reason for thinking that everyone else's attitudes inconsistent with Tännsjö's conclusions are irrational, and we seriously limit the scope of the argument.

What if we go with option (b)? First, it is worth pointing out that is not so obvious that a welfare-consequentialist, for whom "preference satisfaction is after all the coin of the realm,"⁴² can dismiss so many (of so many) people's actual preferences. And it does seem that it should be the attitudes of *everyone* that are relevant. If we are concerned

⁴⁰ Tännsjö (2019): 178.

⁴¹ Goodin (2020): 5-7.

⁴² Ibidem: 8.

with justifiability to each, that is, each member of a society who is potentially affected by healthcare allocation policies, we are going to have to face the fact that *those people's judgements* don't converge on the priorities that Tännsjö argues three philosophical views do converge on. Even among the portion of the population that is sufficiently philosophically-minded, the ecumenical strategy attributes only a *disjunctive* judgement – viz., that utilitarianism (with or without a prioritarian amendment), egalitarianism, or maximin/leximin is true – to a *subset* of *that* set. What about the (presumably) rather large majority that has no view one way or the other on matters of abstract theory? Why should we expect facts about what follows from the disjunction of those three views to make any claim on what is justifiable to those who either have no view on the matter, or who would be, if asked, inclined to *reject* all three views?

We should not necessarily balk at a certain degree of revisionism, nor does it serve to deny that the scope of human irrationality really is quite vast. But if there are too many people whose attitudes have to be *classed* as irrational when deciding which attitudes need to be taken into account when designing fair procedures, it begins to seem a little like fairness itself is not going to get its due. As I have been arguing, that would be a mistake.

Conclusion

I have been advocating contractualism of a Rawlsian variety as applied to healthcare allocation. I have argued that, although Tännsjö's approach is aspirationally ecumenical, that approach, as he conceives it, has a number of drawbacks which in large part stem from its failure to range over non-consequentialist views such as the Rawlsian view I have been discussing. The first problem is simply that positions like the Rawlsian one are attractive alternatives to the forms of hedonic consequentialism that Tännsjö considers and ultimately prefers, and that executing the ecumenical strategy as Tännsjö does is simply dialectically suspect with respect to the non-consequentialist views that it conspicuously omits. Although Tännsjö does make some attempt to include Rawlsian ideas in the approaches that he canvases, I have argued that the view which he chooses to address under that heading does not adequately capture what is distinctive about a Rawlsian approach. Further, Tännsjö appears to be committed to a kind of revisionism that makes it seem unlikely that judgements of fairness – at least judgements of fairness made by those who are not already committed to Tännsjö's preferred range of views – will get their due.

References

- Cohen G.A. (1997), "Where the Action is: On the Site of Distributive Justice," *Philosophy and Public Affairs* 26 (1): 3-30.
- Cohen G.A. (2003), "Facts and Principles," *Philosophy and Public Affairs* 31 (3): 211-245.
- Cohen G.A. (2008), *Rescuing Justice and Equality*, Harvard University Press, Cambridge (MA).
- Daniels N. (2001), "Justice, Health, and Healthcare," *American Journal of Bioethics* 1 (2): 2-16.
- Goodin R. (2020), "Setting Health-Care Priorities: A Reply to Tännsjö," *Diametros*, September: 1-9. <https://doi.org/10.33392/diam.1597>

- Miller D. (2013), "Political philosophy for Earthlings," [in:] D. Miller, *Justice for Earthlings: Essays in Political Philosophy*, Cambridge University Press, Cambridge.
- Rawls J. (1971), *A Theory of Justice* (1st edition), Harvard University Press, Cambridge (MA).
- Rescher N. (1969), "The Allocation of Exotic Medical Lifesaving Therapy," *Ethics* 79 (3): 173-186.
- Roebyns I. (2008), "Ideal Theory in Theory and Practice," *Social Theory and Practice* 34 (3): 341-362.
- Sen A. (2006), "What Do We Want from a Theory of Justice?," *Journal of Philosophy* 103: 215-38.
- Singer P. (2009), "Why We Must Ration Health Care," *New York Times Magazine*. URL = <https://archive.nytimes.com/www.nytimes.com/2009/07/19/magazine/19health-care-t.html> [Accessed 27.12.21].
- Stemplowska Z. (2008), "What's Ideal about Ideal Theory?," *Social Theory and Practice* 34 (3): 319-340.
- Stemplowska Z., Swift A. (2012), "Ideal and Nonideal Theory," [in:] *The Oxford Handbook of Political Philosophy*, D. Estlund (ed.), Oxford University Press: 373-389.
- Tännsjö T. (2019), *Setting Health-Care Priorities: What Ethical Theories Tell Us*, Oxford University Press, Oxford.