

IS WELL-ORDERED SCIENCE A WELL-ORDERED SOCIETY?

DAN HICKS

1. INTRODUCTION

In a previous paper¹, I distinguished between science as theory and science as practice. Insofar as philosophy of science is concerned with science as practice, I said, philosophy of science is interested in giving a regulative ideal for the relationship between science and society. I went on to argue that one prevalent such ideal, which I called the ideal of autonomous science as practice, was objectionable, on the grounds that it was closely related to a controversial political philosophy.

Philip Kitcher has also argued against this ideal² and proposed an alternative, which he calls *well-ordered science*. Kitcher has described well-ordered science as being heavily influenced by the work of John Rawls, among other political philosophers.³ But he has not, as far as I know, explained this influence. In light of my previous paper and Kitcher's own characterisation of his theory, two questions come up:

- In what sense, if any, is well-ordered science 'Rawlsian'?
- Does well-ordered science fare any better than the ideal of autonomous science as practice?

In this paper, I will attempt to answer these two questions. I will start by giving a brief overview of well-ordered science, and presenting the framework from political philosophy within which I will locate it.

Date: May 11, 2008.

¹'On the ideal of autonomous science'

²Kitcher, 2001, especially chs. 3-9, 12, and 13

³*Ibid.*, 211 and 213

2. WELL-ORDERED SCIENCE

Kitcher's well-ordered science is presented in chapter 10 of *Science, truth, and democracy*⁴. Kitcher divides scientific inquiry into three stages: the choice of research programmes, the implementation of these programmes, and the development of technological applications of the resulting knowledge. While well-ordered science strictly speaking includes all three stages, I will focus here exclusively on the first stage. Well-ordered science functions almost identically in the first and last stages⁵ and Kitcher speaks only very briefly of the second stage, so focussing on the first stage is not out of line.

I will refer to this stage as the *research choice stage*, and the procedure well-ordered science gives for this stage the *research choice procedure*. The research choice procedure involves a group of *ideal deliberators* following a complex sequence of steps that results in a choice between available research programmes.⁶ As we shall see below, these steps attempt to incorporate both ethico-political values and objective expert expectations of the likelihood of success and application of the various programmes. As Kitcher does not explicitly number these steps, the interpreter has some freedom in numbering them. I will divide the research choice procedure into four distinct steps. Certain relevant details of the research choice procedure will be skipped in this brief overview. I will, however, give these due consideration in later sections of this paper.⁷

In the first step, the deliberators each individually form their own opinions about which potential outcomes of scientific inquiry are more and less valuable. Concerned about the potential for ignorant deliberators to divert research in less important or even undesirable directions, Kitcher requires that these opinions be *tutored personal preferences*. Individual, personal preferences must be subsumed within the framework of structures Kitcher calls *significance graphs*⁸. Significance graphs articulate

⁴Kitcher, 2001

⁵*Ibid.*, 121

⁶More specifically, the procedure results in an 'assignment of resources' (*op. cit.*) to the research programmes under consideration, not a simple binary choice between pursuing or not.

⁷The summary of the next few pages corresponds to *ibid.*, pp. 119-21.

⁸Cf. *ibid.*, ch. 6, pp. 63-82

the significance or importance of various avenues of research in terms of their contribution to generally-recognised practical and theoretical goals. We shall not go into detail here. The upshot, I think, is clear: at the end of the first step, each individual deliberator possesses a preliminary but reasonable ranked list of the various outcomes.⁹

The second step is interpersonal. The deliberators now exchange their preliminary ranked lists as well as the reasons they have given (in terms, again, of Kitcher's account of significance) for those rankings. After these lists and reasons have been exchanged deliberators debate the merits of various lists, and attempt to formulate a single ranked list on which all deliberators can agree. Ideally this agreement is achieved by consensus but in case of irreconcilable differences Kitcher allows the research choice procedure to temporarily maintain a plurality of ranked lists.

With one or more ranked lists of outcomes, in the third step the deliberators turn to 'experts' to determine 'the possibilities that particular scientific ventures might deliver what the ideal deliberators collectively want'¹⁰. Thus, for each outcome on the list(s), the experts identify research programmes – if any – that are likely to yield that outcome at various levels of funding. Should there be disagreement among the experts (or, indeed, disagreement among the deliberators about who counts as an expert), a plurality of lists associating research programmes with outcomes can be produced. These two (sets of) lists – one ranking the value of various outcomes, the other ranking the chances of success of various research programmes – are now turned over to an arbitrator, who formulates one or more policy proposals designed to maximise the outcomes expected.

Assuming a single list of quantitatively ranked outcomes and a single list of quantitative expected success rates for each research programme, the maximisation can be represented purely mathematically:¹¹ For each research programme p , outcome o , and assignment of resources m , the experts have provided a unique probability $\pi_{p,o,m}$ that p will yield o when m is invested in it. Similarly, the second stage lead to

⁹Note that the tutored personal preferences rank *outcomes* in terms of their desirability, and not *research programmes* that attempt to realise these outcomes.

¹⁰*Ibid.*, 119

¹¹This paragraph is my gloss, not a paraphrase of Kitcher.

the assignment of a unique number v_o of the value attached to o . The expectation value of p is then a function of m :

$$E_p(m) = \sum_o \pi_{p,o,m} v_o.$$

The arbitrator returns the allocation function(s) (that is, research policy proposal(s)) f that assigns resources $f(p)$ to each research programme p maximising the net expectation value

$$\sum_p E_p(f(p)).$$

Presumably there is also a funding constraint – some maximum amount of resources m_0 such that

$$\sum_p f(p) \leq m_0.$$

Should these quantitative assumptions not be met, Kitcher does not explain how the arbitrator’s job is to be done.

Finally, in the fourth step, the deliberators agree on exactly one of the policy proposals produced by the arbitrator. If the arbitrator has formulated exactly one, then the fourth step is moot. Otherwise the deliberators must engage each other in rational debate once again, achieving either consensus or a compromise agreement. The resulting policy proposal consists of a single set of principles, rules, or (in the language of the last paragraph) a function $f(p)$ for the distribution of resources to all the various research programmes under consideration.¹²

Kitcher goes on to define well-ordered science – at least, for the first stage of inquiry – as agreement with the outcomes of this research choice procedure:

For *perfectly* well-ordered science we require that there be institutions governing the practice of inquiry within the society that *invariably* lead to investigations that *coincide* in three respects with the judgments of ideal deliberators, representative of the distribution of the

¹²Note that, while Kitcher does talk about this first stage producing ‘distributions of resources’, his research choice procedure of course only produces *rules or principles* for distributing resources. The ideal deliberators do not actually pass out money or equipment to scientists (and not just because, as we will discuss below, they are not real). This is reflected in the other terms Kitcher often uses to describe the outcome of the procedure: ‘budgetary level[s]’, ‘research agenda[s]’, and ‘assignment[s] of resources’.

viewpoints in the society. First, at the stage of agenda-setting, the assignment of resources to projects is exactly the one that would be chosen through the process of ideal deliberation I have described.¹³

This concludes my overview of the research choice procedure in Kitcher's ideal of well-ordered science.

3. CONTRACTARIANISM

We turn now from philosophy of science to political philosophy. One very common and extremely influential approach to distributive justice is what I will call *contractarianism*. Contractarianism dates back at least to the debates over the legitimacy of the monarchy during and immediately after the English Revolution in the seventeenth century, including the work of John Locke¹⁴ and Thomas Hobbes¹⁵, and is a prominent approach today in, for example, the work of John Rawls¹⁶.

I will characterise as *contractarian* any theory of distributive justice which is based, either implicitly or explicitly, on the following *contractarian thesis*:

A distribution of goods is just if and only if it is compatible with principles that are the outcome of an ideal procedure in which all reasonable individuals come to an agreement in a context for rational deliberation characterised by features of equality and fairness.

This is an intricate definition, and analysing it will be the primary task of the remainder of this paper. For now, simply note that it captures the important basic features of a 'state of nature'/'social contract' approach to distributive justice: a story is told in which roughly equal individuals come together to arrive at a set of principles that will form the basis for their life together in civil society; then, because these principles have the 'consent of the governed', it is claimed that they are legitimate.

The primary claim of this paper is that Kitcher's ideal of well-ordered science – due to the rôle played by the research choice procedure presented in the last section in

¹³*Ibid.*, 122, his emphasis

¹⁴Locke, 1689

¹⁵Hobbes, 1651

¹⁶Rawls, 1971/1990

particular – is contractarian. With this established, we can answer the two questions posed at the end of the introduction: Kitcher’s well-ordered science is Rawlsian because, like Rawls’ just as fairness, it is contractarian; and just as autonomous science as practice was objectionable because libertarianism is objectionable, well-ordered science will be objectionable or not to the extent that justice as fairness is objectionable or not.

My argument for my primary claim is entirely straightforward: I show that Kitcher’s research choice procedure exemplifies the features of the right-hand side of the contractarian thesis. There is, however, also a need for me to show that the contractarian thesis is an adequate definition of contractarianism. I will therefore proceed as follows: Each of the next several sections of this paper will consider one element of the contractarian thesis. I will first comment on the element, and then identify it in Rawls’ *A theory of justice*¹⁷ and Kitcher’s ideal of well-ordered science¹⁸. As *A theory of justice* is an explicit development of the contractarian approach to distributive justice¹⁹, if the key elements of the contractarian thesis are found in the key elements of *A theory of justice*, it is reasonable to conclude that the contractarian thesis is an adequate definition of contractarianism. Then, if the key elements of the contractarian thesis are also found in Kitcher’s well-ordered science, it follows that Kitcher’s well-ordered science is contractarian. I will close my paper by briefly considering some implications of this conclusion.

I will assume that the reader is familiar with at least the general structure of Rawls’ account of justice as fairness – the original position, the veil of ignorance, and so on.

4. ‘IS COMPATIBLE WITH PRINCIPLES’

A set of principles for distributive justice describe the way some goods should be distributed. If a community acts (or would act) according to some set of principles for the distribution of goods – tries to distribute goods in the way the principles

¹⁷Rawls, 1971/1990; citations to Rawls that are otherwise unattributed will be to this book. Unless otherwise noted, emphasis in these quotations is mine.

¹⁸Similarly, citations to Kitcher otherwise unattributed will be to Kitcher, 2001, and emphasis not noted as Kitcher’s is mine.

¹⁹10

say they should be distributed – call those principles the *operational principles* of the distribution. In contractarianism as I have defined it, justice is determined by examining the distribution for compatibility with a certain set of principles – the *ideal principles*. I have not, however, defined compatibility.

Distinguish between *strict compatibility* and *formal compatibility* in the following way: A given distribution of goods is strictly compatible with some ideal principles if and only if the operational principles of the distribution *are in fact* the ideal principles. And a given distribution is formally compatible with some ideal principles if and only if, if the operational principles had been the ideal principles, then, *ceteris paribus*, the distribution would have been produced. Under strict compatibility, the justice of a distribution is determined by examining the principles actually used to realise it. Strict compatibility is satisfied only if the distribution was achieved ‘in the right way’. Strict compatibility is therefore a ‘deontological’ understanding of compatibility. Under formal compatibility, by contrast, the justice of a distribution is determined by examining the distribution, and only the distribution. The operational principles are not considered at all. Formal compatibility is satisfied only if the distribution ‘gets the right result’. Formal compatibility is therefore a ‘consequentialist’ understanding of compatibility. My definition of contractarianism is compatible with both deontological and consequentialist understandings of compatibility.

Rawls describes his vision of a well-ordered society in deontological terms:

Now a well-ordered society is also regulated by its public conception of justice. This fact implies that its members have a strong and normally effective desire to act as the principles of justice require. Since a well-ordered society endures over time, its conception of justice is presumably stable: that is, *when institutions are just (as defined by this conception), those taking part in these arrangements acquire the corresponding sense of justice and desire to do their part in maintaining them.*²⁰

²⁰398

It is not good enough, Rawls thinks, that the right distribution of goods be achieved using ‘suspect’ principles. Noble Lies – widely-accepted falsehoods that, by being widely accepted, help maintain order and stability, including lies about the relative merit of individuals, or religious dogma – are incompatible with Rawls’ theory of justice.

Kitcher’s well-ordered science initially appears consequentialist:

The thought is that, *however inquiry proceeds*, we want it to *match the outcomes* those complex procedures would achieve at the points I’ve indicated So the challenge is to find institutions that generate *roughly the right results*, even though we have no ideal deliberators to make the instantaneous decisions we hope to replicate.²¹

However, the institutions which distribute goods are not the same thing as the operational principles by which goods are distributed. As we shall see below, both Kitcher and Rawls think that their complex *procedures* for producing ideal principles cannot be actualised. But this does not mean that the resulting *principles* cannot be operationalised. The fact that the ideal principles are *justified* by being the outcome of an ideal procedure does not imply that the ideal principles cannot be *used* as the operational principles of real-world agents. Indeed, Kitcher appears to think that the next step in developing his account of well-ordered science is to understand the ways people fail to act according to the ideal principles:

Once the ideal of well-ordered science is recognized, there’s an important need for a political theory of science that will consider the various ways in which the interests of actors and social institutions might easily divert us from the outcomes that would be reached in a state of well-ordered science.²²

The consequentialist reading of the first quotation above was built on a confusion between the ideal principles – the outcomes of the research choice procedure – and the research choice procedure itself. Kitcher’s wording is partly the problem – it is not clear, either from the quoted text or its context, whether the outcomes to be

²¹123

²²133, n. 8

matched are the ideal principles (the outcome of the research choice procedure) or the particular distribution of resources (the outcome of operationalising the ideal principles). But this means the former reading is entirely consistent. Just as Rawls thinks the difference principle can be operationalised without the deliberation in the original position ever having been carried out, Kitcher thinks that the distributive principles produced by the research choice procedure can be operationalised without the research choice procedure ever being actualised.

Hence, both Kitcher and Rawls have a deontological understanding of the compatibility of a distribution of goods with the ideal principles that are the outcome of their respective procedures.

5. 'AN IDEAL PROCEDURE'

As mentioned above, neither Kitcher nor Rawls require that the complex procedures by which they describe principles for distributive justice being generated need be either actual or even actualisable. In particular, Rawls says that

the original position is a *purely hypothetical* situation. Nothing resembling it *need ever take place*, although we can by deliberately following the constraints it expresses simulate the reflection of the parties
*What is necessary is that the principles that would be accepted play the requisite part in our moral reasoning and conduct.*²³

Similarly, Kitcher concedes that 'there's no thought that well-ordered science must *actually institute* the complicated discussions I've envisaged'²⁴. All that is required is that, by reflecting on the ideal procedure, we can arrive at ideal principles against which a distribution is judged to be just or unjust.

As we have already seen the implications of this point, I will not go into further detail here. It suffices to note that Rawls and Kitcher are parallel in this respect.

6. 'ALL REASONABLE INDIVIDUALS'

This noun phrase identifies the ideal deliberators – who it is that is supposed to produce the ideal principles. As this is a crucial element of a contractarian account

²³104

²⁴123

of distributive justice, it requires some careful analysis. In particular, each term in this phrase needs some significant commentary. However, the work each term is doing cannot be understood in isolation from the other two, so the reader should keep in mind that, in this section, we are really analysing the noun phrase ‘all reasonable individuals’ as a unit, with an eye to the total characterisation of the ideal deliberators.

6.1. ‘all’. No-one is excluded from the procedure – the consent of the governed is the consent of *all* of the governed. This is necessary to keep apologists for an unjust status quo from arguing that since reasonable individuals *from the privileged classes* would agree to the principles of distribution, the status quo is just.

On the other hand, there might be good reason to weaken the inclusion of all individuals to the inclusion of *representatives* of all individuals *in the given community*. Kitcher says that ‘The procedure I’ve outlined is indifferent as to whether we suppose one-to-one [direct] representation or proportional representation of groups with a common perspective’²⁵, while Rawls often speaks of the ‘representative man’ for various social and economic classes²⁶.

Even with ‘representative men’, it is clear that Rawls’ veil of ignorance is designed so that deliberators within the original position take every social and economic class into account: behind the veil of ignorance,

no one knows his²⁷ place in society, his class position or social status, nor does any one know his fortune in the distribution of natural assets and abilities, his intelligence, strength, and the like. I shall even assume that the parties do not know their conceptions of the good or their special psychological propensities . . . Since *all* are similarly situated and no one is able to design principles to favor his particular condition, the principles of justice are the result of a fair agreement or bargain.²⁸

²⁵123

²⁶Cf. 66

²⁷Despite the gendered pronoun, it is clear that Rawls thinks sex and gender can also be hidden behind the veil of ignorance. This also goes for matters of race and ethnicity.

²⁸11

Because all these features are obscured, individuals in the original position must consider themselves as potentially occupying any of these social and economic classes, and hence everyone is given due consideration.

Kitcher's tutored personal preferences and an assumed desire for consensus do similar work in the first two steps of his procedure, but in a different way. Ideal deliberators do not forget their identity or conception of the good, but instead, effectively, modify their conception of the good to reflect the interests of others:

They exchange their tutored personal preferences, explaining why they want particular outcomes to particular degrees and listening to the explanations given by others. In this process, I assume each is moved by respect for the preferences of others and aims to arrive at a consensus list in which none of the others is substantially underrepresented.²⁹

Thus, by virtue of having *tutored* preferences and striving to achieve an outcome that all find acceptable, each ideal deliberator takes into account the variety of interests had by individuals throughout all of society. Every social and economic class is taken into account in the research choice procedure.

6.2. **'reasonable'**. However, it's not strictly true that *all* individuals are included in the procedure. There is a crucial qualification. Only those who are competent judges of their own interests and can articulate those interests to other deliberators can participate in a procedure involving rational deliberation and the exchange of reasons for their preferences. As non-reasonable individuals cannot engage in rational deliberation or share reasons for their preferences, it would be functionally impossible to include them in such a procedure. Note that, historically, this qualification has been used to exclude women, children, non-landowners, slaves, non-human animals, the severely mentally impaired, and other individuals who, whether legitimately or not, were not considered 'reasonable'. As we shall see below, this *reasonability requirement* creates problems for both Rawls and Kitcher.

²⁹118

A neo-classical economic conception of rationality as the maximisation of one's ownership of goods (henceforth, the 'maximising conception' of the good) is an important assumption for Rawls. Behind the veil of ignorance, individuals are ignorant of their own conception of the good. They must therefore rely on the assumption that, whatever goods they value, they will probably prefer to have as much of them as possible.

I postulate that . . . they assume that they normally prefer more primary social goods rather than less. Of course, it may turn out, once the veil of ignorance is removed, that some of them for religious or other reasons may not, in fact, want more of these goods. But from the standpoint of the original position, it is rational for the parties to suppose that they do want a larger share Guided by the theory of the good and the general facts of moral psychology, their deliberations are no longer guesswork. They can make a rational decision in the ordinary sense.³⁰

This conception of rationality clearly functions as a restriction on the individuals who can participate in the original position: If some individual's non-maximising – that is, non-rational – conception of the good is *included* in the original position, that is, if h/er conception of the good is not obscured by the veil of ignorance, then 'guesswork' comes back into the deliberation procedure. On the other hand, if her conception is *excluded* by being obscured by the veil of ignorance, but s/he is still included among the deliberators, then the maximising conception must similarly be obscured to avoid question-begging. But then Rawls cannot use this crucial premiss (that rationality is a maximising conception) in arguing for the principles of justice. Hence all non-maximising conceptions of the good must be excluded from the original position *by not even being countenanced among the 'possible lives' deliberators may lead outside the original position*. These irrational individuals cannot be represented from within the original position, not even in the same way all socio-economic classes and other demographic categories are.

³⁰123

Kitcher's reasonability requirement comes in through his theory of significance. Recall, in the first step, when ideal deliberators form tutored personal preferences, they do these by recognising structures that Kitcher calls significance graphs. When reasoning with each other in the second step, they do so by displaying these same significance graphs:

Ideal deliberation must involve presenting the structure of significance graphs, where the multiform sources of significance are revealed They exchange their tutored personal preferences, *explaining* why they want particular outcomes to particular degrees³¹

Within significance graphs, research programmes acquire significance by being connected to 'broad questions'³². Individuals whose preferences cannot be subsumed under these broad questions are dismissed as 'pathological':

tolerance has its limits, and we do count some of our fellows as *pathological*, either because they obsess about trifles or because they are completely dull. In claiming that the sciences ultimately obtain their epistemic significance from the broad questions that express natural human curiosity, I am drawing on this practice of *limited tolerance*, on our conception of "healthy curiosity"³³

Since they are 'pathological', and their interests do not have any 'epistemic significance', these individuals are excluded from the research choice procedure by the requirement of tutored personal preferences.

For both Rawls and Kitcher, then, an individual can be included in the procedure for producing the principles of distribution if and only if s/he can integrate h/er preferences and (general) conception of the good into a public framework that allows individuals to exchange, explain, and understand each others' preferences and conception of the good. That is, if and only if s/he is reasonable. As, presumably, all adult human beings of roughly ordinary cognitive abilities are capable of this, Rawls and Kitcher retain the reasonability requirement while avoiding the racism,

³¹118

³²81

³³*Ibid.*

sexism, and classism of earlier contractarians. One might still be concerned about speciesism, nationalism, and ableism, however. But more on this later.

6.3. **‘individuals’**. Call a theory of the relationship between individual conceptions of the good or the right and community conceptions of the good or the right (distributive justice) *individualist* if individuals’ conceptions are treated as formally prior to the common conceptions. On the other hand, call a theory of such a relationship *communitarian* if the community conceptions are treated as formally prior to individuals’.³⁴

Contractarianism, as I have defined it, is individualist in this sense. The ideal procedure does not presuppose either a particular conception of the community good or a particular conception of distributive justice. Rather, it starts with *only* individuals’ conceptions of the good, and uses these to *develop* the conception of distributive justice. It is individuals, not priests, *Führer*, statesmen, philosopher-kings, or other interpreters of the community good who decide what counts as justice or the good life.

Furthermore, Rawls and Kitcher are both individualists in precisely this sense, and in exactly the same way as contractarianism is individualist. Rawls is quite explicit on both of these points:

justice as fairness, being a contract view, assumes that the principles of social choice, and so the principles of justice, are themselves the object of an original agreement the plurality of distinct persons with separate systems of ends is an essential feature of human societies. . . .³⁵

The principles of justice are not given *ab initio*, but instead are agreed-upon by – are the creation of – individuals with distinct and separate systems of ends (conceptions of the good). The individualism in Kitcher is clear in the first substantive sentence describing the research choice procedure: ‘I envisage *individuals* with *different initial preferences* coming together . . . to discuss the available courses for inquiry to

³⁴While these two are mutually exclusive, they are not, of course, jointly exhaustive. A third possibility might be called *dialectical*, in which individuals’ conceptions are informed by (possibly in reaction to) community conceptions, and community conceptions are at the same time constructed out of the individual conceptions of the community members.

³⁵25

pursue'³⁶. Again, the principles for the distribution of goods are not assumed at the beginning, but instead are the product of discussion and agreement among individuals.

There are two complications to be dealt with, however. One of Rawls' small but important problems is the type and strength of affective ties recognised within the original position. To arrive at what he calls the just savings principle, Rawls must postulate some sort of affective connection between generations. He considers a few possibilities:

We can adopt a motivation assumption [that is, an assumption about the motivations of deliberators] and think of the parties as representing a continuing line of claims. For example, we can assume that they are heads of families and therefore have a desire to further the well-being of at least their more immediate descendants If this is right, we will have succeeded in deriving duties to other generations from reasonable conditions.³⁷

But this is to say that, the conception of the good of a given deliberator is determined, at least in part, by the good *of h/er family* – that is, h/er conception of the good is *not* formally prior to those of the community to which s/he belongs.

This problem can be avoided by casting the affective tie as *shared interests* or *an overlapping conception of the good* rather than an appeal to a *community good*. In particular, I share a certain interest in the material welfare of my children with my children. By my own, individual lights, my welfare is increased if theirs is. The community (common) good we have as a family, if it is still theoretically necessary for the justification of the just savings principle, can then be identified as the intersection of all of our individual conceptions of the good.

Rawls' use of 'representative men'³⁸ for various social and economic classes can be dealt with in much the same way: by virtue of the fact that they occupy the same social and economic classes (or, to the extent that they do), we can expect the conceptions of the good of two waitresses, say, to overlap. Hence they can be

³⁶118

³⁷111

³⁸Cf. §6.1

represented within the original position by a single deliberator without appealing to a common good for social and economic classes. Kitcher's suggestion of 'proportional representation of groups with a common perspective'³⁹ is similarly not incompatible with individualism.

So, all together, both Kitcher and Rawls require all reasonable individuals to be involved in the deliberations leading to the ideal principles for the distribution of goods.

7. 'COME TO AN AGREEMENT'

On my definition of contractarianism, the procedure is monistic, resulting, on each occasion, in a single set of consistent distributive principles. (A pluralist outcome, with multiple, mutually inconsistent sets, is not an agreement; at best, it is disagreement without violence.) While agreement on this single set can be achieved by universal assent, universal consent, or compromise, there is still only one. This is necessary for the set of distributive principles to be operationalised – if two principles disagree with each other on how goods are to be distributed, then they cannot both be satisfied, and hence no distribution of goods can be just by being compatible with them. While this point is never, to my knowledge, made explicit by either Kitcher or Rawls, it is necessary for their deontological, strict understanding of compatibility.

8. 'A CONTEXT FOR RATIONAL DELIBERATION'

The principles are produced by the exchange of proposals backed up with publicly recognised reasons. Put negatively, agreement is not to be reached using threats of force (physical, economic, or otherwise) or other means of coercion. This is intuitively necessary for the agreement reached to be said to be consensual.

Kitcher seems to be getting at an ideal of rational deliberation using the language of explanations: in the second step, the ideal deliberators 'exchange their tutored personal preferences, explaining why they want particular outcomes to particular

³⁹123

degrees and listening to the explanations given by others'⁴⁰. Rawls' picture of the exchange within the original position is negative, but gets at the same idea:

The parties do not seek to confer benefits or to impose injuries on one another; they are not moved by affection or rancor. Nor do they try to gain relative to each other; they are not envious or vain. Put in terms of a game, we might say: they strive for as high an absolute score as possible. They do not wish a high or a low score for their opponents, nor do they seek to maximize or minimize the difference between their successes and those of others.⁴¹

These two descriptions bring together rationality, reasonability, and public reasoning. In Rawls, reasonable deliberators are maximisers of their anticipated individual good, and hence rationally deliberate with reference to this public notion of reasonability. In Kitcher, personal preferences are made reasonable by being replaced with 'tutored' personal preferences, and it is these that are exchanged by the deliberators – that is, personal preferences are made reasonable by being put in the context of public standards of significance.

9. 'FEATURES OF EQUALITY AND FAIRNESS'

The context for rational deliberation is not thrown together arbitrarily. It is designed so as to prevent certain intuitively undesirable principles and distributions from being produced (and hence justified) by the procedure. It is these features which are supposed to make the contractarian approach to distributive justice 'equal' and 'fair'.

In classical contractarianism, the context for rational deliberation is a state of nature that includes two features of equality and fairness. First, the rough equality between individuals in that state of nature:

Nature hath made men so equall, in the faculties of body, and mind; as that though there bee found one man sometimes manifestly stronger in body, or of quicker mind then another; yet when all is reckoned

⁴⁰118

⁴¹125

together, the difference between man, and man, is not so considerable, as that one man can thereupon claim to himselfe any benefit, to which another may not pretend, as well as he. For as to the strength of body, the weakest has strength enough to kill the strongest, either by secret machination, or by confederacy with others, that are in the same danger with himselfe.⁴²

And second, the consensual nature of the social contract, requiring all members of civil society to agree with the conditions of the contract (in some way or another; in contrast with the Common Law tradition, Hobbes believed that contracts made under a threat of violence were valid). Such laws of nature as might also be supposed to hold within the state of nature – such as a right to life and basic property rights – and that cannot be violated by the social contract are also features of equality and fairness.

9.1. **‘equality’**. Call a feature of a context for rational deliberation *of equality* or *equal* if all deliberators are thereby given equal standing within the context for rational deliberation. Everyone can present their views for others to consider, all votes carry the same weight, &c. Put negatively, the fact that one social group is more politically powerful than another, for example, does not mean that the former social group has any more *ab initio* standing within the context for rational deliberation than the latter. Note that this feature refers *only* to equality within the context for rational deliberation; a ‘fair inequality’ in the distribution produced by the resulting set of principles is entirely consistent with contractarianism. Such a feature might also be said to guarantee *equal standing* or *formal equality*, to distinguish it from a feature guaranteeing *material equality*.

Within Kitcher’s research choice procedure, equality is implicitly guaranteed by the procedures of consensus-building and majority voting. In the final step, for example, ‘Perhaps there is consensus among the ideal deliberators If that is not so, then the final resolution must be reached by majority vote’⁴³. While he does not explicitly describe how this voting is to be carried out, nor how consensus is

⁴²Hobbes, 1651, ch. 13

⁴³121

certified, I see no reason to think that he would say that some votes should count more than others. Hence I read him as implicitly including features of equality.

Rawls, as usual, is more explicit than Kitcher:

the parties in the original position are equal. That is, all have the same rights in the procedure for choosing principles; each can make proposals, submit reasons for their acceptance, and so on.⁴⁴

The original position is defined in such a way that it is a state of affairs in which the parties are equally represented as moral persons and the outcome is not conditioned by arbitrary contingencies or the relative balance of social forces.⁴⁵

Hence, in both Kitcher and Rawls, there is at least an implicit guarantee of equal standing.

9.2. **‘fairness’**. Call a feature *of fairness* or *fair* if, *ceteris paribus*, its inclusion in the context for rational deliberation decreases the likelihood of the procedure producing principles of distribution which have the effect of violating Rawls’ difference principle. Recall that the difference principle requires an unequal distribution to be to the benefit of the least advantaged under the proposed distribution, relative to a completely equal distribution.⁴⁶ A feature of fairness is therefore one which prevents a disadvantaged minority from being sacrificed to the advantage of a powerful majority.⁴⁷ A fair feature could thereby also be said to *prevent minority sacrifices*. Note that a fair feature being fair does not imply that it has been included *because* it is fair. Neither does being fair require *significantly* decreasing the likelihood of minority sacrifices. So long as the inclusion of a feature has the effect, intended or unintended, of preventing minority sacrifices to some extent, that feature is called fair. Thus, if debate is brought to an end by majority vote, but rules allow a minority of at least forty percent of the ideal deliberators to block a vote by staging a

⁴⁴17

⁴⁵104

⁴⁶Rawls’ official statement of the difference principle comes in part (a) of the Second Principle on 266, and is the primary topic of §§12-3, pp. 57-73.

⁴⁷By the inclusion of equal features, a minority cannot sacrifice a majority.

filibuster, then this is a fair feature. A requirement of universal assent among the ideal deliberators is an even stronger fair feature.

Rawls, of course, believes that the difference principle is one of the principles of distributive justice produced by his procedure. Hence some feature of the original position must be fair – though I will not take the time to identify which here. Kitcher includes a fair feature in the form of what I call the *principle of charitable engagement*:

each [deliberator] is moved by respect for the preferences of others and aims at a consensus list in which none of the others is substantially underrepresented. The deliberators are committed to seeing the others as having, like themselves, a claim to realize their aspirations, and thus to take seriously the others' descriptions of their preferences and predicaments and the rationales they provide for choosing as they do.⁴⁸

This principle has the clear result of preventing minority interests from being excessively sacrificed for the majority – the majority will get what they want, of course, but the minority will not be completely shut out.

It is probably this principle which Kitcher thinks will solve a problem he considers earlier in *Science, truth, and democracy*⁴⁹. We recognise that certain research programmes, while epistemologically upstanding, are quite likely to have harmful effects on the life chances of members of certain groups that, historically and currently, are subject to discrimination and prejudice. Should we undertake this research? Kitcher argues that, at least in the case that these expectations of harm are specifiable in certain ways, no, we should not undertake this research: ‘If we shouldn’t engage in ventures that can be expected to decrease the well-being of those who are already worse off than other members of society, we should therefore refrain from engaging in *S*’⁵⁰ He indicates⁵¹ that well-ordered science will not engage in these sorts of research programmes. The principle of charitable engagement, perhaps, explains

⁴⁸118-9

⁴⁹Cf. ch. 8

⁵⁰98. Note how closely the antecedent resembles the difference principle.

⁵¹108

why: recognising the strong interest members of these minority groups have in *not* pursuing this research, members of the majority will be disinclined to pursue it as well, and it will be low or even (should the scale allow) negatively ranked on the list of outcomes produced at the end of the second step.

Hence, both Kitcher and Rawls include features which help prevent minority sacrifices. They both, in my terminology, include features of fairness.

This concludes my discussion of the definition of contractarianism and how it is exemplified by Rawls and Kitcher. I have shown that both Rawls and Kitcher exemplify every major feature of the definition. As per the argument presented at the beginning of this discussion, I conclude that Kitcher, like Rawls, is a contractarian. This answers the first of the two questions with which I began this paper.

10. AN OBJECTION

But perhaps this conclusion is too fast. Perhaps I have defined contractarianism too broadly, so that Rawls and Kitcher both exemplify what I have defined as contractarianism, not by both really being contractarians, but by being part of two distinct approaches to distributive justice which happen to have some features in common.

Answering this worry would require a far more careful discussion of the major features of contractarianism than I can offer here. Indeed, just making this objection reasonable would require a far more careful discussion than I can offer here. Instead, I will just consider a significantly narrower variation. This narrower objection points out that the legislative branch of any contemporary representative democracy satisfies my definition of contractarianism: all adult members of the community elect representatives to the legislature, who then (at least, in theory) rationally debate various law and funding proposals.⁵² But contractarianism cannot be identified with representative democracy – the Haudenosaunee (Iroquois Confederacy) and Ancient Athenian democracies were representative, but certainly predate the development of contractarian political philosophy. Hence, the objection concludes, my definition of contractarianism is too broad. In particular, as Kitcher's research choice procedure

⁵²The other features of my definition are exemplified as well, but for the sake of space I will take this as obvious.

is clearly based on an ideal legislative procedure for a representative democracy, the common ground between Kitcher and Rawls is really just the common ground between representative democracy and contractarianism.

My response to this objection is to point to the history of modern representative democracy. As far as I know, the structure, function, and governmental rôle of the legislature of every modern representative democracy except one is based on either or both of two of the earliest modern representative legislatures: the Parliament of the United Kingdom (in which the legislature is the most powerful division of the state), and the Congress of the United States (in which the legislature is much weaker than the executive).⁵³ As is well known, the structure, function, and governmental rôle of both of these systems is heavily influenced by the work of John Locke, who was also one of the first (if not the first) liberal contractarians.

Hence, my response is actually to concede the argument completely, except for the final premiss that ‘contractarianism cannot be identified with representative democracy’. It is true that contractarianism cannot be identified with representative democracy *as such*. But it *can* be identified with representative democracy *as it is practiced today*.⁵⁴ For example, consider the way representatives are chosen. In the Ancient Athenian democracy, representatives were generally chosen using sortition – that is, using a lottery system, much like juries are selected in Common Law countries today. A pure sortition system is not contractarian as I have defined it, because it makes no effort to guarantee that all individuals are represented among the deliberators. Similarly, as pure sortition provides no guarantee that demographic minority interests will not be regularly sacrificed whenever a demographic majority gain a numerical advantage among the representatives (as will, statistically, usually be the case in a pure sortition system), a pure sortition system of representative democracy fails to be contractarian as I have defined it by failing to have features

⁵³The one exception is the Haudenosaunee. However, as the structure of the Haudenosaunee state bears only slight resemblance to the procedures of Kitcher or Rawls or to other representative democracies, I do not consider the Haudenosaunee a strong point against my claim. Indeed, the fact that the resemblance between contractarian-inspired representative democracies and Kitcher and Rawls is significantly greater than the resemblance between contractarian-inspired representative democracies and the Haudenosaunee state actually supports the claim I make in the next sentence and the following paragraph.

⁵⁴Again, except for the Haudenosaunee.

of fairness. But I know of no current national representative democracies that use sortition to choose legislators. All use an electoral system which, at least in theory, requires representatives to take all the various interests of their constituents into account.

11. IMPLICATIONS

I conclude this paper by identifying two implications of the claim that Kitcher's well-ordered science is contractarian. One implication is generally positive; the second is a serious problem for Kitcher.

First, since Kitcher's well-ordered science is contractarian, and indeed shares a number of features with Rawls' theory of justice, then it is compatible with contractarian approaches to distributive justice in political theory proper. It is Rawlsian at least to the extent that both well-ordered science and well-ordered society are highly egalitarian contractarian accounts of the way goods ought to be distributed. Rawls' justice as fairness seems to be generally regarded among philosophers as our current best theory of justice – although of course many philosophers have more or less strong disagreements with Rawls, so it would be false to say that there is agreement that Rawls' theory of justice is *the absolute* best. This makes Kitcher's well-ordered science much less objectionable as an ideal for science as practice than the ideal of autonomous science as practice was, at least with respect to mainstream views on political philosophy.

Second, Kitcher's well-ordered science is vulnerable to the same vulnerabilities as other contractarian theories. Martha Nussbaum has argued, for example, that contractarian political philosophy cannot give a proper account of justice with respect to beings that cannot be considered reasonable or rational or roughly equal to normal adult humans in wealthy nations. In my language, the restriction to all *reasonable* individuals and the procedure of *rational* deliberation mean the interests of beings not capable of reason and rationality cannot be represented within the procedure for generating principles for distributive justice. As Nussbaum puts the point,

In so conceiving of persons, Rawls omits from the situation of basic political choice the more extreme forms of need and dependency that human beings may experience, both physical and mental, and both permanent and temporary. This is no oversight: it is deliberate design Rawls recognizes the problem posed by the inclusion of citizens with unusual impairments, but he argues that this problem should be solved at a later stage, after basic political principles are already chosen.⁵⁵

And this problem is one, she thinks, contractarians cannot deal with: were the procedure to be opened to non-reasonable individuals, the context for deliberation could no longer be characterised by a rational exchange of public reasons.⁵⁶ It is not even clear to me how it could genuinely be said to be a context for deliberation – what else is deliberation but the social use of reason or ratiocination?

And Kitcher is vulnerable to exactly the same problem in exactly the same way. Consider a research programme into autism that is reasonably expected to yield therapies that will enable individuals with the most severe forms of autism to participate more in normal society, but is unlikely to yield any other significant benefits. One might think that interest in such a research programme can be based on the fact that the friends and family of severely autistic individuals share an interest with autistic individuals in the latter's participation in normal society. But Kitcher deliberately defines interests in an entirely subjective way, 'start[ing] from a subjectivist view of individual value (using personal preferences as the basis for an account of a person's welfare)⁵⁷. And it is not clear that severely autistic individuals really do have a subjective interest in increased participation in normal society. Their cognitive abilities may be so impaired that they do not have interests in any subjective sense at all. The *objective* interests of severely autistic individuals in increased participation in normal society cannot be taken into account in Kitcher's research choice procedure.

⁵⁵Nussbaum, 2006, 109

⁵⁶*Ibid.*, 123

⁵⁷116

Again, a particular outcome can only be presented within the research choice procedure if it can be related, via a significance graph, to ‘broad questions’ (or, presumably ‘broad social ends’) that are obviously of (epistemic or political) significance. Can research programmes into therapies for severely autistic individuals find a home in the network of significance graphs? While studying autism *as a human pathology* might have some significance (as one way of increasing our understanding of the relationship between cognitive, social, and linguistic abilities of ‘normal’ adult humans, for example), such pathological research programmes are significantly less likely to lead to therapies for severely autistic individuals than research programmes directly oriented at developing therapies. Furthermore, by framing autism as a pathology or abnormality rather than as an extreme within the range of the normal⁵⁸, pathological research programmes may lead to further devaluing of therapeutic research – better to do research that will benefit ‘normal’ humans than worry about a handful of ‘abnormal’ humans.

Nussbaum identifies similar problems for Rawls, with respect to non-human animals and persistent and pervasive inequalities between different nation-states. And these also apply, *mutatis mutandis*, to Kitcher’s well-ordered science. Kitcher is somewhat sensitive to the fact that global socio-economic inequities will mean that the citizens of wealthy nations – those nations undertaking the vast majority of scientific research – are unlikely to choose research programmes that will significantly improve the lives of the citizens of poor nations while bringing fairly minimal benefits to the citizens of wealthy nations.⁵⁹ His preliminary solution, however, is just to ‘restrict[] membership [in the set of ideal deliberators] to representatives of the citizens [of wealthy nations] but require[] them to become acquainted with the preferences of others’⁶⁰, rather than directly including citizens of poor nations or their representatives in the research choice procedure. But if this is sufficient to accommodate the interests of so-called distant persons, then presumably it is also

⁵⁸According to the CDC, ‘For 2002, of 407,578 children aged 8 years in the 14 surveillance areas, 2,685 (0.66%) were identified as having an ASD [autism spectrum disorder]’ (Rice, 2007, 14). This is roughly 1 in every 150 8-year-old children. Of course, only a small percentage of individuals with an ASD have the severe forms I am considering here.

⁵⁹124-6

⁶⁰*Op. cit.*

sufficient for the ideal deliberators to be the privileged and powerful within a single society, on the condition that they ‘become acquainted’ with the preferences of the less privileged and less powerful – geographically close but economically and politically distant – members of their own society. And this, in turn, seems incompatible with the formal equality that motivates contractarianism.

In short, there is a deep tension within contractarianism between the inclusiveness of *all* individuals and the tempered but still real elitism of only *reasonable* individuals. Both of these features clearly are essential for contractarianism; and yet it is not clear how they could be incorporated into a single account without making a problematic compromise. This is the case whether the contractarianism is for the philosophy of science or for political philosophy. I conclude that, to the extent – and only to the extent – that one accepts Rawls’ theory of justice as fairness as our current best theory of distributive justice, then it seems that one should accept Kitcher’s theory of well-ordered science as our current best theory of normative science policy. This answers the second of the two questions with which I began this paper.

REFERENCES

- Hobbes, T. (1651). *Leviathan*.
- Kitcher, P. (2001). *Science, truth, and democracy*. Oxford.
- Locke, J. (1689). *Two treatises of government*.
- Nussbaum, M. (2006). *Frontiers of justice*. Harvard.
- Rawls, J. (1990). *A theory of justice* (Revised ed.). Harvard. (Original work published 1971)
- Rice, C. (2007, February). Prevalence of autism spectrum disorders – Autism and Developmental Disabilities Monitoring Network, 14 sites, United States, 2002. *Mortality and morbidity weekly report*, 56(SS-1).