Theory Choice and Social Choice: Okasha versus Sen

Jacob Stegenga
University of Utah
jacob.stegenga@utah.edu

A platitude that took hold with Kuhn is that there can be several equally good ways of balancing theoretical virtues for theory choice. Okasha recently modelled theory choice using technical apparatus from the domain of social choice: famously, Arrow showed that no method of social choice can jointly satisfy four desiderata, and each of the desiderata in social choice has an analogue in theory choice. Okasha suggested that one can avoid the Arrow analogue for theory choice by employing a strategy used by Sen in social choice, namely, to enhance the information made available to the choice algorithms. I argue here that, despite Okasha's claims to the contrary, the information-enhancing strategy is not compelling in the domain of theory choice.

1. Introduction

In a recent article in this journal, Okasha (2011) draws an analogy between social choice — the amalgamation of individuals' preferences into a group choice — and theory choice — the consideration of multiple theoretical virtues to choose the best theory. In the domain of social choice, Arrow's theorem states that no method of social choice can satisfy four basic desiderata. Okasha's great insight was to note that given the analogy between social choice and theory choice, an impossibility theorem similar to Arrow's theorem holds for theory choice.

Salvation, Okasha suggests, comes from a strategy that Sen employed for avoiding the impossibility theorem for social choice, namely, to enhance the information made available to the choice algorithms. Sen (1970) showed that by relaxing certain constraints of one of the axioms (the ordinality and non-comparability constraints of the 'independence of irrelevant alternatives' axiom), Arrow's theorem can be avoided. Okasha is hopeful that the same strategy can be employed to avoid the impossibility theorem for theory choice. But his reasons for hope are not compelling. Here I critically evaluate Okasha's preferred strategy for avoiding the impossibility theorem. The information-enhancing

Mind, Vol. 124 . 493 . January 2015
doi:10.1093/mind/fzu180 Advance Access publication 4 February 2015

strategy cannot provide an escape from the impossibility theorem for theory choice.

2. Theory choice

Kuhn 1977 argued that five criteria of theory choice are accuracy, consistency, scope, simplicity, and fruitfulness. The insight of Okasha was to model such theoretical virtues as 'individuals' who have 'preferences' for theories. Consider the theoretical virtues 'accuracy' and 'simplicity', and their employment in comparing two ancient theories of the universe: the homocentric sphere theory of Eudoxus and the epicyclic theory of Ptolemy. The accuracy desideratum can be thought of as a voter (call her Ms Accuracy) who prefers the theory of Ptolemy over that of Eudoxus, and the simplicity desideratum can be thought of as another voter (call him Mr Simplicity) who prefers Eudoxus's theory over Ptolemy's. As in the social choice domain, these preferences are weak ordering relations (reflexive, transitive, and complete).

Okasha asks us to consider the following criteria that one might hope an algorithm for theory choice would satisfy. These include:

Non-Dictatorship (N):
There is no theoretical virtue such that if theory A is ranked above theory B by the theoretical virtue, then A is automatically ranked higher than B by the theory choice algorithm.

Independence of Irrelevant Alternatives (I):
The ordering of theories A and B by a theory choice algorithm can only depend on the theoretical virtues' orderings of A and B, and not on their orderings over other theories.[1]

Unrestricted Domain (U):
A theory choice algorithm must be able to generate an ordering of theories for all possible inputs of theory orderings by the theoretical virtues.

Unanimity (P):
If all theoretical virtues strictly order theory A over theory B, then a theory choice algorithm must order A over B.

[1] I paraphrase Okasha's formulation. Given the weak preference relation xR_iy (individual i weakly prefers x to y), a formal definition of **I** for social choice is:

If two profiles of individual preference orders $<R_1, \ldots, R_n>$ and $<R'_1, \ldots, R'_n>$ are such that for some given pair x,y of alternatives, for every individual i, xR_iy if and only if xR'_iy, and yR_ix if and only if yR'_ix, then the social choice rule, when applied to both profiles, must yield the same social preference for x over y, that is xRy if and only if $xR'y$, and yRx if and only if $yR'x$.

Given the analogy between social choice and theory choice, it is obvious that an analogy to Arrow's theorem holds for theory choice: no theory choice algorithm can jointly satisfy **N**, **I**, **U**, and **P**.[2] Okasha's argument is deep and insightful. His analogy between theory choice and social choice does not end with the analogue to Arrow's theorem, however. He extends the analogy by searching for an 'escape route' to the theorem which is similar to a well-known escape route to Arrow's theorem.

3. Escape

A standard response to any impossibility theorem is to question its axioms in order to find an 'escape route'. With what standard should one evaluate the axioms of the impossibility theorem for theory choice? Possibilities include:

- Normative adequacy (theory choice *ought* to satisfy the criterion)
- Complete descriptive adequacy (theory choice *always* satisfies the criterion)
- Partial descriptive adequacy (theory choice *sometimes* satisfies the criterion)

In the domain of social choice, a case can be made that Arrow's axioms are norms that should be satisfied, whether or not they are actually satisfied in any particular circumstance (this is controversial, but arguing this point here would take me astray).

In the domain of theory choice, the matter is not as straightforward. For reasons stated in Okasha 2011, a case can be made that **U**, **N**, and **P** are norms that always should be satisfied in theory choice.[3] These

[2] For several proofs of Arrow's Theorem, see Geanakoplos 2005, and for exposition of the axioms used in Arrow's Theorem, see Gaertner 2006.

[3] However, Weber (2011) suggests that, in fact, a dictatorship is acceptable for theory choice. The dictator that Weber proposes is 'fruitfulness'. Weber notes that Kuhn himself claimed that fruitfulness is the central reason why scientists adopt a paradigm. Nevertheless, no matter how important fruitfulness is, the other criteria like accuracy and simplicity are also obviously important. Weber claims that 'there is no reason why science should be committed to weight all theory choice criteria equally.' But this is not what **N** demands. **N** is the much weaker condition that even if the theory choice criteria are weighed unequally, the weightings are not *so* unequal that one theoretical virtue fully determines theory choice in all contexts.

conditions satisfy the standard of normative adequacy, because they are very basic norms of inductive inference.

Criterion **I** is more complex. One contribution of Sen 1970 was to show that Arrow's theorem can be avoided by permitting information which is richer than non-comparable preference orderings into a social choice algorithm. Since Okasha's optimism about avoiding the Arrow analogue for theory choice is based on this strategy of Sen's, I describe the basis of Sen's escape route before turning to the analogous escape routes for theory choice.

Sen suggests that a social choice rule ought to take into account a set of individuals' utility functions rather than mere preference orderings. A utility function is an individual's ascription of real numbers to choices; the numbers represent the utility that each choice would bring to the individual. A utility function is more informative than a preference order, since the latter can be derived from the former but not vice versa. Sen showed that to avoid Arrow's theorem, individuals' utility functions must be measured on scales richer than ordinal scales, and these utility functions must be meaningfully comparable between individuals. Thus, following Okasha's lead, it is helpful to analyse the independence of irrelevant alternatives criterion for social choice into three constitutive sub-conditions, which I will call Irrelevance of Alternatives $\mathbf{IA_s}$, Ordinality $\mathbf{O_s}$, and Non-Comparability $\mathbf{NC_s}$.

$\mathbf{O_s}$ and $\mathbf{NC_s}$ can be stated more precisely in terms of the permissibility of transformations to utility functions. $\mathbf{O_s}$ holds that any positive transformation of a utility function is admissible. Alternatively, if only positive linear transformations of a utility function were admissible, then utility would be measured on a cardinal scale; if, in addition to permitting positive linear transformations, there were a natural zero point to the utility function, then utility would be measured on a ratio scale. If only identity transformations were admissible, then utility would be measured on an absolute scale. If utilities are measured on a scale which is more informative than an ordinal scale, then $\mathbf{O_s}$ is not satisfied.

$\mathbf{NC_s}$ holds that within the class of admissible transformations delimited by the choice of scale (ordinal, cardinal, ratio, absolute), any transformation of an individual's utility function is admissible *independent* of any other individual's transformation of their utility function, and thus the output of a social choice rule could vary depending on such transformations. On the other hand, if utility *is* comparable between individuals, then the only permissible transformation to utility functions is that employed by all individuals: each individual must apply the same transformation to their utility functions (within the class of admissible transformations delimited by scale type), and thus the ranked output of the social choice rule would be invariant with respect to such transformations. In effect, to relax $\mathbf{NC_s}$ amounts to imposing an invariance requirement on the social choice rule: if utilities are comparable then the output of a social choice rule must be invariant to transformations of utility functions.[4]

To illustrate, suppose Alexa (A) and Beth (B) must choose between France (F) and Germany (G) for their holiday. They devise ways to measure their respective personal utilities, and find: $A_F = 1$ (Alexa derives 1 unit on her relevant utility scale if France is chosen), $A_G = 3$, $B_F = 2$, and $B_G = 1$. One possible rule for determining their holiday destination is, for all countries (C), determine the sum of $A_C + B_C$ and choose the largest. With this rule, Germany is chosen. To impose the invariance requirement is to stipulate that Alexa can transform her utility function only exactly as Beth does, and vice versa: if Beth multiplies each of her utilities by 2, then Alexa must do the same, and thus, given the stated choice rule, the choice of destination remains invariant. The invariance requirement places what might be seen as a needless constraint on Alexa and Beth, based on an unrealistic supposition that their respective utilities for their choice of country are comparable. Beth might in fact gain much more utility than Alexa does if they go to France, and gain the same amount of utility as Alexa does if they go to Germany, and thus Beth's utility function could be transformed by multiplying her initial utility ascriptions by, say, 3, without a similar transformation to Alexa's utility function (in which case, given the above rule, France would be chosen). To demand that the result of their choice rule be invariant to transformations of their utility functions amounts to barring such possibilities, and assumes that their initial measurements of their utility functions are meaningfully comparable.

Sen proved that to avoid Arrow's theorem it is insufficient to relax only $\mathbf{O_s}$. An Arrowian impossibility theorem can be proven if one relaxes only $\mathbf{O_s}$, but not if one also relaxes $\mathbf{NC_s}$. If utility functions are comparable, and measured on a supra-ordinal scale (say, a cardinal scale), then Arrow's theorem can be avoided. On the other hand, if utility functions are not comparable — that is, if $\mathbf{NC_s}$ is satisfied — then Arrow's impossibility theorem goes through.

[4] For a technical statement of this invariance requirement, see Sen 1977.

Let me return to the domain of theory choice. As above, it is helpful to analyse **I** into three constitutive sub-conditions: Irrelevance of Alternatives (**IA**), Ordinality (**O**), and Non-Comparability (**NC**). **IA** requires the output of a theory choice algorithm to be insensitive to how the theoretical virtues order theories outside the set of theories under consideration. **O** limits the kind of information regarding the support relation between theoretical virtues and theories to ordinal rankings. **NC** stipulates that the support that one theoretical virtue provides to a theory cannot be meaningfully compared to the support that another theoretical virtue provides to a theory (more formally: given a certain measurement scale, any transformation of the 'support measure' of a theoretical virtue is permissible, within the class of admissible transformations delimited by the scale type, independent of the transformation of a 'support measure' from another theoretical virtue, and the output of a theory choice rule can vary with respect to such transformations).

Of the three constitutive sub-conditions of **I**, only **IA** seems like a principled constraint, while **O** and **NC** do not reflect basic norms of inductive inference.

To motivate **IA**, consider this example. **IA** holds that how a theory choice algorithm ranks Copernican heliocentrism to Ptolemaic geocentrism should only depend on how the theoretical virtues rank Copernican heliocentrism to Ptolemaic geocentrism, and should not be sensitive to how the virtues rank Eudoxan geocentrism relative to Ptolemaic geocentrism or Copernican heliocentrism (loosely, the relative rankings of Ptolemaic geocentrism and Copernican heliocentrism should not change simply because Eudoxan geocentrism is also an alternative). **IA** is a desirable requirement for theory choice, and satisfies the standard of normative adequacy.

O and **NC** limit the kind of information that a theory choice algorithm can take into account, and *if* such information is available it would be irrational not to take it into account. If possible, theory choice ought *not* satisfy **O** and **NC**. By the standard of normative adequacy—and in contrast to the other criteria for theory choice algorithms—**O** and **NC** fail. *Is* such information ever available? As Okasha notes, we sometimes have absolute measures of some properties of theories, such as the number of free parameters in a data model. Similarly, we sometimes have supra-ordinal measures of some properties of theories that are meaningfully comparable—the number of free parameters in one data model can be meaningfully compared with the number of free parameters in another data model, for instance. By the standard of complete descriptive accuracy, then, both **O** and **NC** fail for at least some of the theoretical virtues. But in many cases, arguably most cases of interest in science, we only have ordinal and non-comparable measures of the support that a theoretical virtue provides to a theory. By the standard of partial descriptive accuracy, then, **O** and **NC** pass. Moreover, some theoretical virtues *always* satisfy **O** and **NC**, and thus satisfy the standard of complete descriptive accuracy. In the following two sections I argue that **O** and **NC** are ubiquitous, though not universal, features of the support relation between theoretical virtues and theories.

4. Ordinality

O stipulates that information about the extent to which a theory is supported by the various theoretical virtues is limited to measurement on an ordinal scale. As Okasha notes, for some of the theoretical virtues an ordinal scale is the best one can hope for; fruitfulness, for example, might be measurable only on an ordinal scale. One might say that Lavoisier's oxygen theory was more fruitful than Priestley's phlogiston theory, but it would be meaningless to say that the difference between the fruitfulness of Lavoisier's oxygen theory and the fruitfulness of Priestley's phlogiston theory was greater than the difference between the fruitfulness of Aristotle's theory of motion and the fruitfulness of Oresme's theory of motion.

For other theoretical virtues, however, Okasha argues that the appropriate scale might be cardinal, ratio, or absolute, and so **O** would not be satisfied. This is clearly correct. When modelling data, for instance, one way to assess the *simplicity* of a data model is simply to count the number of free parameters in the model. Okasha claims that, at least in the case of model selection, 'we have much more than ordinal information' (p. 103). This is true as far as it goes. But it does not go very far.

Theoretical virtues are generally assumed to be broadly applicable criteria of theory choice. But counting free parameters of data models is applicable to only a narrow domain, namely those situations in which one is modelling data. In the 1940s biologists debated the chemical structure of genes: one theory was that genes were composed of proteins, and another was that genes were composed of DNA. There are obviously no parameters to count in either theory. In the 1440s scholastics debated various theories of motion, of which, again, there was no metric of 'simplicity'. Solar motion could be modelled as either an eccentric structure or an epicyclic structure, and again, to judge the relative simplicity of the competing theories one could not just count parameters. These considerations are not meant to deny the importance of simplicity as a theoretical virtue. Rather, they are meant to deny the relevance of a particular way of quantifying simplicity in one narrow domain to the question of what kind of scale simplicity can be generally and meaningfully measured on.[5]

This point can be made starker by considering which scale should be used when comparing theories from very different domains. Suppose one wants to know which theory is simpler, a data model with two free parameters or Bohr's theory of the atom. Although the simplicity of the former can be quantified, the simplicity of the latter cannot. This does not mean that an ordinal comparison of their respective simplicity cannot be made (even if it turns out that the ordinal comparison is a weak ordering or an equivalence).

One might object: supra-ordinal information is at least sometimes available for some theoretical virtues, and maintaining a commitment to **O** amounts to barring such information when it is available. It may be true that some theoretical virtues, such as simplicity, are sometimes measurable on a super-ordinal scale. But for other theoretical virtues the support that they provide to a theory is only measurable on an ordinal scale. So we are left with a situation in which the appropriate scales for measuring the relevant support relations are mixed: some ordinal, some supra-ordinal. This raises the importance of the non-comparability condition (discussed below). How can an ordinal measure of support from one theoretical virtue be compared to a cardinal measure of support from another theoretical virtue? To make such a comparison the theoretical virtues must be *commensurable*— literally, they must share the same scale. At least one non-arbitrary way to render measurements on different scale types commensurable is to infer an ordinal measure of support from the cardinal measure of support for the latter theoretical virtue, and then compare the two ordinal measures

[5] Same with preferences. For a restricted domain of preferences, a ratio scale might be employed. For example, preferences for some middle-sized consumer products might be measurable based on the price individuals are willing to pay for such products. But this is only narrowly applicable; my preference for my pet puppy over the neighbour's mutt cannot be measured in this way. So although in *particular* domains preferences can be measured on a scale more informative than an ordinal scale, the *general* restriction of preference measures to an ordinal scale is reasonable, generally.

of support (because, obviously, one cannot infer a cardinal measure from an ordinal measure).[6]

Similar considerations apply to accuracy, Okasha's other example of a theoretical virtue which is sometimes quantified on a ratio scale. But I will not labour this point, since although I have given reasons to think that **O** ought generally be satisfied, one of Sen's conclusions was that even if **O** is relaxed, one can still derive an Arrow-type impossibility theorem, as long as **NC** is satisfied.

5. Non-Comparability

NC states that, given a certain measurement scale, any transformation of the support measure of a theoretical virtue is admissible (within the class of admissible transformations delimited by the scale type) independent of the transformation of a support measure from another theoretical virtue, and the output of a theory choice rule can vary with respect to such transformations. **NC** is a reasonable constraint on theory choice.

Consider, for example, accuracy. As Okasha notes, accuracy of fit between data and hypothesis is often measured by 'sum-of-squares'. Though convenient, the choice of sum-of-squares is entirely conventional. One could just as easily measure accuracy by the sum of the fourth power of the differences between the observed data values and their hypothetically expected values. Even if accuracy were measured on a ratio scale based on sum-of-squares (thereby violating **O**), there would be a large degree of freedom in which transformations to the measure are admissible (information-preserving). This would be *entirely independent* of the admissibility of transformations of measures of the support that other theoretical virtues provide to the theory (say, the measurement of simplicity by counting free parameters). There is no reason to think that the output of a theory choice rule must remain invariant given such transformations. In other words, even for scenarios in which one can quantify some support relations between theoretical virtues and theories on supra-ordinal scales, such measures are not meaningfully comparable: **NC** holds.

Here is another way to emphasize the reasonableness of **NC**. The condition **NC** stipulates that the support that one theoretical virtue

[6] This raises an outstanding unresolved technical issue. As far as I know, there is no work in social choice that explores scenarios in which preferences are measured on multiple types of scales.

provides to a theory cannot be compared to the support that another theoretical virtue provides to a theory. This is generally true. Consider scope and fruitfulness, and their relative support for any two theories (T_1 and T_2). It makes no sense to say that the ratio of support that simplicity provides to T_1 compared with T_2 is twice that of the ratio of support that fruitfulness provides to T_1 compared with T_2. The respective support relations are simply not quantifiable in a way which admits of such a comparison. That is, the support relations between this theoretical virtue and theory choice are such that **NC** is satisfied.

Okasha claims that this is 'overly pessimistic' (2011 p. 104). But his two arguments for a converse optimism are thin. First, he claims that if all the theoretical virtues were measurable on an absolute scale, then all the respective support relations would be comparable, and so **NC** would be violated. While this conditional is true, its antecedent is absolutely false. It may be the case that the support that *some* theoretical virtues provide to *some* theories, in *some* limited domains, can be measured on an absolute scale (like counting parameters of a data model). But as already argued, the support that *most* theoretical virtues provide to *most* theories, in *most* domains, cannot be measured on an absolute scale. So the first argument gives no reason to think that **NC** is avoidable. His second argument is similar. Okasha relies on a technical result from Sen to show that if all the theory support relations are measurable on their own ratio scale, then **NC** can be relaxed, and if, in addition, the values of the support relations are all non-negative, then the Arrow analogue for theory choice can be avoided. But again, we have already seen that it is not the case that all the theory support relations are measurable on their own ratio scale (fruitfulness, again, is an example of a theoretical virtue that Okasha himself claims is at best measurable on an ordinal scale). So the second argument gives no reason to think that **NC** is avoidable. Thus, contra Okasha, these considerations do not provide good reason to relax **NC**, and so do not provide an escape route from the Arrow analogue for theory choice.[7]

[7] Note also that the occasional violation of **NC** by some theoretical virtues in some domains could only serve as a general escape route from the Arrow analogue for theory choice if one was willing to relax **U**.

6. No escape

Okasha illustrates the information enhancement strategy for avoiding his impossibility theorem with a discussion of Bayesianism and statistical model selection. These examples, Okasha claims, support the contention that the Arrow analogue can be avoided by violating **I**. Both Bayesianism and statistical model selection involve the assessment of theories with algorithms that assign supra-ordinal and comparable measures of theory support (this reasoning goes), and so both **O** and **NC** are not satisfied, and thus they provide an 'escape route' from the Arrow analogue for theory choice. Unfortunately, neither example is compelling: both involve strategies that amount to denying the relevance of the Kuhnian theoretical virtues that motivated the analogy between social choice and theory choice in the first place.

6.1 Bayesianism

Okasha notes that Bayesians employ two 'criteria' (his term) for choosing between rival theories T_i: the prior probability, $P(T_i)$, and the likelihood, $P(E \mid T_i)$. The theory with the highest product of $P(T_i)$ and $P(E \mid T_i)$ is the most likely to be true, or the most deserving of our belief. Since probabilities are measured on an absolute scale, a Bayesian theory choice algorithm does not satisfy **O** or **NC**, and thus the Arrow analogue is avoided.

What happened to the Kuhnian theoretical virtues (scope, fruitfulness, etc.) that Okasha began with? They do not appear in his discussion of Bayesianism. There are two ways I can see the Kuhnian theoretical virtues relating to the two Bayesian 'criteria', $P(T_i)$ and $P(E \mid T_i)$. First, perhaps the Kuhnian theoretical virtues help *determine* the values of $P(T_i)$ and $P(E \mid T_i)$. But such a determination would require some sort of algorithm to translate the application of the virtues into probabilities. Based on Okasha's own analogy, such an algorithm would face an Arrow analogue, since Arrow's axioms — unanimity, non-dictatorship, etc. — would obviously apply to such an algorithm. Second, perhaps the Kuhnian theoretical virtues are simply meant to be *replaced* by the Bayesian 'criteria'. This, however, is unsatisfactory. Okasha himself laboured to defend the importance of the theoretical virtues, as many others have before him. Worse though, is that the Bayesian criteria are (merely) *post hoc* measures of the goodness of theories, whereas most of the theoretical virtues commonly discussed are substantive properties that *constitute* the goodness of theories. For example, the prior probability of a theory is often said to be constituted, at least in part, by its simplicity: simpler theories are usually thought to be more probable, *ceteris paribus*, than complex theories.[8]

6.2 Statistical model selection

The aim of statistical model selection is to choose the best of multiple hypotheses based on two criteria: simplicity and 'goodness of fit' (how close a hypothesis fits with the data). Okasha describes one popular model selection algorithm known as the Akaike information criterion (AIC). An AIC score is determined by two properties of hypotheses: the number of free parameters of a hypothesis (measuring simplicity), and the probability of the data assuming the hypothesis is true (measuring goodness of fit).[9] Okasha suggests that these two properties can be thought of as 'utility functions' which assign numbers to competitor hypotheses that represent their respective simplicity and goodness of fit. Since these numbers have meaningful content, and that content is only preserved under the identity transformation, the numbers are measures on an absolute scale. Thus **O** and **NC** are not satisfied and so the Arrow analogue can be avoided.

Statistical model selection only applies to scenarios in which one is comparing hypotheses of the relationship between two variables based on noisy data. Its domain of application in science is thus narrow. Its domain of application is further narrowed given its very particular metric of simplicity. As argued above, there are many parts of science in which simplicity cannot be measured by the number of free parameters of a data model, and so the above argument applies with no loss to algorithms for statistical model selection, such as AIC, that employ a metric of simplicity. Finally, one might ask the question raised above for the Bayesian 'criteria': what happened to the (other) Kuhnian theoretical virtues (scope, fruitfulness, etc.) that Okasha began with? AIC scores quantify two theoretical virtues (simplicity and accuracy), but they say nothing about the other theoretical

[8] One might press the replacement strategy by holding that we should forget about theoretical virtues, and instead directly assess the probabilities required of our measures of confirmation by appealing only to our sum total of evidence. How exactly one determines these probabilities is a harder problem than many seem to suppose. But worse, Stegenga 2013 argues that an impossibility theorem analogous to Arrow's theorem also applies to amalgamating our sum total of *evidence*, regardless of theoretical virtues.

[9] Formally, $AIC = 2k - 2\ln(L)$, where k is the number of parameters of a model and L is the likelihood. The lower an AIC score, the better.

virtues. As noted, Okasha himself admits that fruitfulness, say, might only be measurable on an ordinal scale. In short, statistical model selection is not a good example of measuring theory support by theoretical virtues on a supra-ordinal scale.

7. Voter fraud

The analogy between social choice and theory choice holds that theoretical virtues are 'voters'. There are two problems with identifying these voters. First, there is no consensus in the philosophical literature regarding which properties of theories matter to theory choice. Kuhn (1962) had earlier argued that puzzle-solving ability, quantitative precision, novelty of predictions, and plausibility were also important theoretical virtues, in addition to those mentioned by Okasha. Providing lists of theoretical virtues has been a popular pastime for philosophers. Hempel's were simplicity, support by more general theories, prediction of novel phenomena, and plausibility (1966). Van Fraassen 1980 includes elegance, simplicity, completeness, unifying power, and explanatory power. Longino 1994 provides a list of 'feminist theoretical virtues', which include ontological heterogeneity, mutuality of interaction, applicability to human needs, accessibility of ideas, and novelty.

Second, as Okasha notes, a coarse-grained theoretical virtue might in fact be better thought of as a *type* of virtue of which there are multiple sub-types. 'Simplicity', for instance, might include as subtypes 'number of parameters' and 'mathematical tractability'. This manoeuvre allows one to avoid the charge that some theoretical virtues are vague — for instance, it might not be entirely clear which of two theories is 'simpler', but it *is* clear which if any has fewer free parameters. A result of this manoeuvre is a multiplication of sub-types of theoretical virtues. How many sub-types are there? I have no idea, but if types get split at every charge of vagueness, one might expect that the number of virtues is large, perhaps uncountably large, or perhaps just *vaguely* large (exercise: try explicating 'mathematical tractability'). In social choice there is a principled way to determine who the voters are, but in theory choice there is no such principled determination. This is obviously a disanalogy with social choice.[10]

[10] Ultimately this may be a difference that makes no difference, since on any reasonable determination of who the 'voters' are (Kuhn's, van Fraassen's, Longino's, etc. …), and on any degree of specification of the sub-types of theoretical virtues, an Arrow analogue will hold.

8. Conclusion

The analogy between theory choice presented in Okasha 2011 is deep and insightful.[11] However, contrary to Okasha's optimism about finding an 'escape route' by relaxing the ordinality and non-comparability sub-conditions of the independence axiom, I have argued that these sub-conditions are generally applicable in science. The impossibility theorem for theory choice remains puzzling.[12]

References

Arrow, Kenneth 1951: *Social Choice and Individual Values.* New York: John Wiley and Sons.
Gaertner, Wulf 2006: *A Primer in Social Choice Theory.* Oxford: Oxford University Press.
Geanakoplos, John 2005: 'Three Brief Proofs of Arrow's Impossibility Theorem'. *Economic Theory*, 26, pp. 211–15.
Hempel, Carl 1966: *Philosophy of Natural Science.* Englewood Cliffs, NJ: Prentice-Hall.
Kuhn, Thomas 1977: *The Essential Tension.* Chicago, IL: University of Chicago Press.
——1962: *The Structure of Scientific Revolutions.* Chicago, IL: University of Chicago Press.
Longino, Helen 1994: 'In Search of Feminist Epistemology'. *The Monist*, 77, pp. 472–85.
Okasha, Samir 2011: 'Theory Choice and Social Choice: Kuhn versus Arrow'. *Mind*, 477, pp. 83–115.
Schmid, Hans Bernhard, Daniel Sirtes, and Marcel Weber 2011: *Collective Epistemology.* Frankfurt: Ontos Verlag.
Sen, Amartya 1970: *Collective Choice and Social Welfare.* San Francisco: Holden-Day.

[11] One might object to the very idea of modelling theory choice with technical apparatus imported from social choice by noting that scientists rarely engage in large-scale theory choice in the way that Popper, Kuhn, and others seemed to suppose. If so, the Arrow analogue hardly ever applies. While I agree that few scientists engage in explicit large-scale theory choice, nothing about Okasha's framework requires this. The analogy between theory choice and social choice holds even if most 'choices' are implicit, or if 'chosen theory' is just shorthand for 'most confirmed theory' or 'most believed theory', and the analogy holds for 'theories' with narrow scope in a local domain. For pushing me on this I thank Ian Hacking.

[12] I am grateful to Samir Okasha and two anonymous reviewers for detailed commentary, and to an audience at the University of Toronto for fruitful discussion. Financial support was provided by the Banting Postdoctoral Fellowships Program administered by the Government of Canada.

——1977: 'On Weights and Measures: Informational Constrains in Social Welfare Analysis'. *Econometrica*, 45, pp. 1539–72.

Stegenga, Jacob 2013: 'An Impossibility Theorem for Amalgamating Evidence'. *Synthese*, 190, pp. 2391–411.

van Fraassen, Bas 1980: *The Scientific Image*. Oxford: Oxford University Press.

Weber, Marcel 2011: 'Experimentation versus Theory Choice: A Social-Epistemological Approach'. In Schmid, Sirtes, and Weber 2011.

