

# Objectivity, Ambiguity, and Theory Choice\*

Alexandru Marcoci<sup>†</sup>      James Nguyen<sup>‡</sup>

Forthcoming in *Erkenntnis*

## Abstract

Kuhn argued that scientific theory choice is, in some sense, a rational matter, but one that is not fully determined by shared objective scientific virtues like accuracy, simplicity, and scope. Okasha imports Arrow's impossibility theorem into the context of theory choice to show that rather than not fully determining theory choice, these virtues cannot determine it at all. If Okasha is right, then there is no function (satisfying certain desirable conditions) from 'preference' rankings supplied by scientific virtues over competing theories (or models, or hypotheses) to a single all-things-considered ranking. This threatens the rationality of science. In this paper we show that if Kuhn's claims about the role that subjective elements play in theory choice are taken seriously, then the threat dissolves.

Keywords: *Theory choice; scientific rationality; ambiguity; Kuhn; Okasha*

## 1 Introduction

Suppose a scientist is faced with a collection of competing scientific theories, models or hypotheses. Moreover, suppose that she cares about a number of distinct scientific virtues – accuracy, simplicity, and scope for example. How is the scientist to rationally choose the 'best' alternative, all-things-considered? One would like to think that, whatever the details of how the choice is made, some rational procedure is followed to make it.

According to Kuhn (1972), scientists faced with such a choice, even if they agreed about which theoretical virtues should guide their choice, could still rationally disagree about which is the 'all-things-considered' best competing theory. In this sense, there is no 'unique algorithm' that takes how well theories fare according

---

\*Authors are listed in alphabetical order, the paper is fully collaborative. We are grateful to two anonymous referees for their useful comments. Nicholas Baigent, Wulf Gaertner, Michael Morreau, Samir Okasha, and Nicholas Wüthrich also provided valuable feedback. Thanks also to audiences at the first London Philosophy of Science graduate workshop at UCL, the 17th Pitt/CMU graduate conference, MuST15 at Tilburg, and EPSA15 at Düsseldorf.

<sup>†</sup>a.marcoci@lse.ac.uk, Department of Government, LSE.

<sup>‡</sup>jnguyen4@nd.edu, History and Philosophy of Science Program, University of Notre Dame and Centre for Philosophy of Natural and Social Science, LSE.

to the scientific virtues and delivers a ‘winner’. This is not to say that theory choice is ‘a matter of mob psychology’ (Lakatos, 1970, p. 178), but rather that the shared and objective virtues do not determine by themselves a winner, or unique ranking of the theories.

Okasha (2011) uses the formal framework of social choice theory to argue that, rather than there being no unique algorithm for using the objective information supplied by the scientific virtues to rationally choose the best theory, there is *no* such algorithm whatsoever: ‘Where Kuhn saw an embarrassment of riches, Arrow tells us that there is nothing at all’ (Okasha, 2011, p. 93). If theory choice were purely an objective matter, then this would be a significant problem. But Okasha fails to pay attention to Kuhn’s claims about the subjective elements involved in theory choice; in particular, the role scientists play in disambiguating the scientific virtues. We demonstrate how to do this in the framework Okasha proposes, and show via simulation that this blunts the threat posed significantly.

We proceed as follows. In Section 2 we outline Okasha’s argument. Section 3 discusses subjective elements involved in the context of theory choice, i.e. the ambiguity of scientific virtues. In Section 4 we argue that Okasha’s proposed way of dealing with this in the social choice framework is unsatisfactory. In Section 5 we propose an alternative treatment which we then feed into the definition of rationality, in Section 6. Finally, in Section 7 we demonstrate that with this in place, the impossibility result can be bypassed. Perhaps counter-intuitively, the subjectivity involved in disambiguating scientific virtues turns out to be a good thing.

## 2 Okasha’s challenge

Okasha employs a simple but persuasive argumentative strategy. Let  $\mathfrak{V}$  be a finite set of  $m$  scientific virtues, and  $\mathfrak{T}$  a finite set  $n$  of competing theories. Each scientific virtue  $i \in \mathfrak{V}$  provides an ordinal ranking of the elements of  $\mathfrak{T}$ , from most to least virtuous according to  $i$ . These rankings are transitive, reflexive, and complete binary relations.<sup>1</sup> When theory  $x$  is preferred to theory  $y$  by virtue  $i$  we write  $y \prec_i x$ . A theory choice situation is a profile of rankings of theories by virtues, where a profile is an ordered tuple  $\langle \prec_1, \dots, \prec_m \rangle$  for virtues 1 through  $m$ . A theory choice function maps profiles to an all-things-considered binary relation  $\preceq$  defined over  $\mathfrak{T}$ . A theory  $x$  is strictly preferred (all-things-considered) to theory  $y$ , i.e.  $y \prec x$  if and only if  $y \preceq x$  and it’s not the case that  $x \preceq y$ .

What requirements should we impose on a rational theory choice function? Okasha (2011, pp. 92-93) argues that the Arrovian conditions on preference aggregation have clear analogues in the context of theory choice. Unrestricted domain (UD) requires that a theory choice function be applicable irrespective of how theories are ranked by virtues. Weak Pareto (WP) requires that, for all  $x, y \in \mathfrak{T}$ , if  $x$  is ranked above  $y$  according to every virtue, then  $x$  should be ranked above  $y$  all-things-considered. Independence of Irrelevant Alternatives (IIA) requires that

---

<sup>1</sup>In this paper we restrict our focus to *strict* rankings associated with scientific virtues for simplicity. Where relevant our approach can be extended to accommodate rankings allowing for ties. This is of no conceptual importance.

the all-things-considered relation between two theories takes into account only how those two theories are ranked by the scientific virtues, i.e. the overall comparison between the two be insensitive to how virtues rank them with respect to a third theory. Non-Dictatorship (ND) demands that there is no virtue  $i$  such that for every pair of alternatives  $x, y \in \mathfrak{A}$ , whenever  $i$  prefers  $x$  to  $y$ ,  $x$  is ranked above  $y$  all-things-considered. Finally, Overall Rationality (OR) demands that a theory choice function deliver a transitive, complete ranking for every element in its domain.<sup>2</sup> With theories replacing social alternatives, and scientific virtues replacing voters, it is immediate to see that Arrow’s (1951) impossibility result applies. In other words, there is no theory choice function that satisfies UD, WP, IIA, ND and OR. Okasha’s challenge is, that assuming rational theory choice requires the existence of such a function, rational theory choice is impossible.

‘If we agree that U, P, N, and I are conditions on reasonable theory choice, then it is obvious that an Arrovian impossibility result applies. So long as there are at least three alternative theories, there exists no theory choice rule that satisfies all four conditions. This spells bad news for the possibility of making ‘rational’ theory choices’ (Okasha, 2011, p. 93).<sup>3</sup>

Okasha’s argument has generated much discussion. Morreau (2014, 2015) suggest restricting UD. Rizza (2013) and Stegenga (2015) follow up on Okasha’s (2011) own suggestion of enriching the informational basis of scientific virtues by providing a common cardinal scale allowing for inter-virtue comparisons (thereby dropping IIA). Whether all scientific virtues provide such information is questionable as discussed in Okasha’s (2015) response to Stegenga. Relatedly, Gaertner and Wüthrich (2016) suggest *imposing* a cardinality via a scoring rule in a way that captures the spirit of IIA in a cardinal framework. Bradley (forthcoming) suggests that rationality only requires *ruling out* certain alternatives, not a transitive and complete ranking of theories.

In contrast to these authors, rather than attempting to reformulate the conditions, our focus is on one of Okasha’s modelling assumptions. Once profiles of competing theories ranked by virtues are provided, for the purposes of this paper we can grant that Okasha’s result kicks in. What we question instead is whether or not the virtues provide such rankings. If theory choice were perfectly objective then this would be a natural assumption to make. But if Kuhn is correct, then rational theory choice involves subjective elements as well. In particular, at least for our current purposes, subjective elements are required to disambiguate between different ways the same virtue would evaluate competing theories. And if virtues are ambiguous in the sense to be outlined in Section 3, then particular theory choice situations do not supply a unique profile with which to augment a theory choice

---

<sup>2</sup>Note that this condition is usually build into the definition of an aggregation function. In this paper we include it as a separate condition in light of the fact that we are interested in aggregation functions that map to intransitive/incomplete binary relations. We will use the term ‘intransitive’ to refer to the intransitivity of the entire binary relation ( $\preceq$ ), note that this is compatible with  $\prec$  being transitive.

<sup>3</sup>Note that his ‘U’ is our ‘UD’; his ‘P’ our ‘WP’; his ‘N’ our ‘ND’; his ‘I’ our ‘IIA’; and where he assumes our OR in the definition of a theory choice rule we pull it out as a further condition.

function. And as we demonstrate in Section 7, if enough virtues are assumed ambiguous to a certain extent, the threat posed by Okasha's argument dissolves.

### 3 Subjectivity and Ambiguity

Kuhn (1972) goes at great lengths to explain the impact of subjective as well as objective factors in his model of theory change. According to Kuhn, the way scientists evaluate the adequacy of scientific theories is guided by the scientific virtues. These form a shared and, in Kuhn's view, objective list of adequacy conditions according to which every scientist evaluates every theory. Therefore when confronted with a list of competing theories, scientists will produce rankings of these theories according to the virtues taken into consideration.

*Prima facie*, each scientific virtue  $i$  supplies a unique preference ranking  $\prec_i$ . So for a given theory choice situation, a scientist who starts the aggregation procedure is faced with one and only one profile from which to generate an all-things-considered ranking. This is a natural assumption to make in the context of orthodox social choice theory, but it fails to capture the appropriate notion of ambiguity of scientific virtues. Kuhn claims that:

'Individually the [virtues] are imprecise: individuals may legitimately differ about their application to concrete cases' (1972, p. 357)

and relatedly:

'Individuals must then still choose and be guided by the [virtues] when they [choose]. For that purpose, however, each must first flesh [them] out ... and each will do so in a somewhat different way' (1972, p. 364).

This is the sense in which subjective elements, according to Kuhn, enter into the way scientists choose among competing theories, models or hypotheses.<sup>4</sup> For example, two scientists may disagree with respect to how to interpret 'simplicity'. Suppose the competitors are hypotheses in the form of mathematical equations, polynomials for instance. One scientist might believe that equation  $x$  is simpler than equation  $y$  if and only if  $x$  contains strictly fewer parameters than  $y$ . Another might use the order (the largest exponent) of the equations as their guide to simplicity. A third might use the computational labour required to generate solutions to the equations). Alternatively, suppose the competing theories consist of qualitative statements that consist of equal numbers of universally quantified conjunctions. One way of comparing them with respect to accuracy would be to

---

<sup>4</sup>There is another natural way of thinking about the subjectivity involved in choosing between competing scientific theories (Kuhn, 1972, p. 358). When trying to arrive at an all-things-considered ranking of the theories presented with, different scientists may assign different weights to how much the virtues 'count' for the final ranking. For instance, radical empiricists will most likely assign a very high importance to accuracy to the detriment of all the other virtues, whereas others might be interested in a mix of accuracy and simplicity. So, although the virtues form a common template according to which theories are being evaluated, different scientists may disagree to how important some of them are. This second sense in which subjectivity appears in the context of theory choice raises interesting problems, but we will not address it in this paper.

simply count the number of strictly true conjuncts. Another would be to compare the absolute number of falsifying instances (by summing the number of falsifiers across the conjunction). These may plausibly deliver different results. Irrespective of the different reasons for taking virtues to be ambiguous, Kuhn believes that:

‘The considerable effectiveness of [scientific virtues] does not . . . depend on their being sufficiently articulated to dictate the choice of each individual who subscribes to them’ (1972, p. 362).

Therefore, although the virtues according to which decisions are being made are objective and shared by every member of a particular scientific community, the way each virtue ranks theories is not a matter of fact. There are three senses in which this could be the case, mapping to three different ways of understanding the “concrete cases” Kuhn talks about. First, the same scientist is free to adopt a different interpretation in different theory choice scenarios (i.e. where choosing among different sets of competing theories, models, or hypotheses). Second, the same scientist is free to adopt a different interpretation in different theory application scenarios (i.e. where the theories in question are fixed, but applied to different target systems). In some of these scenarios she may interpret simplicity in one way (number of parameters say), and in another she may adopt a different interpretation (computational labour required to make predictions say). In other words, there is no threat of irrationality stemming from this kind of inter-context inconsistency with respect to the interpretation of a virtue.

But more significantly from our current perspective, even in a particular theory choice scenario, where the pertinent virtues, competitors, and application are fixed, different scientists may legitimately disagree about how to interpret each scientific virtue in that context. Such different interpretations can presumably lead to different orderings of theories, and therefore different profiles with respect to which to apply an aggregation procedure. In this sense, even in a particular theory choice context, there is no ‘matter of fact’ with respect to how a scientific virtue orders the competing theories. Different rankings can be equally ‘correct’, and thus a scientific virtue can be ambiguous between them. To explain why someone ranks theories in a particular context according to simplicity, is a question for the sociologist and psychologist, thinks Kuhn, and not for the analytical philosopher of science.

So, to sum up. In particular theory choice contexts, scientific virtues can be ambiguous across multiple orderings of the competitors. Different sociological and psychological facts about the scientists involved can legitimately lead them to disambiguate a virtue in a different manner. Ambiguity thus allows for some freedom of movement between different rankings. How much freedom of movement differs from context to context and also depends on the subjective reasons guiding individual scientists towards particular rankings. In this paper we model this space of movement and investigate the impact this extra dimension of theory choice has on the notion of scientific rationality and ultimately on bypassing the impossibility Okasha discusses.

## 4 Okasha’s treatment

Okasha claims that ‘[d]isambiguation can always be carried out by sub-dividing an ambiguous [virtue]’ (2011, p. 85) into non-ambiguous distinct preference rankings, each of which are included in the profile of rankings that a theory choice function takes as an argument. So if, for example, simplicity is ambiguous between two distinct rankings  $\prec_{si}$  and  $\prec_{si'}$ , then both are included in the profile.

This approach is unsatisfactory as it does not capture Kuhn’s claims about legitimate disagreement with respect to how virtues apply to concrete cases. Rather than modelling  $\prec_{si}$  and  $\prec_{si'}$  as competing disambiguations of simplicity, Okasha’s approach treats them as being compatible alternatives. So multiple scientists who disagreed with respect to how to disambiguate simplicity in a concrete instance of theory choice would be treated as agreeing that each disambiguation should be used for the purpose of generating an all-things-considered ranking. Okasha’s approach thus multiplies the objective element of theory choice – an additional virtue for each disambiguation of a virtue – rather than recognising Kuhn’s claim that there is a subjective element involved in turning these less than fully articulated notions into ones that can guide choice.

We therefore believe that an alternative account of ambiguity in theory choice is worth pursuing. We propose such an account below and argue that it formally captures the idea that virtues are ambiguous in Kuhn’s sense, thus capturing the subjective element involved in theory choice, and that it blunts the threat posed by Okasha’s argument.

## 5 A Kuhnian construal of ambiguity

We have argued that on Kuhn’s construal of theory choice scientific virtues cannot provide an objective (or in any sense ‘true’) ordering of theories. Therefore, given their ambiguity scientists can have some freedom of movement in between different orderings of the same set of theories under the same virtue. How different can two rankings be so that a scientist treats them as being different disambiguations of the same virtue? This is sensitive to how many different ways there are to rank the alternatives, which is sensitive to the size of the choice set under consideration. For  $n$  alternatives, there are  $n!$  distinct (strict) rankings over them. So, with respect to our example of choosing between  $\mathfrak{T} = \{x, y, z\}$ , there are 6 possible rankings over  $\mathfrak{T}$ . A virtue that specified a precise ranking would be unambiguous. A virtue which was maximally ambiguous between all of these 6 would be uninformative. Between these two extremes, there are multiple ways of setting a sensible threshold.

One obvious way to proceed is to define a notion of ‘closeness’ between two rankings in terms of the number of their pairwise disagreements. Two rankings are ‘close’, for instance, if and only if they differ on only one pair of alternatives. To understand this notion of ‘closeness’ better assume  $\mathfrak{T} = \{x, y, z, u, v\}$  and consider the four rankings which are close to the ranking  $v \prec_i u \prec_i z \prec_i y \prec_i x$ .

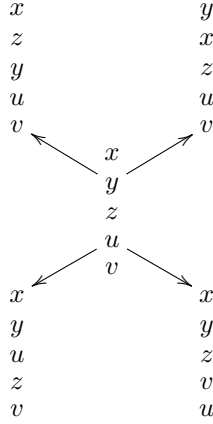


Figure 1

This notion can be extended in a natural way to profiles. To illustrate this consider the following example. Suppose a scientific community is faced with the choice out of  $\mathfrak{T} = \{x, y, z, u, v\}$ , by means of 3 virtues,  $i, j, k$  and that  $x \prec_j v \prec_j u \prec_j z \prec_j y$  and  $y \prec_k x \prec_k v \prec_k u \prec_k z$ . In the below set all profiles are close to the profile in which  $v \prec_i u \prec_i z \prec_i y \prec_i x$ .

$$\left\{ \begin{array}{l} \langle x \ y \ z \rangle \\ \langle y \ z \ u \rangle \\ \langle z \ , \ u \ , \ v \rangle \\ \langle u \ v \ x \rangle \\ \langle v \ x \ y \rangle \end{array} , \begin{array}{l} \langle x \ y \ z \rangle \\ \langle z \ z \ u \rangle \\ \langle y \ , \ u \ , \ v \rangle \\ \langle u \ v \ x \rangle \\ \langle v \ x \ y \rangle \end{array} , \begin{array}{l} \langle y \ y \ z \rangle \\ \langle x \ z \ u \rangle \\ \langle z \ , \ u \ , \ v \rangle \\ \langle u \ v \ x \rangle \\ \langle v \ x \ y \rangle \end{array} , \begin{array}{l} \langle x \ y \ z \rangle \\ \langle y \ z \ u \rangle \\ \langle u \ , \ u \ , \ v \rangle \\ \langle z \ v \ x \rangle \\ \langle v \ x \ y \rangle \end{array} , \begin{array}{l} \langle x \ y \ z \rangle \\ \langle y \ z \ u \rangle \\ \langle z \ , \ u \ , \ v \rangle \\ \langle v \ v \ x \rangle \\ \langle u \ x \ y \rangle \end{array} \right\}$$

Figure 2

There are two different ways in which the definition of ‘closeness’ can be relaxed. Firstly, we have so far assumed two rankings are ‘close’ if they differ in at most the ordering of a single pair of theories. This makes more sense if the set of alternatives is small than it does for larger sets. Consequently, in general, we may consider two rankings as being close even if they differ in the ordering of  $\beta$  pairs of alternatives. In the example presented in Fig. 1  $\beta = 1$ , which meant that for any given ranking there are four others which are close to it. But if  $\beta = 2$  then this grows to include 9 more additional rankings. And if  $\beta = 10$  then this corresponds to a trivial notion of closeness in the sense that any way of ranking the alternatives is close to any other. In general, for a set of  $n$  alternatives, the maximum value  $\beta$  can take is  $\frac{n(n-1)}{2}$ . As before, this generalized relation of closeness can be naturally extended to deliver a notion of closeness for profiles.

Secondly, there is no reason why closeness between profiles be judged in reference to only one virtue. That is, for a given profile there may be profiles which are close to it in the sense in which they differ from the original with respect to pairwise disagreements on multiple virtues (or one which disagrees on virtue  $i$  and another that disagrees on virtue  $j$ ). We will use  $\alpha$  to denote the number of virtues assumed in the definition of closeness. In the example in Fig.2,  $\alpha = 1$ , and this delivered four profiles. If  $\alpha = 2$  then there would be 20 additional profiles.

We can combine the two generalisations and obtain the notion of  $\alpha$ - $\beta$ -closeness. As  $\alpha$  and  $\beta$  grow the size of the set of profiles close to a particular profile grows as well.<sup>5</sup> This notion of closeness can help us model Kuhn’s idea of ambiguity.

If virtues really are ambiguous, as Kuhn suggests they are, then can the profiles which majority voting<sup>6</sup> maps to an intransitive relation be avoided by replacing them with other profiles  $\alpha$ - $\beta$ -close to them? Such a move would be permissible since, as per Kuhn, the ranking of competing theories based on scientific virtues is not set in stone, and each virtue may be disambiguated in different ways, where different disambiguations may result for reasons specific to each individual scientist. The point then, is that although a scientist may have well justified (sociological or psychological) reasons to disambiguate each virtue the way that she does, she cannot hold each disambiguation at the same time, on pain of her favoured aggregation function mapping to an intransitive/incomplete result. But, if she cannot hold each disambiguation at the same time, then this provides evidence that something has ‘gone wrong’ in the way that she has disambiguated the virtues, at least when considered together. And thus, there is nothing stopping her from revisiting them and adopting alternative close disambiguations for (at least some) of the virtues.

It bears noting here that we are not claiming that our notion of closeness, in terms of pairwise disagreements with the original disambiguations, is the only (or even the best) way of thinking about ambiguity in the formal framework under consideration. This (similar to our choice of focusing on pairwise majority voting) is a modelling assumption, and it would be an interesting question to consider different ways of capturing how a scientific virtue could be ambiguous over multiple rankings. However, it does seem plausible that if a scientist were to be forced to revisit how she disambiguated a certain scientific virtue in a given theory choice situation, then adopting an alternative disambiguation that radically disagreed with her original one would be undesirable. It could even be taken as a sign that her initial attempt at interpreting the virtue was not sufficiently well grounded. Alternatively, the sociological or psychological reasons that led to the original disambiguation might be seen as having an ‘anchoring’ effect. The claim is that when faced with an intransitive/incomplete all-things-considered value, alternative disambiguations that

---

<sup>5</sup>Notice that if we increase  $\alpha$  and  $\beta$  too much, i.e to  $\alpha = m$  and  $\beta = \frac{n(n-1)}{2}$ , then this set collapses into the domain that profile belongs to.

<sup>6</sup>Pairwise majority voting can be defined thus: for a set of virtues  $\mathfrak{V}$ , let  $\Delta^+ =_{df} \{i \in \mathfrak{V} : y \prec_i x\}$ ,  $\Delta^- =_{df} \{i \in \mathfrak{V} : x \prec_i y\}$ . Then:  $y \prec x$  if and only if  $|\Delta^+| \geq |\Delta^-|$ . For the remainder of this paper we will focus on pairwise majority voting and assume there is an odd number of virtues. This means that the output of the aggregation, under pairwise majority, will always be a strict ordering (if an ordering, at all). We use pairwise majority vote for its simplicity in illustrating our argument. We do not thereby suggest that it be the actual aggregation function used; clearly, for example, one may want to weight different virtues differently.



largely agree (in terms of pairwise disagreements) with the original disambiguations are to be preferred.

To illustrate the potential of such an approach to bypass Arrow's impossibility result, consider the following case: assume three virtues are used to rank three alternatives. It is trivial that majority voting maps some profile 3-3-close to any Condorcet-like profile<sup>7</sup> into a transitive ranking, since any element of the domain under consideration is 3-3-close to any other (so think of the profile in which all virtues agree). This is not very useful as these particular large values of  $\alpha$  and  $\beta$  do not appear plausible for such a small number of theories and virtues.

Nevertheless as the number of competing theories and virtues grow, higher values of  $\alpha$  and  $\beta$  become intuitively plausible, since for more competitors and virtues, fixed values of  $\alpha$  and  $\beta$  span proportionally less of the corresponding universal domain. A virtue faced with a large number of theories may be such that multiple pairwise disagreements between competing disambiguations are allowed. And if there are a larger number of virtues we may allow for more of them to be ambiguous.<sup>8</sup> Notice we are not suggesting replacing the troublesome profiles (those that instantiate a Condorcet-pattern of preferences) with simply any other profile in the domain. We want to restrain the possible replacements as much as possible, and we do so by only looking at profiles which are  $\alpha$ - $\beta$ -close to the troublesome ones, for plausible values of  $\alpha$  and  $\beta$ .

But what precisely are plausible values? This is a difficult question to answer in the abstract for two reasons. First, as noted above, it does seem reasonable to assume that the plausibility of certain values of  $\alpha$  and  $\beta$  depends on the number of alternatives ( $n$ ) and virtues ( $m$ ) involved in the theory choice situation. Secondly, and more importantly, even fixing the numbers  $n$  and  $m$ , whether or not particular values of  $\alpha$  and  $\beta$  are plausible will still presumably depend on the particular field of research in question, and the individual scientist doing the aggregation. A scientist with deeply entrenched reasons for disambiguating in a way that led her favoured aggregation function delivering an intransitive/incomplete value may well resist considering alternative disambiguations which disagreed significantly with her original disambiguation (thereby restricting the value of  $\beta$ ). And she may well resist the idea that she has to revisit a large number of the virtues she has been using to guide her choice (thereby restricting the value of  $\alpha$ ). As such, a detailed discussion of which  $\alpha$  and  $\beta$  are plausible would require a detailed discussion of particular theory choice scenarios, which we cannot do in this paper. Instead, we aim to offer a proof of concept highlighting the viability of such an approach to the conceptualisation of scientific rationality and show that it blunts the threat raised by Okasha. In the following section we explain how  $\alpha$ - $\beta$ -closeness can lead to a weakened notion of rationality in the context of theory choice. And then we present some results obtained by applying this relaxed notion of rationality to several simple theory choice scenarios.

---

<sup>7</sup>A Condorcet profile is simply a profile that pairwise majority voting maps to an intransitive value.

<sup>8</sup>We conjecture, that for each domain, there are minimal values of  $\alpha$  and  $\beta$  such that majority voting will succeed somewhere in each neighbourhood of profiles, and that the values of  $\alpha$  and  $\beta$  are low enough to make the rationality of theory choice nontrivial. Proving this conjecture is beyond the scope of this paper, but is a viable avenue for future research.

## 6 Weak rationality

Suppose that scientific virtues, such as simplicity, accuracy, scope, etc. do deliver complete rankings of competing theories. Then, for any scientist, a theory choice situation can be represented by a profile collecting all the ranking of theories according to the individual virtues. Okasha construes the rationality of theory choice in the following way:

**Rationality** Theory choice is rational only if there exists a function that respects UD, WP, ND, IIA and that takes every profile in UD to a transitive and complete all-things-considered ordering.

If one were to accept Okasha’s principles governing theory choice discussed in Section 2, then Arrow’s result indeed shows that theory choice is irrational. In other words, there is no such function that outputs a transitive and complete all-things-considered ordering no matter what rankings the virtues supply. Using the notion of  $\alpha$ - $\beta$ -closeness introduced in the previous section, we aim to offer a double-tier weakening of Rationality.

**Weak Rationality** Theory choice is rational only if there exists a function that respects U, WP, ND, IIA and that for at least  $\gamma$  profiles, takes those profiles, or profiles  $\alpha$ - $\beta$ -close to them, into transitive and complete, all-things-considered orderings.

*Rationality* and *Weak Rationality* differ in two respects. Firstly, whereas *Rationality* stipulates that the aggregation function ought to deliver a transitive and complete all-things-considered ranking for any profile, *Weak Rationality* is concerned only with  $\gamma$  profiles. For simplicity, we will express  $\gamma$  as a percentage of profiles in the domain for which *Weak Rationality* holds. Secondly, *Weak Rationality* not only checks the behaviour of a function applied to a profile, but also its behaviour applied to all profiles  $\alpha$ - $\beta$ -close to it. So, in case we try to apply pairwise majority voting, say, to a Condorcet profile, *Weak Rationality* will be satisfied if pairwise majority voting can deliver, for at least one profile  $\alpha$ - $\beta$ -close to the Condorcet profile, a transitive, all-things-considered ranking. In contrast, Condorcet profiles represent the counterexamples to satisfying *Rationality* with pairwise majority voting.

The first weakening is motivated by the following idea. Arrow’s theorem tells us that theory choice is not rational according to *Rationality* because for any function there will be at least one profile that function will not map to a transitive and complete all-things-considered ordering. But what if there really is a single such profile for a particular function? Let  $f$  denote your favourite aggregation function. There is a sense in which  $f$  would be *less* ‘rational’ (in an intuitive sense) if it generated a transitive and complete all-things-considered ordering from only 1 out of 216 profiles (the space of all profiles formed out of three virtues ranking over three theories), than it would be if it did so from 215. If a scientist used the function in the former case she would be acting ‘irrationally.’ But she wouldn’t if she did so in the latter. In fact, not using  $f$  in such a scenario simply because

of the threat posed by the single profile  $f$  fails to map to an ordering, would be ‘irrationally’ cautious. Suppose a bookie offered you a choice between two bets. the first bet returns £100 with probability 1, the second £200 with probability .5 and £0 otherwise. Preferring the first bet in this instance seems rational. However, as the probability of winning £200 increases, the second bet becomes more appealing. There seems to be a point at which preferring the first to the second bet becomes irrationally cautious. Analogously, a scientist refraining from using  $f$  because of its failure in only 1 out of 216 cases appears irrationally cautious. As such, rationality can be treated as a matter of degree.

The second weakening is inspired by Kuhn’s idea that the way in which virtues rank theories is subjective. We understand this as saying that in case a particular profile delivers, under an aggregation function, an intransitive (or incomplete) all-things-considered relation, moving to an  $\alpha$ - $\beta$ -close profile is permissible. We have discussed Kuhn’s idea of ambiguity in Section 3 and the meaning of  $\alpha$ - $\beta$ -closeness in Section 5. The issue of what values  $\alpha$  and  $\beta$  should take is still beyond the scope of this paper, but in the next section we present a proof of the viability of this proposal for saving the rationality of theory choice, if this is construed as *Weak Rationality*.

## 7 Ambiguity to the rescue

To illustrate the viability of moving from *Rationality* to *Weak Rationality*, consider a very simple theory choice scenario. Take  $\alpha$  and  $\beta$  to be 1. That is, treat two profiles as being close if they differ in how they rank only one pair of adjacent theories according to a single virtue. Also take majority voting to be the method used to aggregate the individual rankings supplied by the virtues into an all-things-considered ranking. Then, for three virtues and three theories every Condorcet-like profile is close to a profile for which majority voting succeeds in mapping that profile to a transitive and complete all-things-considered ranking.

What this means is that scientists who are willing to revise the way one ambiguous virtue ranks a pair of theories then the intransitive all-thing-considered ranking can be avoided. Given that the values of  $\alpha$  and  $\beta$  are minimal and  $\gamma=1$  in this case, the theory choice among 3 theories based on 3 virtues appears to be *Weakly Rational*. And pairwise majority voting is the witness of this result. The below table documents some more values as a function of the number of theories and virtues under consideration (note that we have not included the, plausibly common, instance of choosing between only two competitors. As Okasha (2011, pp. 94-95), the Arrovian result does not hold in this case).<sup>9</sup>

---

<sup>9</sup>These numbers were computed using *Mathematica 10*. Please contact the authors for a copy of the notebooks used.

		$n$		
		3	4	5
$m$	3	1	.9826	.8836
	5	.9907	.9375	

**Table 1**

The columns in the above table denote the number of theories which need to be ranked, whereas the rows denote the number of virtues according to which these theories are assessed. Every cell contains the proportion of 1-1-close to at least one profile that majority votes maps to a transitive ordering.

Notice that for other cases than three theories ranked according to three virtues (so cases of more virtues or more theories) the behaviour of majority voting is not as ‘nice’. Only .88 of the profiles comprised of 3 virtues ranking 5 theories will be such that they are at most 1-1-close to a profile that maps to a transitive all-things-considered ranking under majority voting. Is this value of  $\gamma$  good enough for claiming that theory choice using 3 virtues to rank 5 theories is *Weakly Rational*? This seems like a subjective decision inherent in theory choice and we do not wish to argue either way. Instead, it is more fruitful to observe that the situation of majority voting is improved if we relax the definition of closeness. The below tables document these improvements:

3 virtues		$\alpha$		
4 theories		1	2	3
$\beta$	1	.9826	1	1
	2	1	1	1
	3	1	1	1

**Table 2**

3 virtues		$\alpha$		
5 theories		1	2	3
$\beta$	1	.8836	.9385	.9539
	2	.9748	1	1
	3	.9972	1	1

**Table 3**

5 virtues		$\alpha$		
3 theories		1	2	3
$\beta$	1	.9907	1	1
	2	1	1	1
	3	1	1	1

**Table 4**

5 virtues		$\alpha$		
4 theories		1	2	3
$\beta$	1	.9375	.9831	.9935
	2	.9805	.9995	1
	3	.9982	1	1

**Table 5**

Each table above corresponds to a cell in Table 1 except for the top left cell. So, Table 2 corresponds to the case of three virtues ranking four theories, Table 3 three virtues ranking five theories, and so on. The columns in these tables denote the number of ambiguous virtues considered, whereas the rows denote the number of pairwise disagreements allowed. The numbers in the cells denote the proportion of profiles in the appropriate domain which have at least one profile,  $\alpha$ - $\beta$  close to

them for which majority voting delivers a transitive all-things-considered ranking. For the case of 1 ambiguous virtue and 1 pairwise disagreement the numbers are identical to those in Table 1. However, as we move away from that cell,  $\gamma$  increases and also reaches 1 in each instance.

For a more interesting example, consider Table 3. Here we record the results for the case in which we need to rank 5 competing theories by means of 3 scientific virtues. Now, narrow in on the intersection between line 1 and column 2 in this table. This cell represents the situation in which we construe two profiles as being close if they differ in the rankings of at most 2 virtues ( $\alpha = 2$ ) and on each of these virtues the difference between the two profiles can be in the ordering of at most one pair of adjacent theories ( $\beta = 1$ ). The value in this cell is .9385. This tells us that only approx. 6% of all profiles of 5 theories ranked by 3 virtues do not have a 1-2-close profile for which majority voting delivers an overall transitive ranking. Allow two profiles to be close even if they differ with respect to the ordering of two pairs of adjacent theories, i.e. move a column to the right, and all profiles in the space have a profile 2-2-close to them for which majority voting delivers an overall transitive ranking.

In other words, assuming closeness is lax enough, all of the simple cases of theory choice are such that all profiles (even the problematic ones) will have at least one profile close to them that is mapped to a transitive all-things-considered ranking under majority voting. And notice that this possibility result has been achieved without trivialising the definition of closeness, i.e. for small values of  $\alpha$  and  $\beta$ . For instance, in the case of 4 theories and 3 virtues (Table 2) as well as in the case of 3 theories and 5 virtues (Table 4), this is the case for any value of  $\alpha$  and  $\beta$  greater than 1. In the case of 5 theories and 3 virtues (Table 3), an  $\alpha > 1$  and a  $\beta > 1$  are sufficient. The case of 4 theories and 5 virtues (Table 5) is more demanding, but since there are also more virtues and theories, slightly higher values for  $\alpha$  and  $\beta$  are still non-trivial. Finally, even in cases in which there are still profiles which are not close to any profiles for which pairwise majority voting delivers an overall transitive ranking (such as Table 5, line 1), the probability of ending up with a non-transitive profile is much lower (under any construal of closeness, i.e. any values of  $\alpha$  and  $\beta$ ) than in the case in which we take virtues to be non-ambiguous.

The interpretation we assign to this result is that as long as communities of scientists include enough subjective disagreement about how to disambiguate the shared and objective virtues, then at least some of them will be able to rationally aggregate the competing alternatives. And theory choice ends up being rational, albeit weakly so.

## 8 Conclusion

Okasha (2011) aims to prove that the situation of theory choice, if we are to construe it in Kuhnian terms, is even worse than Kuhn anticipated. It is not the case the objective elements of theory choice alone do not supply an unique algorithm for arriving at an all-things-considered ranking of theories, but rather they supply no such algorithm at all. This poses a threat to the objectivity of theory choice. In this paper we argue that Kuhn's ideas regarding the subjectivity of the rankings

generated by the ambiguous virtues offer a solution to Okasha’s challenge. Taking the ambiguity of scientific virtues seriously, what Okasha shows is that across the universal domain, regardless of the function used, it is not the case that every disambiguation will yield a rational choice. But he doesn’t show that no disambiguation will do this. We investigate some simple cases of theory choice and prove that rational aggregation is possible as long as the definition of ambiguity is sufficiently relaxed and more virtues are treated as being ambiguous. It therefore seems that the considerable effectiveness of scientific virtues depends on them being sufficiently unarticulated.

## References

- Arrow, K. (1951). *Social Choice and Individual Values*. John Wiley.
- Bradley, S. (forthcoming). Constraints on rational theory choice. *British Journal for the Philosophy of Science*.
- Gaertner, W. and Wüthrich, N. (2016). Evaluating competing theories via a common language of qualitative verdicts. *Synthese*, 193(10):3293–3309.
- Kuhn, T. (1972). Objectivity, value judgment, and theory choice. In *The Essential Tension*. University of Chicago Press.
- Lakatos, I. (1970). Falsification and the Methodology of Scientific Research Programmes. In Imre Lakatos and Alan Musgrave, editors, *Criticism and the Growth of Knowledge*, pages 91–195. Cambridge University Press.
- Morreau, M. (2014). Mr. Fit, mr. Simplicity and mr. Scope: From social choice to theory choice. *Erkenntnis*, 79(6):1253–1268.
- Morreau, M. (2015). Theory choice and social choice: Kuhn vindicated. *Mind*, 124(493):239–262.
- Okasha, S. (2011). Theory choice and social choice: Kuhn versus Arrow. *Mind*, 120(477):83–115.
- Okasha, S. (2015). On Arrow’s theorem and scientific rationality: Reply to Morreau and Stegenga. *Mind*, 124(493):279–294.
- Rizza, D. (2013). Arrow’s theorem and theory choice. *Synthese*, 191(8):1–10.
- Stegenga, J. (2015). Theory choice and social choice: Okasha versus Sen. *Mind*, 124(493):263–277.