Is Theory Choice Using Epistemic Virtues Possible?

Kit Patrick and Kate Hodesdon

January 5, 2015

Abstract

According to the popular 'epistemic virtue account' (EVA) of theory choice, we should choose between scientific theories on the basis of their epistemic virtues —empirical fit, simplicity, unity etc. Specifically, we should use a rule that aggregates theories' virtues into a ranking of the overall goodness of the theories. However, an application of Arrow's impossibility theorem shows that given plausible premises there is no rule that can aggregate theories' virtues into a theory ranking. The EVA-supporter might try to avoid the impossibility result of Arrow's theorem by asserting that we have more fine-grained distinctions between theories' epistemic virtues than initially supposed. We show that implausibly fine-grained distinctions between virtue quantities are necessary to escape the impossibility result. This is shown via novel proofs of Arrow's theorem for cases in which the quantities to be aggregated are measured on *any* combination of different scales of information, as is typically the case when aggregating epistemic virtues.

Contents

Introduction

According to the epistemic virtue account of scientific theory change (EVA) we ought to choose between scientific theories on the basis of their epistemic virtues. One theory is preferable to another if is better with respect to some collection of epistemic virtues. An important virtue is 'empirical fit': how well the theory fits the observable data. But there are also other considerations important in theory choice such as simplicity, unifying power and capacity to make novel predictions. Given information about the extent to which each theory possesses each of the epistemic virtues, EVA claims that we can apply a rule to a collection of theories that ranks them in order of overall goodness and allows us to select the best one.

This paper presents a powerful argument against EVA. Given certain plausible assumptions, an application of Arrow's theorem shows that there is *no rule* that an EVA-supporter could use to aggregate the information about how virtuous each theory is into a ranking of theories' overall goodness. This is the 'impossibility result'. The application of Arrow's theorem to theory choice takes theories to be ordered linearly according to how much of each of a number of epistemic virtues they possess. The proof of Arrow's theorem shows that there is no aggregation of theories on the basis of these virtue orderings which satisfies certain plausible assumptions. Howard Darmstadter (1974) has used Arrow's theorem to argue that there is no way to aggregate scientists' personal preferences. In a recent paper, Samir Okasha (2011) argues that the problem of theory choice on the basis of epistemic virtues is "formally identical" to the social choice problem that Arrow's impossibility theorem is typically applied to. Okasha argues, furthermore, for the plausibility of the premises of Arrow's theorem when applied to theory choice on the basis of epistemic virtues. After outlining the core commitments of EVA (§1), we show here that the EVA-supporter is committed to all of the plausible assumptions required to reach the impossibility results, except for two (§2). First of these exceptions is the 'unrestricted domain' assumption: that theories with any combination of epistemic virtues can be ranked. Second, the 'informational assumption': that for any two theories and a given virtue we have (only) ordinal information about the theories' possession of that virtue. This is to say, we know which one has more of that virtue, or that they have an equal amount of it.

In his paper, Okasha (2011) argues that in order to avoid the impossibility result yielded by Arrow's theorem, the best hope for the scientific realist is to deny the informational assumption. Since this assumption is not implied by an EVA supporter's core commitments, denying it is one possible way to save EVA. Moreover, the EVA-ist has independent reason to deny the informational assumption of Arrow's theorem, for it does seem too strong: it seems that we can know *more* than simply the relative orderings of theories for each epistemic virtue. Thus the informational assumption is perhaps the least plausible of the plausible assumptions, and some weakening of it is independently motivated. The contribution of this article is to demonstrate that every plausible weakening of the informational assumption fails to avoid the impossibility result, given EVA's core commitments (§3). This, we take it, renders problematic Okasha's conclusion that the EVA-ist had better deny the informational assumption of Arrow's theorem.

Proofs of Arrow's theorem in which all the choices to be aggregated are measured with the same strength of information are well-known to game-theorists. But what has not been proven is what aggregation rules, if any, are permissible when the choices to be aggregated are measured on mutually different information scales—some giving ordinal-scale information, some to ratio-scale, and so on. The original, paradigmatic application of Arrow's theorem is the aggregation of voters' preferences. In this case, the capacity to carry out aggregation over choices measured with mutually different scales is not an obvious desideratum. For, if we have a way of measuring one voter's choice using a cardinal-scale, then in a fair election we ought to be able to measure every voter's preference in this way. However, aggregation over input orderings measured on a number of different scales is precisely what is needed for an application of Arrow's theorem to the problem of scientific theory choice, since there is no reason to think that every epistemic virtue is measured to the same scale, or that it is meaningful to assign the same richness of information to measurements of each virtue. We argue that since our proofs that Arrow's theorem holds for the aggregation of epistemic virtues measured on any combination of informational scales, in order to avoid the impossibility result the EVA-ist must either deny the unrestricted domain assumption, or give up the EVA account altogether.

1 The Epistemic Virtue Account

The view that scientific theory choice is determined by the epistemic virtues of contending theories has been a dominant view in philosophy of science at least as far back as the scientific revolution. McMullin (2008) gives a historical survey of those philosophers and scientists who have endorsed the account. The method by which theories' virtues are aggregated to determine overall theory choice has seldom been explicitly considered; debate has focused almost exclusively on what the virtues are. However, where virtue aggregation has been considered, a view much like EVA has been endorsed. Kuhn assumes that scientists aggregate theories' virtues as EVA describes (Kuhn, 1969; Okasha, 2011). Graham Priest (forthcoming) asserts that EVA is standardly assumed in philosophy of science. Furthermore, formal methods of scientific theory choice in economics also assume that virtues are aggregated according to an algorithm much like EVA (see, e.g., Brock and Durlauf, 1999; Bonilla, 2012).

EVA has been used to defend realism from the underdetermination of theories by evidence (Psillos, 1999). But even anti-realists may endorse EVA: for example, constructive empiricists choose which empirically equivalent theory to accept on the basis of pragmatic virtues such as simplicity and explanatory power (van Fraassen, 1989, p.88). If the constructive empiricist requires there to be a rule for aggregating these virtues, then she is endorsing EVA.[1]

Many details of EVA remain unsettled. There's widespread disagreement on what the epistemic virtues are,[2] and almost nothing has been said about the aggregation stage. This oversight is understandable—it's natural to think that until we know what the epistemic virtues are we won't know how to combine them. However, our arguments apply no matter what the epistemic virtues are and no matter how numerous. We restrict our consideration to virtues that are independent of one another in the sense that it is possible for theories to possess any combination of either.

According to EVA, theory choice is a two-stage process. First, we evaluate the extent to which each theory possesses each epistemic virtue. Second, we apply a rule that aggregates this information into an overall 'ranking' of theories, which lists every theory in order of preference. This ranking allows us to select the best theory. Thus EVA-supporters are committed to the following claims:

EVA-1: There is more than one epistemic virtue.

For if there were only one epistemic virtue, then the 'aggregation' stage would be redundant.

EVA-2: The ranking of theories is determined by their epistemic virtues alone.

We intend this in a strong sense: two theories' relative ranking depends only on those two theories' virtues, and not on what other theories have been thought up. EVA-2 distinguishes EVA from Kuhn's (1969; see also Okasha, 2011) and Duhem's views (1954, see also Ivanova, 2009). They both endorse accounts of theory choice that involve epistemic virtues, but deny that there is a single rule or 'algorithm' that could be used to settle theory choice.

[1] Van Fraassen (1989) doesn't subscribe to EVA since his voluntarism denies that we should follow any rule at all. Changes in one's degrees of belief can be rational even if one is following no hard and fast epistemic rule in making those changes.

[2] See e.g. (Ladyman and Ross, 2007; Psillos, 1999; Baumann, 2005)

We can see EVA in use in Kepler's preference for Copernicus' theory over both Ptolemaic astronomy and Tycho Brahe's geocentric model. While empirical fit is an important consideration when scientists make such theory choices, empirical fit alone doesn't suffice to establish Copernicus' theory as best, for the fit of Copernicus' theory was worse than that of both Brahe and Ptolemy's theories. Copernicus' theory predicted unobserved stellar parallax and contradicted observed changes in the size of Venus; nevertheless, Kepler preferred Copernicus' theory because it was simpler. This historical case-study suggests a third commitment:

EVA-3: A small enough deficit in one epistemic virtue can always be outweighed by a big enough surplus in other epistemic virtues.

2 Arrow's Theorem

The EVA-supporter needs to find a permissible aggregation rule that will tell them which theory is best. For this, the rule must give a total ordering of the theories; i.e. for every pair of theories considered, either they are ranked equally—in which case we will say the rule is indifferent between them—or one is ranked above the other. We reserve the term 'ranking' to describe a list of theories outputted by the aggregation rule, and denote it by '>'. The ranking is transitive, so a cyclical ranking where A>B>C>A is no good. We will now give a deductive argument from premises describing constraints on permissible aggregation rules to the 'impossibility result' which says that no rules jointly satisfy them. Unlike many proofs in the social choice literature where the theorem originates, our proof is diagrammatic and intended to be non-technical.

2.1 Premises of Arrow's Theorem

Although Arrow originally intended his theorem to apply to democratic voting systems, the EVA-supporter will find the premises are plausible when applied to theory choice. As stated above, Okasha has also argued for this claim. Two of the premises, however, are not implied by the EVA account. We present these first. For each other premise we indicate the aspect of the EVA account (EVA-1 to EVA-3) that implies it.

The Informational Assumption: the only information that can be used in compiling the overall ranking is the order in which the virtues place the theories.

Unrestricted Domain (UD): the rule must provide a ranking for theories with *any combination* of virtues.[3]

[3] Michael Morreau (forthcoming) explores whether UD or something weaker is true in the theory choice case.

Pareto Indifference (PI): for any two theories, T_1 and T_2, if they are ordered as equal with respect to every virtue (i.e. neither has strictly more of any virtue) then the rule must rank them indifferent.

Weak Pareto (WP): For any two theories, if one has more of every virtue than another then the rule must rank it higher.

If either PI or WP fail, then despite two theories being ordered in the same way with respect to each virtue, the rule may rank them in the opposite way. In this case, we can hardly be said to be choosing theories on the basis of virtues alone, thus violating EVA-2.

Independence of Irrelevant Alternatives (IIA): for any two distinct theories, T_1 and T_2, the rule must determine their relative output ranking independently of all information other than the amount of virtue of T_1 and T_2.

In particular, the ranking does not depend on the virtues of any third theory, T_3. If IIA fails, then whether or not we consider T_3 will affect how we rank T_1 and T_2. This violates our strong reading of EVA-2.

Non-dictatorship (ND): A rule is impermissible if there is some virtue such that if T_1 has more of that virtue than T_2, then the rule always ranks T_1 above T_2.

If there were a dictating virtue, then EVA-3 would be violated: a theory with a slight deficit in the dictating virtue compared to another theory could never be ranked higher than that second theory, no matter how much of the other virtues it had.

2.2 Proof of Arrow's theorem for two epistemic virtues

In this section we prove Arrow's theorem for two epistemic virtues; no rules exist that aggregate the two virtues according to the constraints of §2. The proof is adapted to theory choice from the proofs of Blackorby (1984) and Geanakoplos (1996). The primary aim of this section is to give a non-technical demonstration of Arrow's theorem, and to introduce principles we will use in later arguments.

We represent how much of a particular epistemic virtue theories have by plotting the theories on a smooth line, which we label V_j. Since the informational assumption asserts that the only information that can be used by the aggregation rule is the order of theories' virtue, the aggregation rule must give the same result for any two ways of plotting theories that put them in the same order along each virtue axis. In other words, transforming one way of plotting the theories' virtues to another way of plotting which orders them the same must not yield a different aggregation rule, and we will say that the rule is 'blind' to such information-preserving transformations. So far as the aggregation rule is concerned, *order-isomorphic ways of plotting are equivalently aggregated.*

Given two epistemic virtues we can plot a theory T_1 in the two-dimensional virtue space. Divide theories around T_1 into four quadrants: I, II, III and IV (see diagram 1).

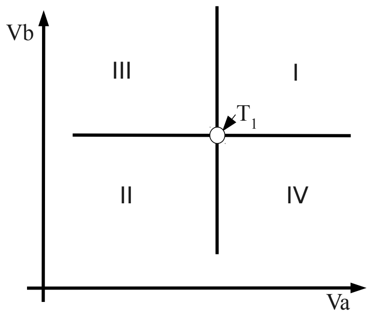


Diagram 1

A: The rule must rank each theory in quadrant I above T_1.

By Weak Pareto, if one theory has more of every virtue than another then the rule must rank it higher. Therefore, every theory in I is ranked above T_1, and likewise, every theory in II is ranked below T_1 (diagram 2).

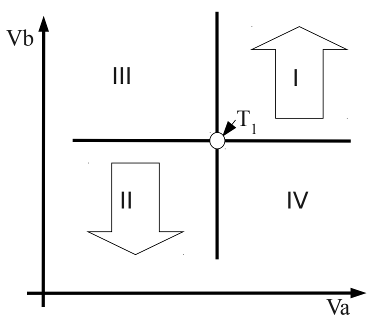


Diagram 2

B: The rule must rank each theory in quadrant III equivalently with respect to T_1.

Consider the transformation from T_1 to itself, and $T_2 \rightarrow T_2*$ where T_2* is any theory in quadrant III (see diagram 3). This transformation is order-preserving along both the V_a and the V_b axes.[4] By the informational assumption, the rule

[4]Talk of mapping a theory T to a theory T' does not commit us to claiming that there

must therefore must rank T_2 and T_2* identically with respect to T_1. A parallel argument establishes that the rule must rank all theories in IV identically with respect to T_1.

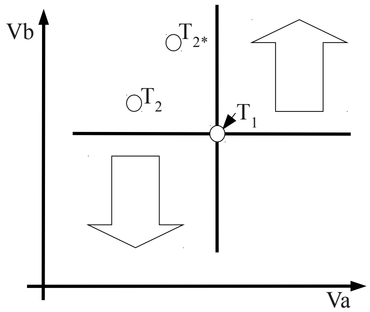


Diagram 3

C: The rule cannot rank all theories in quadrant III indifferent to T_1.

Suppose to the contrary that all theories in III are ranked indifferent to T_1. By transitivity of ranking, all theories in III must be ranked indifferent to one another. However, there are pairs of theories in III such that one has more of every virtue than the other, and so must be ranked above it (e.g. T_2 and T_2* in diagram 3). By contradiction then, theories in III cannot all be indifferent to T_1.

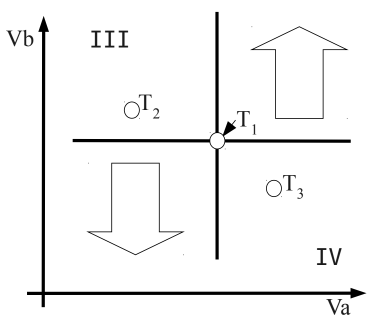


Diagram 4

D: The rule must rank theories in quadrants III and IV opposite with respect to T_1.

actually exists a theory T' distinct from T with the corresponding virtue scores: we are just using the way of speaking to describe rescaling of the virtue axes.

For any theories T_2 in III and T_3 in IV, the transformation that takes $T_2 \rightarrow T_1$, and $T_1 \rightarrow T_3$ (diagram 4) preserves the order between theories. By the informational assumption, applying this transformation does not change theories' ranking. Thus, if T_2 is ranked above T_1, then T_1 must be ranked above T_3, and vice-versa, if T_2 is ranked below T_1, then T_1 must be ranked below T_3. But T_2 is ranked above T_1 iff all theories in III are ranked above T_1, and T_3 is ranked below T_1 iff all theories in IV are ranked below T_1. Thus, all theories in III are ranked above T_1 if and only if all theories in IV are ranked below it.

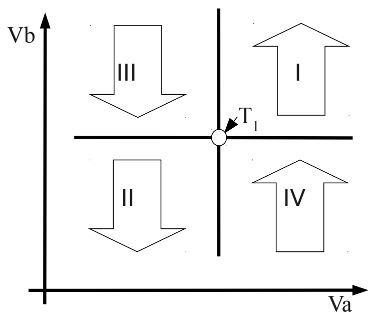


Diagram 5.1

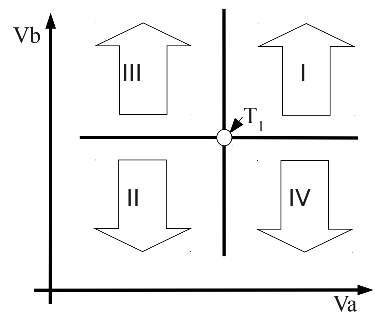


Diagram 5.2

Conclusion: the rule must make one virtue a dictator

Two options remain for the ranking of theories in these quadrants, described in diagrams 5.1 and 5.2. In 5.1, if any theory has more V_a than T_1 then it is ranked higher, whereas in 5.2 the same holds for V_b. Thus, the only available rules are dictatorships and these are impermissible. We leave it to the reader to show that any virtue that is dictator over T_1 must also be dictator over all theories in the space.

Our proof above showed that no rule which aggregates theories on the basis of two epistemic virtues satisfies the premises of Arrow's theorem. For conciseness, we shall just briefly indicate how to extend the proof to the case that there are any number n of epistemic virtues. We may generally represent a theory as a point in n-dimensional space. The two-dimensional proof establishes that on any two-dimensional pane, one virtue always dictates.[5] An argument from transitivity of dictatorship establishes that one of these 'plane-dictators' must dictate over the entire space of theories.

[5] A $V_i - V_j$ plane, for virtues V_i, V_j, is comprised of all the points of the form $(p_1, p_2, \ldots, p_i, \ldots, p_j, \ldots p_n)$ for p_1 to p_n all fixed except p_i, and p_j , which can vary. Thus it is parallel to the axes, which is a more specific notion of plane than is standardly used in geometry.

2.3 The impossibility result is fatal to EVA

The arguments above establish the impossibility result only if we can consider three or more theories at a time (this was required for step D). Okasha (2011, pp.12-13) suggests that scientists can avoid the impossibility result by only choosing between two theories at a time. However, this strategy cannot save EVA. There are two circumstances in which we decide between only two theories at a time: (i) where there are many rival theories but we restrict choice to just two at a time: e.g. scientists chose between Aristotelian physics and Newtonian physics, and only later between Newtonian physics and Relativity; and (ii) where there are only two rival theories possible: e.g. a choice between geocentrism and non-geocentrism.

In the 'many rival' case, no rule is permissible: whatever rule we use will run into trouble. The arguments of §2.1 and §2.2 show that when our rule considers all rival theories, it either contradicts one of the criteria UD, PI, WP, IIA, ND, or it ranks theories cyclically. If our rule contradicts one of the criteria when it considers all rival theories then it will still contradict that criterion when it considers only two theories. For example, if a rule ranks higher any theory with more virtue V_i when there are three theories, then it must also rank higher any theory with more V_i when there are two theories. Whether a theory considers many rival theories or just two has no bearing on whether or not that rule is a dictatorship.

On the other hand, if the rule ranks some theories cyclically when it considers all rivals at once then it won't rank those theories cyclically when it considers them two at a time, since cyclicity requires more than two theories. But, in this case the rule will be impermissible for other reasons: it will make which theory we ought to endorse depend on the order in which we consider theories. For example, consider the cyclical ranking: Relativity>Newton>Aristotle>Relativity. Suppose we start by considering Aristotle and Newton; then the rule says we ought to endorse Newton. Next we consider Newton and Relativity theory, and the rule says we ought to endorse Relativity theory. However, if we start by considering Newton vs. Relativity, then the rule says we ought to endorse Relativity. Next we decide between Relativity and Aristotle, the rule says we ought to endorse Aristotle. And surely any rule which makes which theory we endorse depend on the order in which we consider theories is impermissible. This argument undermines Baumann's claim in (2005) that we can live with cyclical ranking because discovering more evidence usually resolves the cyclical ranking.

Where there are only two rival theories available some rules will be permissible. But this won't save EVA, because EVA isn't reliable when there are just two rivals and will often lead us to endorse theories even when those theories are badly wrong. For example, non-geocentrism is less virtuous than geocentrism because non-geocentrism is the disjunction of every theory that contradicts geocentrism. It is not clear that this disjunction should be considered a theory at all, but if it is, then it will certainly not be simple and will make very few empirical claims. So if simplicity and empirical fit are important epistemic virtues, then EVA will probably, wrongly, recommend geocentrism. EVA is only a plausible account of theory choice when there are many rival theories.

3 Weakening the Informational Assumption

The EVA-supporter may try to avoid the impossibility result by weakening the 'informational assumption' criterion. Indeed, Okasha suggests that this is 'arguably the most attractive 'escape route" from the impossibility result (2011). The original informational assumption does seem too strong. It asserts that the *only* information available to the aggregation rule is theories' ordering for each virtue. But for some commonly suggested epistemic virtues, we seem to have *more* information that we can use in the aggregation process to decide the best theory. For example, we seem to know more than that Copernican theory is simpler than than Ptolemaic theory; we also know that Copernican theory is *much* simpler than Ptolemaic theory. And we don't simply know that Ptolomeic theory had better fit than Copernican theory; we know it has a *little* better fit. This extra information leads us to believe that EVA-supporters who take simplicity to be an epistemic virtue should weaken the informational assumption: it is not merely the ordering of virtues that decides the output ranking, but sometimes relative amounts virtue that weigh in the balance.

If the EVA-supporter weakens the information assumption enough then they can avoid the impossibility result. Here's one way. Suppose we could have absolute numerical scores for each theory's simplicity, empirical fit, unity etc.; Copernican theory might have 100 'simplons' and 45 'empiricons' compared to Ptolemaic theory's 40 simplons and 50 empiricons. Given such rich information, it's easy to find permissible, non-cyclic rules. One permissible rule is to multiply the absolute numbers of each virtue a theory possesses. The theory with the highest total score wins.

If the EVA-supporter weakens the informational assumption to this extent, their account must implausibly presuppose that we have access to information about theories' virtues that is just not attainable. Clearly, we don't have absolute numerical values for each theory's simplicity, empirical fit, unity etc. It's not clear whether this is meaningful. In §4 - §6 we argue for our main result, that there is no desirable 'Goldilocks' level of information: information is either too impoverished to avoid Arrow's impossibility result or too rich to obtain. In the current section we lay the groundwork for this argument by describing four weakenings of the informational assumption corresponding to four information scales.

The weakened informational assumptions are presented in a similar way to the original assumption of ordinality. We introduced this 'ordinal-scale' information assumption with the constraint that the rule give the same ranking for any two ways of plotting theories in the virtue-space that preserved their order on each virtue axis. We now present weakened informational assumptions alongside their corresponding equivalence classes of plottings that any rule constrained by the assumption must treat in the same way. If we have richer information about theories' virtues, we can make finer-grained distinctions between different ways of plotting theories. Transforming all theories equivalently is just a kind of rescaling of virtue scores, like moving from degrees Farenheit to degrees Celcius.

Cardinal-Scale Measurability: *The information available to the aggregation rule is the order in which theories fall and the ratios of their differences in virtues.*

A virtue is cardinal scale measurable if we have information about the order of theories with respect to that virtue and about facts of the form 'the difference between T_1's virtue and T_2's virtue is three times the difference between T_1's virtue and T_3's virtue. If a virtue is cardinal-scale measurable then any two ways of plotting a profile of theories are equivalent only if they are related by a transformation from v to v+b. These are just those ordinal-scale transformations defined by linear equations.

Unit-Scale Measurability: *The information available to the aggregation rule is the order in which theories fall and absolute differences between virtue scores.*

A virtue is unit-scale if in addition to all the information available for cardinal-scale virtues we have access to facts such as 'the difference between T_1's virtue and T_2's virtue is 4.5'. For example, converting temperature in Celsius to Kelvin preserves the absolute differences in temperatures. If a virtue is unit-scale measurable then two ways of plotting a profile are equivalent if they are related by a transformation of the form $v \rightarrow v+b$. These are just the cardinal-scale transformations with a gradient of 1.[6]

Ratio-Scale measurability: *The information available to the aggregation rule is the order in which theories fall and ratios between their virtue scores.*

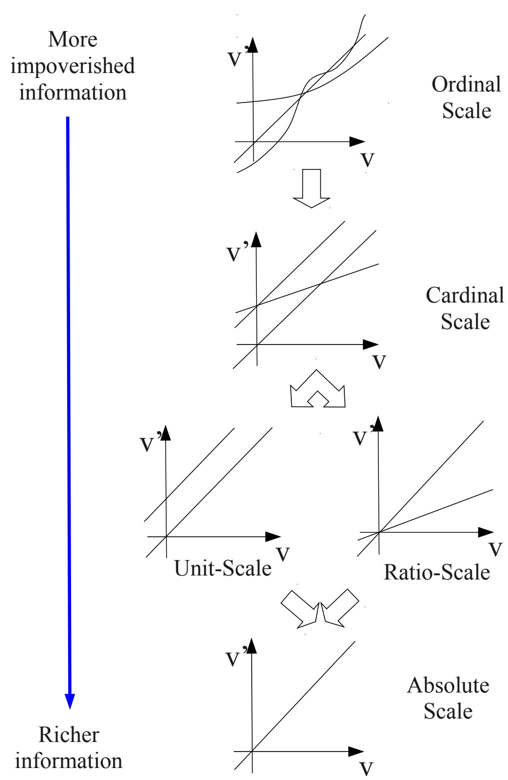
A virtue is ratio-scale measurable if in addition to all the information available for cardinal-scale virtues, we have information of facts of the form 'T_1 has twice as much virtue as T_2'. Unlike ordinal-scale, cardinal-scale and unit-scale virtues, ratio-scale virtues may take the value 0. For example, Kilograms and pounds give different representations of mass, and are related by a ratio-preserving

[6]Having all virtues unit-scale is a special case of Sen's 'cardinal-scale-unit-comparability'.

transformation: a 10lb box is twice as heavy as a 5lb one, and remains twice as heavy when we weigh them in kg.

If a virtue is ratio-scale then any two ways of plotting a profile are equivalent only if they are related by a transformation of the form v→av for any a $>$ 0. These are just the cardinal-scale transformations for which b $=$ 0.

There are important logical relations between the different scales: cardinal-scale transformations are a limiting case of ordinal-scale transformations. Both unit-scale and ratio-scale transformations are limiting cases of ordinal-scale transformations. Finally, absolute-scale transformations are limiting cases of both unit-scale and cardinal-scale. These logical relations can be seen using graphical examples:



For the rest of this paper we assume that we have unit-scale, ratio-scale or weaker information for all epistemic virtues. This seems very plausible for the commonly suggested epistemic virtues such as simplicity and unity. Indeed, it seems overgenerous to the EVA-supporter. It's difficult to find any epistemic virtues that we plausibly have even ratio-scale or unit-scale information about. Okasha (2011) provides an example of how we might get ratio-scale information about empirical fit. However, his example concerns line-fitting exercises, and it's not clear how we can find such specific information for choices between more complex theories such as Copernican and Ptolemaic theory. At any rate, the burden of proof lies with the EVA-supporter to show that we have richer information than this.

4 No rules are permissible for cardinal-cardinal theory choice

If both virtues are all cardinal-scale, then no aggregation rule is permissible, since all transformations in the Arrow proof we gave initially (§2.2) can be cardinal-scale. Thus a parallel argument shows that when there are more than two cardinal-scale virtues, there is a dictator over any plane. The same argument also implies that there is no aggregation rule for theory choice on the basis of any combination of cardinal-scale and ordinal-scale virtues.

5 No rules are permissible for unit-cardinal theory choice or ratio-cardinal theory choice

In this section, we consider theory choice on the basis of *different* strengths of information about the different virtues. This hasn't been discussed in the extensive literature on Arrow's theorem. As mentioned above, the reason for this is probably that the Arrow's theorem literature has focused on democratic voting, and we naturally assume that we have the same depth of information for all voters. If we find a way to get say ratio-scale information about one voter, then we can also get ratio-scale information about any other voters. Quite the contrary is true for the theory choice case: *in this context, there is no reason to suppose that we have the same depth of information about simplicity, unity, empirical fit, scope etc.* It is therefore worth investigating whether any rules are available when the virtues have different information scales. We present proofs investigating this novel situation of mixed information scales. Of course, our results apply to any situation in which an aggregated ordering is decided upon on the basis of factors that are measured on different scales. In the following sections we describe the availability of aggregation rules for each two-virtue combination of the scales defined above. We will use the naming convention 'unit-cardinal measurability' (or simply 'unit-cardinal') to name the scale of information available to the aggregator in a two-dimensional virtue space in which one virtue is measured on a unit-scale axis and the other on a cardinal-scale axis. The same convention applies for any pair of information scales.

5.1 Unit-Cardinal Theory Choice

We show that when there is one unit-scale virtue V_u, and one cardinal-scale virtue V_c, no rule is permissible. Let T_1 be a theory at any point in the virtue score space. Draw quadrants I-IV around T_1 as usual.

A: The rule must rank all theories in quadrant I above T_1 and everything in quadrant II below T_1.

This follows step A in §2 exactly.

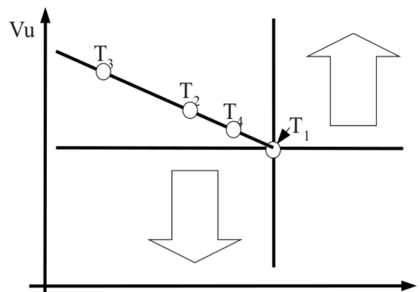


Diagram 7

B: There is a straight line through III of theories ranking identically with respect to T1

Let T_2 be a theory in quadrant III. Consider the transformation from $T_1 \rightarrow T_2$ and $T_2 \rightarrow T_3$, where T_3 is chosen such that the difference between T_1 and T_3 is twice the difference between T1 as T_2 on both virtue axes (see diagram 7). This transformation preserves absolute differences of virtue scores, so is a unit-scale transformation on both virtues V_u and V_c. Since unit-scale transformations are a special case of cardinal-scale transformations, this transformation also is cardinal-scale on Vc. Therefore, the rule must rank T_1 with respect to T_2 the same as it ranks T_2 with respect to T_3. If T_2 is ranked above T_1 then T_3 must be ranked above T_2; thus $T_3 > T_2 > T_1$. By transitivity, $T_3 > T_1$. Likewise, if T_2 is ranked below or indifferent to T_1, then T_3 must also be ranked below or indifferent to T_1 respectively.

Repeat this procedure to construct a densely packed set of points along a straight line from T_1 passing through T_2 in quadrant IV. We will impose a stipulation of completeness, or 'no sudden jumps' on our rule. This says that if there are two theories T' and T" on the line, both of which are ranked the same with respect to T_1, such that there is a theory lying between them and infinitesimally close to both, then that infinitesimally close theory must be ranked the same as T' and T" with respect to T_1. With this stipulation in place, every theory on this line must be ranked identically with respect to T_1.

For example, consider the transformation from $T_1 \rightarrow T_4$ and $T_4 \rightarrow T_2$ where T_4 is halfway between T_1 and T_2 on both virtue axes. This transformation preserves absolute differences of virtue scores. Therefore, T_4 and T_2 must be ranked identically with respect to T_1.

C: Every point in quadrant III must be ranked the same with respect to T_1. Let T_5 be any point on the line of points ranked identically with respect to T_1. Consider the transformation that fixes T_1 and sends T_5 to T_6, for T_6 any point in quadrant III with the same amount of V_u as T_5 (see diagram 8). This transformation doesn't change theories V_u scores and is a cardinal-scale transformation on the V_c axis. Therefore, the rule must rank T_5 and T_6 identically with respect to T_1. A parallel argument shows that any point in quadrant III with the same V_u score as some point on the line must be ranked identically with respect to T_1. That is, the rule must rank all points in quadrant III identically with respect to T_1.

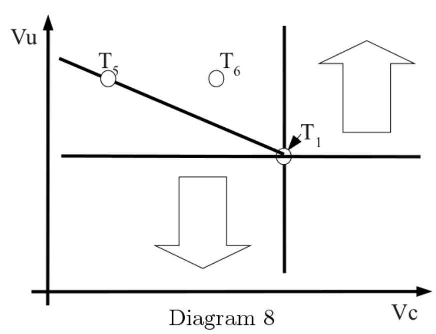


Diagram 8

From here the proof proceeds exactly as from stage C in §2. Since cardinal-scale transformations are a special case of ordinal-scale transformations, the same proof will also show that there is no permissible rule for unit-ordinal theory choice.

5.2 Ratio-Cardinal Theory Choice

When there is one ratio-scale virtue, V_r, and one cardinal-scale virtue, V_c, no rule is permissible. Let T_1 be a theory at any point in the virtue score space. Draw quadrants I—IV about T_1 as usual. Let T_2 be a theory in quadrant III. For brevity, we introduce a labelling convention: for theory T_1 and virtue V_a, label T_1's virtue V_a score 'T_{1a}'.

A: The rule must rank all theories in quadrant I above T_1 and everything in quadrant II below T_1.

This follows A in §2.

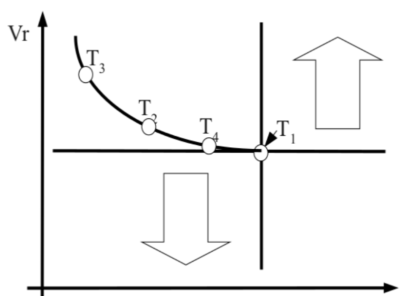


Diagram 9

B: There is a line through III of theories ranking identically with respect to T_1.

Consider any point T_2 in quadrant III relative to T_1. The ratio-scale co-ordinate for T_1 is T_{1r} and the ratio-scale co-ordinate for T_2 is T_{2r}. We can express T_{2r} as a multiple of T_{1r} to obtain the ratio between them: trivially, $T_{2r} = \frac{T_{2r}}{T_{1r}} T_{1r}$. Since T_2 has less of V_r than T_1, $\frac{T_{2r}}{T_{1r}}$ is less than 1.

Without loss of generality, let T_2 be ranked above T_1 . We will show that there is a line starting from T_1 and passing through T_1 such that every point on it is ranked above T_1. We can plot the point T_3 that is such that the ratio between T_1 and T_2 is the same as the ratio between T_2 and T_3 along the ratio-scale axes. There is thus a transformation mapping T_2 to T_3 and T_1 to T_2 that is ratio-preserving along all axes. Since $T_{2r} = \frac{T_{2r}}{T_{1r}} T_{1r}$, in order to preserve the ratio of $\frac{T_{2r}}{T_{1r}}$ between points on the line we have $T_{3r} = \left(\frac{T_{2r}}{T_{1r}}\right)^2 T_{1r}$ (see diagram 8). But this means that $T_{3r} = \frac{T_{2r}}{T_{1r}} T_{2r}$. By the same process of preserving ratios we can find the co-ordinate for T_3's virtue score for the cardinal-scale virtue. Since ratio-scale transformations are a special case of cardinal-scale transformations, the transformation $T_1 \rightarrow T_2$, $T_2 \rightarrow T_3$ is also is cardinal-scale on V_c. For this reason, the rule must rank T_1 with respect to T_2 the same as it ranks T_2 with respect to T_3. By transitivity, T_2 and T_3 are ranked identically with respect to T_1, i.e. both above it.

To show that there is a line of points all of which are ranked above T_1 we need to show that there are theories in between T_1 and T_2 that are densely packed along the line and are ranked above T_1. This is achieved iteratively by taking midpoints of points already known to be on the line. So, we need to find a point $T_4 = (T_{4r}, T_{4c})$ in between T_1 and T_2 which is such that the ratio between T_1 and T_2 and between T_2 and T_3 is the same along the ratio-scale axes. Given this requirement, T_{4r} must therefore be $(T_{2r} \;/\; T_{1r})^{\frac{1}{2}}\ T_{2r}$.

We can find the virtue score of T_4 for the other axis by likewise finding the point in between T_{1c} and T_{2c} that is related to these in the same ratio-preserving way. All these points trace out a line, c=(T_{1r} T_{1c})/r.

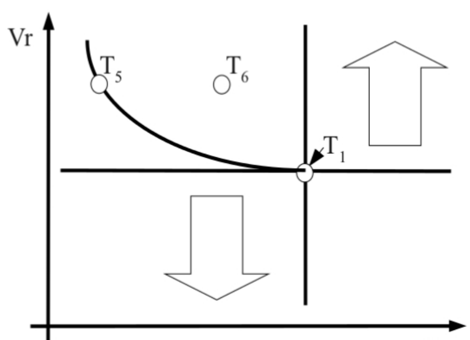


Diagram 10

C: Every point in quadrant III must be ranked the same with respect to T_1.

Consider the transformation fixing $T_1 \rightarrow T_1$ and taking $T_5 \rightarrow T_6$ where T_5 is any theory on the line, and T_6 is any theory in quadrant III with the same V_r score as T_5 (see diagram 10). This transformation yields no change in V_r , and is a cardinal-scale transformation on V_c. Therefore, the rule must rank T_5 and T_6 identically with respect to T_1; in this case, both are ranked above it.

A parallel argument shows that any point in quadrant III with the same V_r score as some point on the line must be ranked identically with respect to T_1. So the rule must rank all points in quadrant III with positive scores for ratio-scale virtue V_r identically with respect to T_1. From here the proof proceeds exactly as from stage C in §2. Since cardinal-scale transformations are a special case of ordinal-scale transformations, the same proof will show that there is no permissible rule for unit-ordinal theory choice.

5.3 Intermediate conclusions

We've already established enough to entail conclusions that will be unacceptable to the EVA-supporter. For brevity, distinguish between 'weak' virtues—those about which our information is at most as rich as cardinal-scale; and 'strong' virtues—those about which we have richer than cardinal-scale information, up to and including unit- and ratio-scale information. If all virtues are weak there is no permissible rule, as established in §4. If there is only one strong virtue and all other virtues are weak then there can be no permissible rule either. For in this case, there will only be two sorts of planes: those involving two weak virtues and those involving one strong and one weak virtue. We can use the arguments above to show that on both sorts of planes one of the virtues is a 'dictator', and then use the n-virtue generalisation of Arrow's theorem to show that there is no permissible aggregation rule.

We have not yet proved anything about virtue spaces with more than one strong virtue. To recap, there is never a rule to aggregate between pairs of weak virtues: but it is possible that there is a rule to aggregate between two strong ones. We now show that if there is more than one strong virtue and all other virtues are weak, then a theory's weak virtues almost never make any difference to that theory's ranking. Weak virtues are 'marginalised': that is, they are only used to decide between two theories that are judged indifferent on the basis of all their strong virtues (and in other very rare circumstances specified below). This is to say that a lexicographic ordering results, which the strong virtue dominates. To prove this, it's sufficient to consider the simple three virtue case with one weak virtue 'V_w' and two strong virtues 'V_{s1}' and 'V_{s2}'. On the V_w-V_{s1} or V_w-V_{s2} planes there are three possible combinations of plane-dictatorships:

(i) The weak virtue plane-dictates over both strong virtues.
(ii) The weak virtue plane-dictates over one strong virtue, and is plane-dictated over by the other.
(iii) The weak virtue is plane-dictated over by both strong virtues.

Any rule that satisfies (i) is a dictatorship. Any rule that satisfies (ii) has cyclical output rankings. But any rule that satisfies (iii) marginalises the weak virtue. We prove each claim in turn.

We know from the generalized proof of Arrow's theorem that amongst weak virtues there is always a transitive chain of plane-dictators, with one as topmost virtue. The dictator on the V_{s1}-V_{s2} plane is either the dictator over all weak virtues or it is plane-dictated over by such a virtue. Either way, we have a dictating virtue for each plane.

Claim (i): Any rule where the weak virtue plane-dictates over both strong virtues is a dictatorship of V_w. To show this, we'll prove that for two theories T_1 and T_2 such that T_2 has more of the weak virtue V_w than T_1 does, T_2 is ranked higher than T_1. For, consider the theory T' that has the following properties; T' is on the same V_{s1}-V_{s2} plane as T_1, and T' is on the same V_w-V_{s2} plane as T_2. T' has more of virtue V_{s2} than T_1 does, but less of V_w than T_2 does.

Then, since T' is identical to T_1 except for having greater V_{s2}, it is ranked above-or-indifferent to T_1, by Weak Pareto. But since T' is on the same V_w-V_{s2} plane as T_2, and has less V_w than it, T' is ranked below T_2. And so, we have $T_2 > T' \geq T_1$, establishing, by transitivity, the result we wanted.

Claim (ii) Whenever the weak virtue plane-dictates over one strong virtue, and is plane-dictated over by the other, the only aggregation rules give cyclical rankings, and so are impermissible.

To generate such cyclical rankings, consider a single V_{s1}-V_{s2} plane. Find two theories, T_1 and T_2 that are not ranked indifferent to one another, such that one has more V_{s1} and the other more V_{s2}. Such pair of theories exist by the original Arrow proof (see §3 D). Without loss of generality, let T_2 have less V_{s1} and more V_{s2} than T_1, and be ranked above T_1. We now show that there is some theory T' that must be ranked above T_2 and below T_1. Choose T' so that it has greater V_w than T_2, and the same V_{s2}. Since V_w dictates over V_{s2}, T_3 must be ranked above T_2. Since V_{s1} dictates over V_w, T_1 must be ranked above T_2. In summary, $T_1 < T_2 < T' < T_1$; a cylical output ranking. The same argument applies *mutatis mutandis* to show that when T_2 is ranked below T_1, there must be a cyclical output ranking.

Claim (iii): If the weak virtue is plane-dictated over by both strong virtues, then the weak virtue is marginalized: we only appeal to the weak virtue to decide the ranking of theories which have identical scores for both the strong virtues.

There are three options available to the EVA-supporter. All are unattractive. First, they could bite the bullet and accept that virtues such as simplicity and unifying power which were previously thought important in theory choice are marginalized. But EVA is attractive largely because it seems plausible that simplicity and unifying power *do* typically influence theory choice. Second, EVA-supporters could claim that simplicity and unifying power are strong virtues; we have richer than cardinal-scale information about them. This seems implausible: claims such as 'T_1 has 10 simplons more than T_2' or 'T_1 is 1.5 times as simple as T_2' seem unknowable or meaningless. Third, EVA-supporters could argue that there is an absolute scale virtue. We rejected this claim as implausible in §3.

6 No rules for unit-unit theory choice, ratio-ratio theory choice and unit-ratio theory choice

In this section, we argue that there are no rules permissible for the EVA-supporter for unit-unit, ratio-ratio and unit-ratio theory choice. There are rules compatible with the original Arrow premises (§2.2) in these cases, but all are in serious tension with two assumptions that the EVA-supporter will find hard to deny. These assumptions are described and motivated in §6.1. Then we describe all rules compatible with one assumption (§6.2) and show that they are all incompatible, with the other assumption (§6.3).

6.1 Two assumptions plausible to the EVA-supporter

The EVA-supporter will find it hard to deny:

Trade-off assumption: Two theories with different virtue scores can be ranked indifferent to one another.

Given the EVA-supporter's other commitments, denying trade-off is implausible: the EVA-supporter is already committed to saying that for any two theories, the one with slightly less of some virtue can still be ranked higher if it has sufficiently more of other virtues, else the rule is a dictatorship. If a slight deficit in one virtue can be overcompensated for by a big surplus in other virtues, it seems likely that a deficit in one virtue can be compensated for by a slightly smaller surplus in other virtues. The EVA-supporter will also find it hard to deny:

Strong Pareto (SP): For any two theories, if one has more of at least one virtue and the same amount of all others, then the rule must rank it above the other theory.

Strong Pareto is independently plausible. If, say, simplicity is an epistemic virtue then it seems that a theory's simplicity should also be a reason to prefer it to another theory with the same amount of other virtues. Strong Pareto will be especially plausible to those who appeal to EVA to avoid the underdetermination of theory by evidence. According to all versions of the underdetermination of theory by evidence, there are sometimes two theories that make importantly different claims about the world but have equal empirical fit. This is supposed to challenge realism: if we decide between theories solely on the basis of empirical fit then we cannot know which theory to endorse. A standard realist response has been to appeal to extra-empirical virtues such as simplicity and unity. Thus, such EVA-supporters are committed to something much like Strong Pareto. In fact, even weakening Strong Pareto significantly won't save EVA, as shown in §6.3.

6.2 All rules compatible with trade-off

Given the premises in §2.2 and the trade-off assumption for unit-unit, ratio-ratio and unit-ratio theory choice respectively, we now show which rules are permissible.

6.2.1 Rules for unit-unit theory choice

The only permissible rules where both virtues are unit-scale are *weighted utilitarian rules*. A weighted utilitarian rule works by adding up the scores of a theory for each of the virtues. One theory is ranked above another in the output just in case that theory has a higher weighted utility. The virtues may be differently weighted so that an increase in, say, simplicity might be less significant than the same increase in empirical fit. Dictatorship is a limiting case of the weighted utilitarian view, where one virtue has all the weight and the other virtues have none.

The trade-off assumption asserts that some theories with different virtue score are ranked indifferent to one another. Choose T_1 and T_2 to be two such theories. Construct a straight line passing through T_1 and T_2 in exactly the way described in §5.1 (see diagram 11). This is an 'indifference line': theories on this line can be expressed as having the same weighted utility score W=Ax+By, where A and B are constants determined by the gradient of the line. Since they have the same weighted utility score, every theory on this line must be ranked the same with respect to T_1, namely, indifferent to it. And since indifference is transitive, every theory on the line must also be ranked indifferent to every other theory on the line.

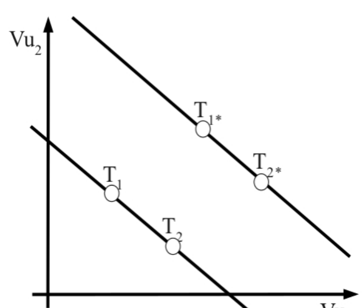


Diagram 11

Suppose that V_{u1} and V_{u2} (on diagram 11) are two unit-scale measurable virtues. Take T_{1*} any point not on the indifference line, and consider the transformation from $T_1 \rightarrow T_1{}^*$ and $T_2 \rightarrow T_2{}^*$, where $T_2{}^*$ is chosen so that the absolute difference between T_1 and $T_1{}^*$ is the same as that between T_2 and

T_2*. Because the rule is blind to these transformations it follows from the fact that T_1 is ranked indifferent to T_2 that T_1* is ranked indifferent to T_2*. We can now construct a new indifference line passing through T_1* and T_2*: any theory on this line must be ranked indifferent to every other. This line will be parallel to the first line, but will have a higher weighted utility score. By repeating this step, we can establish indifference lines parallel to the first throughout the virtue score space. Since each indifference line corresponds to a weighted utility score, Weak Pareto entails that theories on indifference lines with a higher weighted utility score must be ranked higher than those theories with a lower weighted utility score.

In the unit-unit case where the trade-off assumption is false, the only permissible rules are dominated weighted utilitarian rules. For theories that do not share a weighted utility score, a dominated weighted utilitarian rule is just a weighted utilitarian rule: any theory with a higher score (on a higher line) is ranked above a theory with a lower score. But since denying the trade-off assumption means that theories sharing a weighted utility score (those on the same line) cannot be indifferent to one another there must be some virtue such that for any two theories with the same weighted utility score whichever theory has more of that virtue is ranked above the theory with less of that virtue.

6.2.2 Rules for ratio-ratio theory choice

The ratio-ratio case follows almost exactly the same as the unit-unit case. The only permissible rules are Cobb-Douglas rules according to which theories with a higher score W=$x^a y^b$ (for constants a and b) are ranked higher than those with a lower W score. Given the trade-off assumption, theories with exactly the same W score are ranked indifferent to one another. For proofs that these are the only rules compatible with Arrow's assumptions see (Baumann, 2005; Tsui and Weymark, 1997).

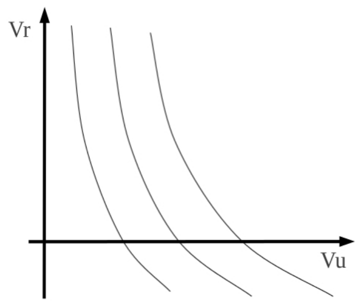


Diagram 12

6.2.3 Rules for unit-ratio scale theory choice

The case where one virtue is unit-scale and the other ratio-scale follows the same pattern as that of ratio-ratio theory choice and unit-unit theory choice. In each of these cases, theory ranking is preserved under transformations that preserve, respectively, the ratios or unit differences between all theories with respect to each axis. In ratio-unit theory choice theory ranking is preserved under those transformations that are ratio-preserving along the V_r-axis, and unit-preserving in the V_u-axis. Given two theories ranked indifferent by ratio-unit aggregation, we construct an indifference line by transforming the indifferent points in a ratio-preserving way along the ratio axis, and a unit-preserving way along the unit axis. This gives us indifference lines of the form $x = \frac{1}{x-n}$ where y is the unit-scale variable and x the ratio-scale variable (Diagram 12). Note that this line is invalid when $x = n$, so the unit-ratio rule is going to violate unrestricted domain.

6.3 These rules are impermissible to the EVA-supporter

All these rules are either incompatible with Strong Pareto, or incompatible with the motivation for Strong Pareto. The rules described in §6.1 entail:

Continuity:[7] For any theory T_1, if we draw a line through virtue space such that theories at one end of the line are ranked above T_1, and theories at the other end are ranked below T_1 then there is some point along this line at which theories are ranked indifferent to T_1. To see that this follows from the rules

[7]This notion of continuity is strictly weaker than the notion of continuity given by Sen (1977) and others, *viz.* that for any theory T_1 the set of theories ranked above T_1 and the set of theories ranked below T_1 are both closed sets. Several discussions of Arrow's theorem treat Sen's notion as equivalent to the notion of continuity we use (e.g. Roemer, 2004).

described in §6.2, notice that any such line must pass through T_1's indifference line. Thus, there must be some point on it indifferent to T_1.

Continuity is incompatible with Strong Pareto. Our proof for this is adapted from Tsui and Weymark (1997). They discuss only the case where all virtues are ratio-scale. We extend this proof beyond ratio-ratio theory choice, and weaken the assumptions required for the proof to work. Suppose there are two virtues: 'V_r' which is ratio-scale or weaker, and 'V_s' which is either unit-scale, ratio-scale or weaker. We obtain a contradiction by showing that any theory 'T_0' with none of virtue V_r can be neither above, indifferent to, nor below any theory 'T_+' with positive amounts of both V_r and V_s.

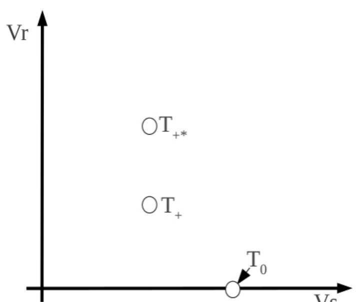


Diagram 13

Claim: T_0 cannot be ranked indifferent to T+

Suppose the contrary: T_0 is ranked indifferent to T_+. Let T_{+*} be any theory that has the same amount of V_s and more V_r than T_+ (see diagram 13). The transformation from $T_+ \rightarrow T_{+*}$ and $T_0 \rightarrow T_0$ preserves the ratios of theories on V_r and involves no transformation on V_s. Therefore the rule must rank T_0 versus T_+ the same as it ranks T_0 versus T_{+*}. Since by supposition, T_0 is ranked indifferent to T_+, T_0 must also be ranked indifferent to T_{+*}. Because indifference is transitive, T_+ must be ranked indifferent to T_{+*}. On the other hand, Strong Pareto entails that T_{+*} is ranked above T_+, since it has more of one virtue and less of none. In summary, if T_+ is ranked indifferent to T_0, then T_{+*} is ranked both indifferent to and above T_+. This is impermissible. Therefore, the aggregation rule cannot rank T_+ indifferent to T_0 after all.

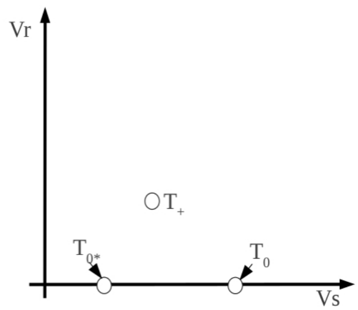


Diagram 14

Claim: T_0 cannot be ranked above T+

Suppose the contrary: T_0 is ranked above T+. Consider any theory T_{0*} that has less Vs than T_0 and T_+, and V_r (see diagram 14). By weak Pareto, T_{0*} will be ranked below T+ since it has less of every virtue. Therefore, there is a line from T_0 to T_{0*}, at one end of which theories are ranked above T_+ and at the other end of which theories are ranked below T+. The continuity assumption asserts that there must be a point along this line indifferent to T_+. But, by the argument in 1 above, there is no such point. Therefore, T_0 cannot be ranked higher than T_+.

Claim: T_0cannot be ranked below T_+

Suppose the contrary: T_+. is ranked above T_0. Strong Pareto entails that T_0* is ranked below T_0, since it has less of one virtue and more of none. Therefore, there is a line from T_0 to T_0*, at one end of which theories are ranked above T_0 and at the other end of which theories are ranked below T_0. The continuity assumption asserts that there must be a point along this line indifferent to T_0. But this cannot be: we already know by the argument in A that there is no theory with positive simplicity and unity that is indifferent to T_0. And the only point on this line that does not have positive simplicity and unity is T_0*. But we know by Strong Pareto that T_0* is not ranked indifferent to T_0.

Thus, theory T_0 cannot be ranked above, indifferent, or below T_+. This contradicts Unrestricted Domain, which requires any rule to rank all the theories against all other theories. Therefore, given the trade-off assumption and Strong Pareto, there is no permissible rule for ratio-ratio and unit-ratio theory choice.

This proof doesn't show that there is no permissible rule for unit-unit theory choice. In unit-unit theory choice, the weighted utilitarian rue is the only rule compatible with the trade-off assumption. And, although this rule entails continuity, continuity is only inconsistent with strong Pareto when at least one virtue is ratio-scale or weaker. In unit-unit theory choice, no virtues are ratio-scale or weaker. Nevertheless, the unit-unit rule undermines many EVA-supporters' reason for accepting Strong Pareto. For, as we argues above, many EVA-supporters accept EVA to avoid underdetermination of theory by evidence. If the weighted utilitarian rule is accepted then an even worse form of underdetermination arises, for each theory sits on an indifference line dense with mutually indifferent theories. This is considerably stronger and more threatening to realism than the relatively harmless trade-off assumption, which claimed that there are at least two theories that ought to be ranked indifferent to one another: for if the weighted utilitarian rule is correct then every theory ought to be ranked indifferent to numerous others. Thus a weighted utilitarian account of theory choice replaces underdetermination of theory by evidence with the even more threatening underdetermination of theory by empirical virtues. This will be unacceptable to many EVA supporters. And we suggest that they should declare the weighted utilitarian rule impermissible.

In §6 we have argued that for unit-unit, ratio-ratio and unit-ratio theory choice all rules are implausible. This has broader consequences. Suppose that there are more than two virtues, al of which are unit-scale or ratio-scale or weaker. Arguments parallel to those showing that there is no permissible rule with unit-unit, ratio-ratio and unit-ratio can be used to show that any plane in this multi-virtue theory choice must have a 'dictator'. So, no matter how many virtues there are or how much information we have about them, there are no aggregation rules.

7 Conclusion

We have shown that the EVA-supporter can only avoid the impossibility result either by denying the informational assumption or by denying unrestricted domain. We have also shown that the EVA-supporter cannot plausibly avoid the impossibility result by denying the information assumption. For information is either too impoverished to avoid the impossibility result, or too strong to plausibly obtain. Therefore, it seems that the best options are either to deny unrestricted domain or abandon EV A altogether. Our argument might be extended to apply to other instances of theory choice using virtues, such as truthlikeness (Zwart and Franssen, 2009), inference to the best explanation (Lipton, 1993) and even choices between philosophical theories *Ladyman and Ross, 2007, p.17).

References

Baumann, P. (2005). Theory choice and the intransitivity of 'is a better theory than'. *Philosophy of Science 72*, 231–240.

Blackorby, Charles, D. D. and J. A. Weymark (1984). Social choice with interpersonal utility comparisons: A diagrammatic introduction. *International Economic Review 25*, 327–356.

Bonilla, Z. J. (2012). The economics of scientific knowledge. In U. Mäki (Ed.), *Handbook of the Philosophy of Science: Philosophy of Economics*. Elsevier.

Brock, W. A. and S. N. Durlauf (1999). A formal model of theory choice in science. *Economic Theory 14*, 113–130.

Darmstadter, H. (1974). Better theories. *Philosophy of Science*.

Duhem, P. (1954). *The Aim and Structure of Physical Theory*. Princeton University Press, Princeton.

Geanakopolos, J. (1996). Three brief proofs of arrow's impossibility theorem". *Economic Theory 26*, 211–5.

Ivanova, M. (2009). Pierre duhem's good sense as a guide to theory choice. *Studies in History and Philosophy of Science Part A 41*, 58–64.

Kuhn, T. (1969). *The Structure of Scientific Revolutions*. Chicago: University of Chicago.

Ladyman, J. and D. Ross (2007). *Every Thing Must Go: Metaphysics Naturalized*. Oxford University Press.

Lipton, P. (1993). *Inference to the Best Explanation* (2 ed.). Routledge.

McMullin, E. (2008). The virtues of a good theory. In S. Psillos and M. Curd (Eds.), *The Routledge Companion to the Philosophy of Science*, pp. 499–508. Routledge.

Morreau, M. (forthcoming). Theory choice and social choice: Kuhn vindicated. *Mind*.

Okasha, S. (2011). Theory choice and social choice: Kuhn versus arrow. *Mind 120*, 83–115.

Psillos, S. (1999). *Scientifc realism: how science tracks truth*. Routledge, London.

Roemer, J. E. (2004). Eclectic distributional ethics". *Politics, Philosophy and Economics 3*, 267–281.

Sen, A. (1977). On weights and measures: Informational constraints in social welfare analysis. *Econometrica 45*.

Tsui, K.-y. and J. A. Weymark (1997). Social welfare orderings for ratio-scale measurable utilities. *Economic Theory 10*, 241–56.

van Fraassen, B. (1989). *Laws and Symmetry*. Clarendon Press, Oxford.

Zwart, S. and M. Franssen (2009). An impossibility theorem for verisimilitude. *Synthese 158*, 75–92.