
SECOND THOUGHTS ABOUT MY FAVOURITE THEORY

BY

JOHAN E. GUSTAFSSON

Abstract: A straightforward way to handle moral uncertainty is simply to follow the moral theory in which you have most credence. This approach is known as My Favourite Theory. In this paper, I argue that, in some cases, My Favourite Theory prescribes choices that are, sequentially, worse in expected moral value than the opposite choices according to each moral theory you have any credence in. In addition, this problem generalizes to other approaches that avoid intertheoretic comparisons of value, such as My Favourite Option, the Borda Rule, Variance Normalization and the Principle of Maximizing Expected Normalized Moral Value.

There are droves of moral theories. Even so, you probably find one of them more plausible than the others. But you're not sure that theory is correct: You find that some of its rivals have at least some plausibility. Suppose, then, that you have some credence in two or more moral theories. What do you do when these theories prescribe different, incompatible courses of action? For instance, it might be that, according to one of these theories, you *ought* to do something, but, according to another, you *ought not* to do it. A straightforward way to handle such moral uncertainty is simply to follow the theory in which you have most credence. This approach is known, disparagingly, as 'My Favourite Theory':¹

¹Gracely 1996, p. 331. The name is due to Lockhart (2000, p. 42). In Gustafsson and Torpman 2014, pp. 167–170, Torpman and I put forward some more complicated versions of My Favourite Theory to handle different kinds of ties between options and between theories. These complications, however, won't matter for the argument of this paper. This is so, because, in the cases we'll consider here, there is no choice node where there is a tie between moral theories (at each choice node, there is a theory that you have more credence in than any other theory) or a tie between options (at each choice node, the theory with the highest credence prescribes a single option). So these complications cannot block the argument of this paper, because those more complicated versions of My Favourite

Pacific Philosophical Quarterly 103 (2022) 448–470 DOI: 10.1111/papq.12408

© 2022 The Authors

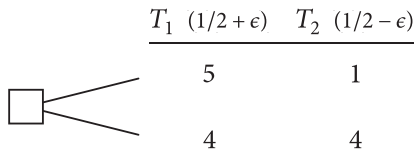
Pacific Philosophical Quarterly published by University of Southern California and John Wiley & Sons Ltd. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

My Favourite Theory An option is a morally conscientious choice in a situation S if and only if that option is permitted by a moral theory such that there is no moral theory in which the agent in S has more credence.

My Favourite Theory captures the way that many people handle moral uncertainty: They identify with the theory they like best (they call themselves Kantians, libertarians, utilitarians, and so on), and they aim to follow the dictates of that theory.

A standard objection to My Favourite Theory is that it's insensitive to the relative sizes of the stakes on different moral theories. Suppose, for instance, that you only have credence in two moral theories, T_1 and T_2 . And you have slightly more credence in T_1 than in T_2 . More precisely, there is some positive constant ϵ less than $1/4$ such that your credence in T_1 is $1/2 + \epsilon$ and your credence in T_2 is $1/2 - \epsilon$. Consider the decision problem in the following diagram, where the table gives the value of each outcome according to each moral theory (with your credence for each theory in parenthesis):²

Case One



The square represents a choice, where you can either go up or go down. Suppose that T_1 and T_2 are maximizing moral theories.³ Going up has a value of 5 according to T_1 and a value of 1 according to T_2 , whereas going down has a value of 4 according to both theories.

The nerve of the different-stakes objection is that, since there is much more at stake on T_2 (the three-unit difference between 4 and 1) than on T_1 (the one-unit difference between 5 and 4), the higher stakes on T_2 should outweigh your slightly higher credence in T_1 . So, according to this line of

Theory prescribe the same options in the cases we will consider as the simpler version stated here. Tarsney (2017, pp. 215–219) puts forward the following variant of My Favourite Theory:

Coarse-Grained My Favourite Theory Aggregate theories as far as content-based comparisons will allow, and when no further content-based comparisons are possible, do what the most plausible comparability class of theories says, in aggregate.

Assuming that all moral theories belong to different comparability classes in the cases we will discuss, the argument of this paper also works against Coarse-Grained My Favourite Theory.

²Hudson 1989, p. 224, Lockhart 2000, p. 84, and Gustafsson and Torpman 2014, p. 160.

³Moreover, we will assume that every moral theory in the examples we will discuss measure moral value on an interval scale. See Roberts 1979, pp. 64–65.

thought, the morally conscientious choice in Case One is to go down. My Favourite Theory, however, prescribes going up, because you have most credence in T_1 and, according to T_1 , going up is better than going down.

The main rival to My Favourite Theory is based on expectations of moral value:⁴

The Principle of Maximizing Expected Moral Value An option is a morally conscientious choice in situation S if and only if that option has at least as great expected moral value in S as any alternative option.

This approach yields a more compelling result in Case One. The expected moral value of going up is $(1/2 + \epsilon) \cdot 5 + (1/2 - \epsilon) \cdot 1 = 3 + 4\epsilon$, and the expected moral value of going down is $(1/2 + \epsilon) \cdot 4 + (1/2 - \epsilon) \cdot 4 = 4$. Since ϵ is less than $1/4$, the Principle of Maximizing Expected Moral Value prescribes going down in Case One.

The different-stakes objection may seem fatal for My Favourite Theory (and to be strong evidence in favour of the Principle of Maximizing Expected Moral Value). But there is a problem with this objection: The claim that the stakes are higher on T_2 than on T_1 relies on intertheoretic comparisons of value differences – that is, comparisons of the value differences of options according to one moral theory with the value differences of options according to another moral theory. Such comparisons seem arbitrary because moral theories typically don't state how their evaluations compare with those of other theories.⁵ So having credence in two or more moral theories does not seem to commit us to any particular exchange rate between the units of moral value in these theories.⁶ If we can't make sense of the claim that the stakes are higher on T_2 than on T_1 , then the different-stakes objection to My Favourite Theory can't get off the ground.

In fact, part of the appeal of My Favourite Theory is that it doesn't rely on any intertheoretic comparisons of value. The reliance on such intertheoretic comparisons is the main drawback of the, otherwise compelling, Principle of Maximizing Expected Moral Value. Without intertheoretic comparisons of value differences, we can't calculate the needed expectations of moral value across moral theories.

⁴A variation of this approach was put forward by Lockhart (2000, p. 82).

⁵Consider, for example, total and average utilitarianism. Equating the difference of one unit of average well-being with the difference of one unit of total well-being is implausible, because it would make average utilitarianism count for almost nothing compared with total utilitarianism for most choices; see Broome 2012, p. 185. And any other exchange rate between the two theories seems arbitrary and still more implausible. For discussions of some alleged solutions to the problem of intertheoretic comparisons of value differences, see Ross 2006, pp. 761–765 and Gustafsson and Torpman 2014, pp. 160–165.

⁶The problem is not whether there's an overarching scale of moral value to which we could convert the moral value from all moral theories. The problem is whether there are any non-arbitrary exchange rates between the theories. Consider an analogy with monetary currencies. The exchange rates between national currencies do not rely on any conversions to some overarching international currency.

A further, *alleged*, advantage of My Favourite Theory is that, unlike many of its rivals, it is immune to value pumps.⁷ In this paper, I will argue that My Favourite Theory is vulnerable to a similar problem: My Favourite Theory prescribes, in some situations, choices that are, sequentially, worse in expected moral value than the opposite choices according to each moral theory you have any credence in. Thus there are situations where one must, in order to follow the approach, follow a plan that has a worse expectation of moral value than some other available plan according to every moral theory in which one has some credence.⁸ What's more, the argument

⁷Gustafsson and Torpman 2014, pp. 160, 172. Of course, if the theory you have most credence in is vulnerable to value pumps, then My Favourite Theory will prescribe behaviour that is vulnerable too. But, in that case, it's the theories you have credence in that are to blame rather than My Favourite Theory. MacAskill *et al.* (2020a, p. 104) question the cogency of value-pumps arguments, pointing approvingly to Ahmed's (2017) self-regulation response to value pumps. Ahmed's approach, however, can be rebutted with a very minimal form of backward induction. See Gustafsson and Rabinowicz 2020, p. 585n13.

⁸The other commonly raised objections to My Favourite Theory aren't very worrying. The most common objection is that My Favourite Theory is sensitive to the individuation of moral theories which seems arbitrary. This challenge can be met by combining the approach with a principle of theory individuation, such as the following (suggested in Gustafsson and Torpman, 2014, p. 171):

The Principle of Fine-Grained Individuation Regard moral theories T and T' as versions of the same moral theory if and only if you are certain that you will never face a situation where T and T' yield different prescriptions.

MacAskill (2014, p. 25; 2016, p. 975n18) proposes a counter-example to this combination. In the example, you are almost certain in prioritarianism and have some credence in utilitarianism, but you are unsure about the shape of the concavity of prioritarianism. So your credence is split between a lot of versions of prioritarianism, each with less credence than utilitarianism, and these versions might make different prescriptions in some future choices. Hence these versions of prioritarianism should be treated as different theories. Suppose that all versions of prioritarianism recommend one option and that utilitarianism recommends another option. The above approach would still recommend that you follow utilitarianism, which may seem counter-intuitive.

But this objection doesn't work. It implicitly relies on either (1) My Favourite Option – that is, the approach of doing what is most likely to be right – or (2) the intuition that these versions of prioritarianism should be regarded as a combined unit. If the objection relies on (1), it runs into the same value-pumps as My Favourite Option; see Gustafsson and Torpman 2014, pp. 165–166. If the objection relies on (2), it suggests that there is a favoured way of individuating theories, because theories that should be regarded as a unit could then, non-arbitrarily, be regarded as versions of the same theory. And then there would be a non-arbitrary principle of individuation, which My Favourite Theory could be combined with instead. It may be objected that the objection could instead rely on the idea that prescriptions under moral uncertainty shouldn't be driven by a moral theory which you have only minuscule credence in. But that idea is just a weaker form of My Favourite Option: It leads to the same kind of cyclic behaviour and vulnerability to value pumps as My Favourite Option – just like supermajority rule leads to the same kind of voting paradoxes as majority rule (although the paradoxes become rarer; see Balasko and Crès, 1997). It may next be objected that the objection could be based on the idea that theories with some arbitrary element (like the concavity in prioritarianism) are biased against, because you are unlikely to have any significant credence in any specific version. This, however, doesn't seem to be a clear drawback of My Favourite Theory, because arbitrariness is a drawback of moral theories. That is: if there is a bias, it may be founded.

Most other objections to My Favourite Theory rely on, seemingly arbitrary, intertheoretic comparisons of value – see, for example, Hudson 1989, p. 224, Lockhart 2000, p. 84, Hedden 2016, p. 106, and MacAskill and Ord 2020, p. 11.

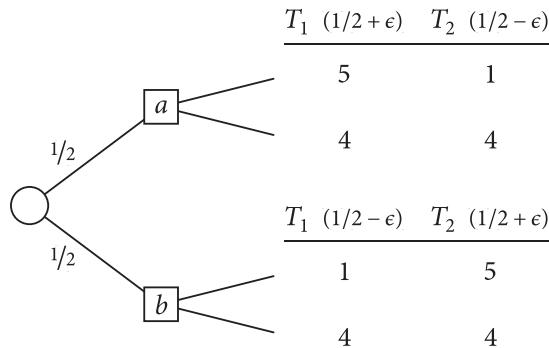
generalizes to other approaches that avoid intertheoretic comparisons of value, such as My Favourite Option, the Borda Rule, and the Principle of Maximizing Expected Normalized Moral Value.

1. The problem of future moral progress

Suppose that you start off with a $1/2$ credence in each of two mutually exclusive moral theories, T_1 and T_2 . That is, you're certain that one of these theories is correct, but you find them equally likely to be so. These are maximizing theories; they prescribe, in each situation, the option that has (according to the theory) the best outcome. Suppose, in addition, that you know that you will soon learn something new that will make one of T_1 and T_2 seem more credible than the other but, currently, you don't know which. You find it equally likely that the news you are about to receive will favour T_1 as that it will favour T_2 . Let ϵ be the size of this foreseen change in your credences and suppose that the shift in your credences between T_1 and T_2 will be less than $1/4$. That is, we suppose that $0 < \epsilon < 1/4$.

Now, consider the dynamic decision problem depicted in the following diagram, where the tables (one for each of the two choice situations) give the value of each outcome according to each of the moral theories (with your credence for each theory in that choice situation in parenthesis):

Case Two



The circle represents a chance node, and the squares represent choice nodes. At the initial chance node, there is, for each of the two choice nodes, a $1/2$ chance that you will face that node. At each of the choice nodes, you have a choice between going up and going down. At node a , you have more credence in T_1 than in T_2 ; more precisely, you have a $1/2 + \epsilon$ credence in T_1 and a $1/2 - \epsilon$ credence in T_2 . And, at node b , you have more credence in T_2 than in T_1 ; more precisely, you have a $1/2 + \epsilon$ credence in T_2 and a $1/2 - \epsilon$ credence in T_1 .

The chance node represents, in part, your uncertainty about which theory will be supported by the news you are about to receive. Suppose that you know, from reliable sources, that a newly published paper has made a breakthrough in the debate regarding T_1 and T_2 and contains a new, compelling argument in favour of one of T_1 and T_2 , but you don't know which.

Depending on what news you get, you will face different choice situations. (This last stipulation may seem strange; we will make do without it in a more complex case in Section 4.⁹) At the initial chance node, your conditional credence in T_1 given that T_1 is supported by the new argument is $1/2 + \epsilon$. And your conditional credence in T_2 given that T_2 is supported by the new argument is also $1/2 + \epsilon$. But, before you learn the news, your unconditional credence in each of T_1 and T_2 is $1/2$.

Let a *plan* at a node n be a specification of what to choose at each choice node that can be reached from n while following the specification. Let us say that one *follows a plan* from a node n if and only if, for each choice node n' that can be reached from n while choosing in accordance with the plan, one would choose in accordance with that plan if one were to face n' . Moreover, let us say that one *intentionally follows a plan* from a node n if and only if one follows the plan from n and, for all nodes n' such that n' can be reached from n by following the plan, if one were to face n' , one would either form or have formed at n' an intention to choose in accordance with the plan at every choice node that can both be reached from n and be reached from n' by following the plan. Finally, let us say that a *plan is available* at a node n if and only if the plan can be intentionally followed from n .¹⁰

In order to follow My Favourite Theory, one must follow *the Up-Up Plan*, that is, the plan of going up at node a and going up at node b . At node a , you have most credence in T_1 and, according to T_1 , going up has a value of 5 whereas going down has a value of 4. Accordingly, My Favourite Theory prescribes going up at node a . And, at node b , you have most credence in T_2 and, according to T_2 , going up has a value of 5 whereas going down has a value of 4. Accordingly, My Favourite Theory also prescribes going up at node b .

Consider, at the initial chance node, the expectation of moral value of the Up-Up Plan conditional on each theory. Let *Learn Up* denote that the chance node resolves upwards, that is, that you get information that favours T_1 . And let *Learn Down* denote that the chance node resolves downwards, that is, that you get information that favours T_2 . The expectation of moral value for the Up-Up Plan conditional on T_1 is $P(\text{Learn Up}|T_1) \cdot 5 + P(\text{Learn Down}|T_1) \cdot 1$. We define the conditional credence in C given A , where $P(A) > 0$, in the usual way:

⁹The assumption involves the kind of normative-descriptive dependence that's discussed in Podgorski 2020, pp. 48–50.

¹⁰Gustafsson 2021, p. 28.

$$P(C|A) =_{\text{df}} \frac{P(A \& C)}{P(A)}.$$

So

$$P(T_1 \& \text{Learn Up}) = P(T_1 | \text{Learn Up}) \cdot P(\text{Learn Up}).$$

Since $P(T_1 | \text{Learn Up}) = 1/2 + \epsilon$ and $P(\text{Learn Up}) = 1/2$, we find that $P(T_1 \& \text{Learn Up}) = (1/2 + \epsilon) \cdot 1/2$. And, since $P(T_1) = 1/2$, we find that

$$P(\text{Learn Up} | T_1) = \frac{P(T_1 \& \text{Learn Up})}{P(T_1)} = \frac{(1/2 + \epsilon) \cdot 1/2}{1/2} = 1/2 + \epsilon.$$

Note that, while the unconditional credence $P(\text{Learn Up})$ is $1/2$, the conditional credence $P(\text{Learn Up} | T_1)$ is, as we have just seen, $1/2 + \epsilon$.

Similarly, since $P(T_1 | \text{Learn Down}) = 1/2 - \epsilon$ and $P(\text{Learn Down}) = 1/2$, we find that $P(T_1 \& \text{Learn Down}) = (1/2 - \epsilon) \cdot 1/2$. So

$$\begin{aligned} P(\text{Learn Down} | T_1) &= \frac{P(T_1 \& \text{Learn Down})}{P(T_1)} = \frac{(1/2 - \epsilon) \cdot 1/2}{1/2} \\ &= 1/2 - \epsilon. \end{aligned}$$

Accordingly, the expectation of moral value for the Up-Up Plan conditional on T_1 is equal to $P(\text{Learn Up} | T_1) \cdot 5 + P(\text{Learn Down} | T_1) \cdot 1 = (1/2 + \epsilon) \cdot 5 + (1/2 - \epsilon) \cdot 1 = 3 + 4\epsilon$. Because of symmetry, the expectation of moral value for the Up-Up Plan conditional on T_2 is the same.¹¹

Compare these theory-conditional expectations of the Up-Up Plan with the theory-conditional expectations of the opposite plan, namely, *the Down-Down Plan* – the plan of going down at node a and going down at node b . The expectation of moral value for the Down-Down Plan conditional on T_1 is $P(\text{Learn Up} | T_1) \cdot 4 + P(\text{Learn Down} | T_1) \cdot 4 = (1/2 + \epsilon) \cdot 4 + (1/2 - \epsilon) \cdot 4 = 4$. Again, because of symmetry, the expectation of moral value for the Up-Up Plan conditional on T_2 is the same.

Hence, from the above calculations, we find that the expectations of moral value at the initial node for each of the theories are the following:

Table One

	T_1	T_2
The Up-Up Plan	3 + 4 ϵ	3 + 4 ϵ
The Down-Down Plan	4	4

Since ϵ is less than $1/4$, we find that the Up-Up Plan has a worse expectation conditional on each moral theory with positive credence than the

¹¹That the theory-conditional expected value is the same given both theories has of course no real meaning because we don't assume intertheoretical comparisons. But, given the symmetry in numbers, the calculations will be the same.

Down-Down Plan. Accordingly, since following My Favourite Theory requires following the Up-Up Plan, following My Favourite Theory requires violating the following principle:¹²

The Weak Principle of Theory-Conditional Plan Dominance If (1) p and p' are plans that are available in situation S and (2), for each moral theory T that the agent in S has some positive credence in, p has a greater expectation of moral value conditional on T than p' in S , then p' is not followed from S .

The problem with violating this principle is that, when you do so, you're certain that there is a plan p such that, no matter which moral theory turns out to be correct, the plan you follow will have a worse expectation of moral value than plan p . An adequate approach to moral uncertainty shouldn't lead to a lower expected moral value according to *every moral theory* in which you have some credence.

Consider, once more, Table One. Your moral uncertainty consists in not knowing whether the moral expectations of the plans are those in the T_1 column or those in the T_2 column. Your moral uncertainty does not extend to how the expectations compare within each column. Your moral uncertainty doesn't prevent you from knowing that, for each moral theory, the moral expectations are greater for the Down-Down Plan than for the Up-Up Plan. You're still certain that the Up-Up Plan has a worse moral expectation than the Down-Down Plan. So an adequate approach to moral uncertainty shouldn't prescribe the choices of the Up-Up Plan. But following My Favourite Theory requires following the Up-Up Plan. So we should reject My Favourite Theory.

Note that this objection to My Favourite Theory does not rely on intertheoretic comparisons of value. The expectations of moral value in Table One do not rely on comparisons of value between moral theories. These expectations are calculated conditional on each theory being correct. So, when we calculate each of these expectations, we assume that one particular moral theory is true. Hence these expectations only rely on *intratheoretic* comparisons of value.

It may be objected that we could avoid these violations of the Weak Principle of Theory-Conditional Plan Dominance if we adopt resolute

¹²This is a weak variant of the following, logically stronger, principle:

The Strong Principle of Theory-Conditional Plan Dominance If (1) p and p' are plans that are available in situation S , (2), for each moral theory T that the agent in S has some positive credence in, p has an at least as great expectation of moral value conditional on T as p' in S , and (3), for some moral theory T' that the agent has some positive credence in, p has a greater expectation of moral value conditional on T' than p' in S , then p' is not followed from S .

While this principle is also plausible, it is stronger than needed for the objection to My Favourite Theory. It may, however, be the closest we can get to the Principle of Maximizing Expected Moral Value without relying on intertheoretic comparisons of value.

choice. Resolute choice is an approach to sequential choices where you adopt a plan at an initial, privileged node and then, at later nodes, you stick to this plan even if it's no longer optimal when it's evaluated at the later nodes.¹³ Suppose that the initial node is the privileged node in Case Two. At that node, you divide your credence between T_1 and T_2 . If we apply My Favourite Theory to the available plans at the initial node, we see that neither of these theories permit the Up-Up plan: T_1 requires the Up-Down plan, and T_2 requires the Down-Up plan. Thus the Up-Up plan is not permitted by any of the moral theories in which you have some credence. So, far from being the plan you must follow in order to follow this resolute version of My Favourite Theory, the Up-Up plan would be prohibited.

This resolute response, however, has two problems. First, a resolute version of My Favourite Theory would require, in Case Two, that you ignore the new moral evidence you receive between the initial chance node and the choice nodes. It doesn't seem morally conscientious to ignore moral evidence. Second, on the resolute approach, there would be one point in time that is privileged in the sense that your plans are always calculated relative to that time. It's hard to see why one point in time would have this privileged status. Why would the expectation of a plan calculated relative to an earlier node have any special significance at a later node?

It may next be objected that it's strange to consider different plans at an initial chance node where you have no immediate choice to make. Note, however, that we could add an earlier choice between the decision tree in Case Two and another decision tree just like it. Then, at the new initial choice node, the plans that involve going up at each choice node after the initial node would still be dominated by the plans that involve going down at those nodes. My Favourite Theory would require that you follow one of the plans that involves going up at those nodes. Hence we still get the same problem.

2. *Conditionalization versus imaging*

It might seem strange that we calculate the theory-conditional expectations at the initial node using a conditional credence for chance going up or for chance going down, rather than the unconditional credence, that is, $1/2$. (If this doesn't seem strange to you, skip ahead to the next section.)

It may be objected that we should use imaging rather than conditionalization. To get the image of a credence distribution P on A , we transfer the credence in each world W where A is false to the world closest to W where

¹³McClennen 1990, p. 13.

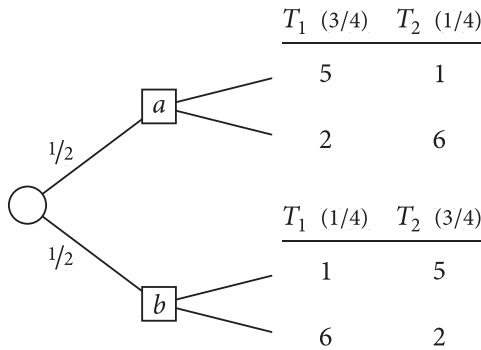
A is true.¹⁴ The crucial thing for our purposes is that your credences in the possible resolutions of the chance nodes in our decision trees are unaffected by imaging on any of the moral theories in which you have some credence. So, for example, your credence in chance going up at the initial chance node of Case Two after imaging on T_1 is $1/2$ (the same as your unconditional credence).

If we use imaging rather than conditionalization, we should replace conditionalization with imaging in the Weak Principle of Theory-Conditional Plan Dominance. If we do so, we get the following principle:

The Weak Principle of Theory-Imaged Plan Dominance If (1) p and p' are plans that are available in situation S and (2), for each moral theory T that the agent in S has some positive credence in, p has a better expectation of moral value after imaging on T than p' in S , then p' is not followed from S .

But this revised principle is implausible. It rules out the Principle of Maximizing Expected Moral Value, which is a plausible approach to moral uncertainty *given that non-arbitrary intertheoretic comparisons of value could be made*. To see this, assume that non-arbitrary intertheoretic comparisons of value can be made and consider the following case:

Case Three



The circle represents a chance node, and the squares represent choice nodes. You start off with a $1/2$ credence in each of T_1 and T_2 . You think that, at the chance node, it's equally likely you will get information that favours T_1 (the node resolves upwards) as that you will get information that favours T_2 (the node resolves downwards). After the chance node, your credence in the theory that is favoured by the evidence rises to $3/4$.

¹⁴Lewis 1976, pp. 310–311.

At node a , the Principle of Maximizing Expected Moral Value prescribes going up, because the expected moral value of going up is $3/4 \cdot 5 + 1/4 \cdot 1 = 4$, whereas the expected moral value of going down is $3/4 \cdot 2 + 1/4 \cdot 6 = 3$. Similarly, the Principle of Maximizing Expected Moral Value also prescribes going up at node b . Hence, following the Principle of Maximizing Expected Moral Value requires following the Up-Up Plan, that is, the plan of going up at node a and up at node b .

The expectation of the Up-Up Plan after imaging on T_1 , or imaging on T_2 , is $1/2 \cdot 5 + 1/2 \cdot 1 = 3$. And the expectation of the Down-Down Plan (that is, the plan of going down at node a and down at node b) after imaging on T_1 , or imaging on T_2 , is $1/2 \cdot 2 + 1/2 \cdot 6 = 4$. Hence the Principle of Maximizing Expected Moral Value violates the Weak Principle of Theory-Imaged Plan Dominance.

Note that the Principle of Maximizing Expected Moral Value does not violate the Weak Principle of Theory-Conditional Plan Dominance. In Case Three, the expectation of the Up-Up Plan conditional on T_1 , or conditional on T_2 , is $3/4 \cdot 5 + 1/4 \cdot 1 = 4$. And the expectation of the Down-Down Plan conditional on T_1 , or conditional on T_2 , is $3/4 \cdot 2 + 1/4 \cdot 6 = 3$. So following the Up-Up Plan doesn't violate the Weak Principle of Theory-Conditional Plan Dominance.

3. *Other approaches without intertheoretic comparisons of value*

Do other approaches in the literature which avoid intertheoretic comparisons of value fare any better than My Favourite Theory in Case Two? And do they satisfy the Weak Principle of Theory-Conditional Plan Dominance? As we shall see, they do not.

First, consider the following approach, which prescribes the option that is most likely to be right.¹⁵

My Favourite Option An option is a morally conscientious choice in situation S if and only if the agent in S has at least as high credence in that option being right as any alternative option.

In Case Two, going up is the option that is most likely to be right at node a , because going up has a $1/2 + \epsilon$ chance of being right whereas going down has a $1/2 - \epsilon$ chance of being right. Likewise, going up is the option that is most likely to be right at node b , because going up has a $1/2 + \epsilon$ chance of being right and going down has a $1/2 - \epsilon$ chance of being right. Hence following My Favourite Option requires following the Up-Up Plan. So, just

¹⁵Lockhart (1992, pp. 35–36) defends a similar principle. The name 'My Favourite Option' is due to Gustafsson and Torpman (2014, p. 165).

like My Favourite Theory, My Favourite Option violates the Weak Principle of Theory-Conditional Plan Dominance.

Next, we will consider two approaches that compare options relative to a certain *Comparative Set*. We can specify this set in a number of ways. We start with a specification based on availability:

The Availability Specification The Comparative Set in a situation is the set of available options in that situation.

Given this specification of the Comparative Set, consider the following approach:¹⁶

The Borda Rule

The *Borda Score* of option *A* in situation *S* according to theory *T* is equal to the number of available options in the Comparative Set which are, according to *T*, worse than *A* minus the number of options in the Comparative Set which are, according to *T*, better than *A*.

The *Credence-Weighted Borda Score* of an option *A* in situation *S* is the sum, for all theories *T*, of the Borda Score of *A* according to *T* multiplied by the credence that the agent in *S* has in *T*.

An option is a morally conscientious choice in a situation *S* if and only if that option has an at least as high Credence-Weighted Borda Score in *S* as any alternative option.

The idea is that each moral theory assigns a Borda Score to each option in the Comparative Set – that is, giving a score of 1 to the worst option, a score of 2 to the second worst, and so on. These scores are multiplied by the agent's credence in the theory, and then these credence weighted scores are added up for each option. Finally, the Borda Rule prescribes the option with the greatest Credence-Weighted Borda Score, that is, the sum total of the credence weighted scores.

To see how this works, consider Case Two.

At node *a*, we find that T_1 gives going up a Borda Score of 1 and going down a Borda Score of -1 , whereas T_2 gives going up a Borda Score of -1 and going down a Borda Score of 1. The Credence-Weighted Borda Score for going up is then $1 \cdot (1/2 + \epsilon) + (-1) \cdot (1/2 - \epsilon) = 2\epsilon$. And the Credence-Weighted Borda Score for going down is $(-1) \cdot (1/2 + \epsilon) + 1 \cdot (1/2 - \epsilon) = -2\epsilon$. So the Borda Rule prescribes going up at node *a*.

At node *b*, we find that T_1 gives going up a Borda Score of -1 and going down a Borda Score of 1, whereas T_2 gives going up a Borda Score of 1 and going down a Borda Score of -1 . The Credence-Weighted Borda

¹⁶MacAskill 2016, p. 989 and MacAskill *et al.* 2020a, p. 73. The Borda Score originates from Borda's (1995, pp. 84–85) voting method.

Score for going up is then $(-1) \cdot (1/2 - \epsilon) + 1 \cdot (1/2 + \epsilon) = 2\epsilon$. And the Credence-Weighted Borda Score for going down is $1 \cdot (1/2 - \epsilon) + (-1) \cdot (1/2 + \epsilon) = -2\epsilon$. So the Borda Rule prescribes going up at node b .

Hence following the Borda Rule requires following the Up-Up Plan in Case Two. So the Borda Rule violates the Weak Principle of Theory-Conditional Plan Dominance.

Next, consider the following normalization approach:¹⁷

The Principle of Maximizing Expected Normalized Moral Value

Normalize the value scales for each moral theory with positive credence so that the best option in the Comparative Set according to each theory is equally good as the best option in the Comparative Set according to the other theories and the worst option in the Comparative Set according to each theory is equally good as the worst option in the Comparative Set according to the other theories.

An option is a morally conscientious choice in a situation S if and only if that option has at least as great expected normalized moral value in S as any alternative option.

Basically, the idea is to first normalize the scales of moral value for each moral theory so that the difference between the best and the worst option is the same on all theories. Then, given this normalization, the Principle of Maximizing Expected Normalized Moral Value prescribes the option with the greatest expected moral value.

Let's see how this works in Case Two, using 1 for the maximum value on each theory after the normalization and 0 for the minimum value.

At node a , we find that T_1 gives going up a normalized value of 1 and going down a normalized value of 0, while T_2 gives going down a normalized value of 1 and going up a normalized value of 0. Then, the expected normalized moral value of going up is $(1/2 + \epsilon) \cdot 1 + (1/2 - \epsilon) \cdot 0 = 1/2 + \epsilon$. Likewise, the expected normalized moral value of going down is $(1/2 + \epsilon) \cdot 0 + (1/2 - \epsilon) \cdot 1 = 1/2 - \epsilon$. Hence the Principle of Maximizing Expected Normalized Moral Value prescribes going up at node a .

Similarly, at node b , we find that T_1 gives going down a normalized value of 1 and going up a normalized value of 0 and that T_2 gives going up a normalized value of 1 and going down a normalized value of 0. Then, the expected normalized moral value of going up is $(1/2 - \epsilon) \cdot 0 + (1/2 + \epsilon) \cdot 1 = 1/2 + \epsilon$. And the expected normalized moral value of going down is $(1/2 - \epsilon) \cdot 1 + (1/2 + \epsilon) \cdot 0 = 1/2 - \epsilon$. Hence the Principle of

¹⁷This approach is a variation of Lockhart's (2000, p. 581) Principle of Equity among Moral Theories. Lockhart's principle also takes care of cases where all options are equally good according to some theories with positive credence but not according to some others. This complication doesn't matter for the argument of this paper.

Maximizing Expected Normalized Moral Value prescribes going up at node *b*.

Thus following the Principle of Maximizing Expected Normalized Moral Value requires following the Up-Up Plan in Case Two. So the Principle of Maximizing Expected Normalized Moral Value also violates the Weak Principle of Theory-Conditional Plan Dominance.

Next, consider the following somewhat more technical proposal:¹⁸

Variance Normalization

Normalize the value scales for each moral theory with positive credence so that the variance (that is, the average of the squared differences in moral value from the mean moral value) of the moral value of the options in the Comparative Set is the same for all theories.

An option is a morally conscientious choice in a situation *S* if and only if that option has at least as great expected normalized moral value in *S* as any alternative option.

The idea is that, if you have the same credence in two moral theories, they should have an equal say – in the sense that their valuations have the same distance (given a Euclidean measure) to a theory that gives the same moral value to all options.

If there are only two options in the Comparative Set, then Variance Normalization is equivalent to the Principle of Maximizing Expected Normalized Moral Value. So following Variance Normalization requires following the Up-Up Plan in Case Two. Hence, like earlier approaches, the Principle of Maximizing Expected Normalized Moral Value violates the Weak Principle of Theory-Conditional Plan Dominance.

At this point, it may be objected that the Borda Rule, Variance Normalization, and the Principle of Maximizing Expected Normalized Moral Value would avoid this problem if they were revised so that they took into account more options than those that are available in the choice situation. One could revise the Borda Rule by replacing the Availability Specification with a specification based on all possible options:¹⁹

The Possibility Specification The Comparative Set in a situation is the set of all possible options in all possible situations.

A problem with this revision is that there seem to be infinitely many possible options. So the revised Borda Score would be undefined for most options, being equal to infinity minus infinity. Hence this revision of the Borda Rule breaks down.

¹⁸MacAskill *et al.* 2020a, p. 86 and 2020b, p. 67. For the concept of variance, see Fisher 1918, p. 62.

¹⁹MacAskill 2014, pp. 123–125.

In the same way, we could revise the Principle of Maximizing Expected Normalized Moral Value so that we normalize the best and worst possible options across all moral theories with positive credence.²⁰ A problem with this revision is that, on some moral theories, there's no upper or lower limit on the moral value of possible options. Consider, for instance, total utilitarianism: For every possible option that realizes a certain sum total of happiness, there is another possible option which realizes an even greater sum total of happiness. Hence there would be no best option among all possible options. And then this revision of the Principle of Maximizing Expected Normalized Moral Value breaks down.²¹

Likewise, for Variance Normalization combined with the Possibility Specification, the variance of moral value will be undefined for some moral theories that lack an upper or lower limit on moral value, such as total utilitarianism.²²

To avoid these problems with infinity, we could let the Comparative Set be not all possible options but all potential options in the decision tree (rather than the available options). That is, the idea is to adopt the following specification of the Comparative Set:

The Resolute Specification The Comparative Set in a situation is the set of all plans that are available from a certain privileged node.

Given that we take the privileged node to be the initial node in Case Two, neither the Borda Rule nor the Principle of Maximizing Expected

²⁰Sepielli 2010, p. 163; 2013, p. 588.

²¹Sepielli 2010, pp. 163–164; 2013, p. 588.

²²MacAskill *et al.* (2020a, pp. 100–101) and MacAskill *et al.* (2020b, pp. 69–70) get around this problem with the help of a probabilistic measure over all possible options. With a probabilistic measure, we may still get, even for unbounded theories, a weighted form of variance over the value of possible options. But, if this probabilistic measure corresponds to your current credences in options being chosen or available, then we are effectively back to the Availability Specification (and the earlier problems). Nevertheless, rather than your current credences over possible options, MacAskill *et al.* (2020b, p. 69) suggest that the relevant probabilistic measure is your fundamental prior credence distribution over possible options. It's implausible, however, that what's morally conscientious for you to do now would depend on what priors you had in the past. (This problem is analogous to the problem we discussed earlier with resolute choice.) And, as Pivato (2022, p. 156) points out, MacAskill *et al.* neither motivate nor justify their probabilistic measure; it's an ad hoc addition they need for their theory to work. It's needed to avoid sensitivity to the individuation of options. If an option is replaced by two more specific versions with the same value as the original according to all moral theories you have any credence in, this may affect the variance; see MacAskill *et al.* (2020a, pp. 96–98). Given that the weight for the original option is the same as the sum of the weights for the versions, the weighted variance stays the same. Yet this suggestion does not solve the analogous problem that, given the Availability Specification, the variance may be affected by the addition of a duplicate of some option in the set – duplicate in the sense that it has the same moral value as the original option according to all moral theories you have any credence in. Your fundamental prior credence in the original option is, of course, lower than the sum of your fundamental prior credences in the original and in the duplicate. What it is morally conscientious for you to do shouldn't be affected by an addition of duplicates, but it may be so given the Availability Specification and Variance Normalization – even given the weighted version.

Normalized Moral Value violates the Weak Principle of Theory-Conditional Plan Dominance, because following them then requires following the Down-Down Plan.

With this revision, however, these approaches both violate the following separability principle:²³

Decision-Tree Separability Whether an option is a morally conscientious choice in a situation does not depend on options that are no longer feasible in that situation.

With the Resolute Specification and the initial node as the privileged node, the Borda Rule and the Principle of Maximizing Expected Normalized Moral Value both violate Decision-Tree Separability because what option is a morally conscientious choice at either choice node in Case Two depends in part on what options are available at the other choice node. It's implausible that options that can no longer be reached at a choice node should matter for what a morally conscientious person would choose at that node. Moreover, there's no plausible reason why the initial node should have any special significance at later choice nodes.

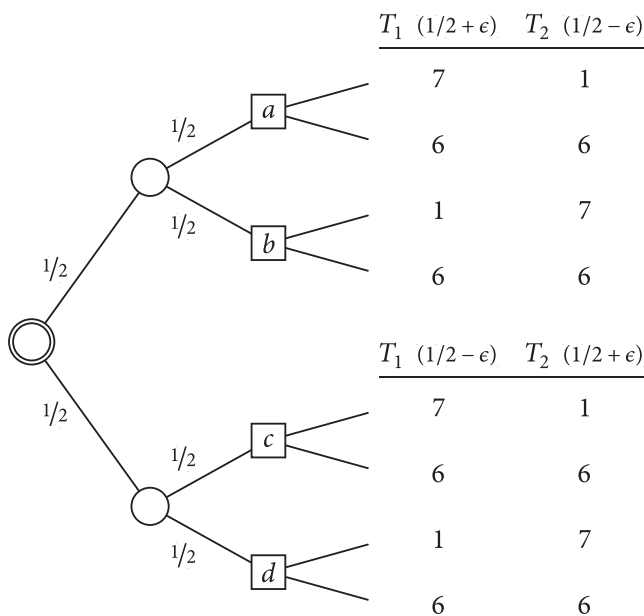
We have seen that My Favourite Theory and its rivals that don't rely on intertheoretic comparisons of value violate either the Weak Principle of Theory-Conditional Plan Dominance or Decision-Tree Separability. Can we do better under moral uncertainty without relying on intertheoretic comparisons of value? We cannot. To see the more general problem, consider Case One. If we don't rely on intertheoretic comparisons of value, all we can say about Case One is that there are two options and two moral theories, which give opposite prescriptions, and one of the moral theories has slightly more credence than the other. Once we describe Case One this way, it's hard to see how a morally conscientious person could do anything else in that case than to follow the slightly more credible theory – that is, to go up. Any approach that satisfies Decision-Tree Separability and prescribes going up in Case One must also prescribe going up at node *a* of Case Two. Then, by symmetry, the approach must also prescribe going up at node *b* of Case Two. But then we find that following that approach requires following the Up-Up Plan in Case Two, which violates the Weak Principle of Theory-Conditional Plan Dominance.

²³McClennen 1990, p. 122. This principle may look less plausible if we were to reject consequentialism. But we can suppose that, in our examples, T_1 and T_2 are different forms of consequentialism. So we can suppose that you are certain that some form of consequentialism is true. And then Decision-Tree Separability should be compelling.

4. Choice-independent stakes

A strange feature of Case Two is that you will face very different stakes depending on how your moral credences will change. Although possible in principle, this may seem unrealistic. This worry, however, can be sidestepped with a, somewhat more complex, variation of Case Two. Suppose, as before, that T_1 and T_2 are two maximizing moral theories. And suppose, again, that you know that you will soon learn something that will make one of T_1 and T_2 seem more credible than the other but currently you don't know which. And let ϵ be the size of this foreseen change in your credences, and suppose that the shift in your credences between T_1 and T_2 will be less than $1/3$. That is, we suppose that $0 < \epsilon < 1/3$. Now, consider the following case:

Case Four



The double circle represents a learning node, the other circles represent chance nodes, and the squares represent choice nodes. The initial learning node models the uncertainty about which theory you will come to have more credence in. A *learning node* differs from a standard chance node in that, at a learning node, no random event occurs – the agent merely learns a piece of information from a set of alternative pieces of information. But the agent has credences in advance about which of the alternative pieces of information they will receive, just like agents have credences about how standard chance nodes will resolve.

As in Case Two, you start with a $1/2$ credence in each of T_1 and T_2 , but you think that, at the learning node, it is equally likely that you will get information that favours T_1 (the learning node resolves upwards) as that you will get information that favours T_2 (the learning node resolves downwards). After the learning node, your credence in the theory that were supported by the evidence rises to $1/2 + \epsilon$. After the learning node, you reach one of the two standard chance nodes. These chance nodes depend on a random event, which you think is equally likely to resolve upwards or downwards.

Note that the learning node just represents your uncertainty about which theory you will get evidence for. The learning node does not depend on a random event. There is a probabilistic dependence between how the learning node resolves and the two moral theories. Suppose, as before, that you know from reliable sources that a newly published paper has made a breakthrough in the debate regarding T_1 and T_2 and contains a new, compelling argument in favour of one of T_1 and T_2 , but you don't know which. The learning node just represents that you regard it as equally likely that the new argument favours T_1 as that it favours T_2 . When the learning node resolves, you learn which theory the new argument supports and you adjust your credence in the two theories accordingly. As in Case Two, your conditional credence in T_1 , at the initial node, given that T_1 is supported by the new argument is $1/2 + \epsilon$. Your conditional credence in T_2 given that T_2 is supported by the new argument is also $1/2 + \epsilon$. And your unconditional credence in each of T_1 and T_2 is still $1/2$.

Following My Favourite Theory, in Case Four, requires following the plan of going up at node a , down at node b , down at node c , and up at node d . Let us call it *the Up-Down-Down-Up Plan*.

Likewise, following My Favourite Option also requires that one follows the Up-Down-Down-Up Plan, and so do the Borda Rule, Variance Normalization, and the Principle of Maximizing Expected Normalized Moral Value given the Availability Specification. Consider the expectation of moral value at the initial learning node of the Up-Down-Down-Up Plan conditional on each theory. The expectation of moral value for the Up-Down-Down-Up Plan conditional on T_1 , or conditional on T_2 , is $(1/2 + \epsilon) \cdot (1/2 \cdot 7 + 1/2 \cdot 6) + (1/2 - \epsilon) \cdot (1/2 \cdot 6 + 1/2 \cdot 1) = 5 + 3\epsilon$. Compare the expectation of the Up-Down-Down-Up Plan with the expectation of one of the alternative plans, namely, *the Down-Down-Down-Down Plan* – that is, the plan of going down at each choice node. The expectation of moral value for the Down-Down-Down-Down Plan conditional on T_1 , or conditional on T_2 , is $(1/2 + \epsilon) \cdot (1/2 \cdot 6 + 1/2 \cdot 6) + (1/2 - \epsilon) \cdot (1/2 \cdot 6 + 1/2 \cdot 6) = 6$. So we find that the expectations of moral value at the initial node for each of the theories must be the following:

Table Two

	T_1	T_2
The Up-Down-Down-Up Plan	$5 + 3\epsilon$	$5 + 3\epsilon$
The Down-Down-Down-Down Plan	6	6

Since ϵ is less than $1/3$, we find that the Up-Down-Down-Up Plan has a worse expectation conditional on each moral theory with positive credence than the Down-Down-Down-Down Plan. Following My Favourite Theory, My Favourite Option, or one of the Borda Rule, Variance Normalization, and the Principle of Maximizing Expected Normalized Moral Value given the Availability Specification requires following the Up-Down-Down-Up Plan. Consequently, these approaches all violate the Weak Principle of Theory-Conditional Plan Dominance. And, as we have seen, we can show this without assuming that the stakes you will face in the future depend on your moral credences.

5. The arbitrariness of the Principle of Maximizing Expected Moral Value

As we have seen, approaches that avoid intertheoretic comparisons of value either violate the Weak Principle of Theory-Conditional Plan Dominance or some other plausible requirement. But, as mentioned earlier, the trouble with the Principle of Maximizing Expected Moral Value is its need for intertheoretic comparisons of value differences. If intertheoretic comparisons of value differences are arbitrary, then the prescriptions of the Principle of Maximizing Expected Moral Value are also arbitrary. But it's hard to see how non-arbitrary intertheoretic comparisons of value differences could be made. In the rest of this section, we will consider two recent proposals for how to make these comparisons.²⁴

It may be objected that the intertheoretic comparisons needed for the Principle of Maximizing Expected Moral Value can be established via a variation of John C. Harsanyi's social-aggregation theorem.²⁵ Let the overall moral value of an option be the agent's overall moral evaluation of the option given their moral uncertainty. Then the idea is that, if the agent's judgements about overall moral value under moral uncertainty satisfy the axioms of Expected Utility Theory and a compelling dominance condition,

²⁴I won't cover the common-ground approach and the reactive-attitude approaches. For a critical discussion of Ross's (2006, pp. 764–765) common-ground approach and Sepielli's (2010, p. 184) reactive-attitude approach, see Gustafsson and Torpman 2014, pp. 162–164 and MacAskill *et al.* 2014, pp. 142–149.

²⁵Riedener 2020, based on Harsanyi 1955. See also Ross 2006, p. 763.

there is a unique expected-utility representation of these value judgements. The dominance condition is, roughly, that options are equally good in terms overall moral value if the options are equal in moral value according to all moral theories with some positive credence, and an option is better in terms of overall moral value than another option if the first option has at least as much moral value as the second option according to all theories with some positive credence and the first option has more moral value according to some theory with some positive credence.

The problem with this argument is that, given the moral theories you have some credence in, there are lots of potential judgements about overall moral value that would satisfy the axioms of Expected Utility Theory and the dominance condition. Different sets of these judgements about overall moral value would have different expected-utility representations, which may prescribe different choices. Since the choice between these different sets of judgements about overall moral value seems arbitrary, the prescriptions of the Principle of Maximizing Expected Moral Value would still be arbitrary.²⁶

It may next be objected that there could be a universal scale for moral value. The idea is that the moral theories we have credence in assign moral value to options on one and the same scale. So these moral theories are all theories about the same absolute moral value quantities.²⁷ And, once we have a universal scale for moral value, we have a way to make non-arbitrary intertheoretic comparisons of value, since different moral theories can then use the same universal scale.

It's far from clear, however, that there is a universal intertheoretic scale for moral value.²⁸ Nevertheless, in the following, I will grant – for the sake of the argument – that there is such a scale. But I will raise a new objection. For this objection, we need to introduce some technical notions. Let $V_T(x)$ be the moral value of option x according to moral theory T . A theory T' has the same *cardinal structure* as theory T'' if and only if, for all possible options x , it holds that $V_{T'}(x) = kV_{T''} + c$, where c and k are constants such that $k > 0$. And, if in addition $k > 1$, T' is an *amplified* variant of T'' .²⁹

Suppose we grant that all moral theories in which we have some credence all grade options in terms of moral value on the same universal scale. Then intertheoretic comparisons of value are non-arbitrary. So far so good. The trouble is that we have traded the old problem of the arbitrariness of

²⁶Another problem, put forward by MacAskill *et al.* (2014, p. 146), is that we would like to know what a morally conscientious person would choose under moral uncertainty. Judgements about the choice-worthiness of options under moral uncertainty is what we would like an approach to moral uncertainty to provide. But, on the Harsanyi-based approach, these judgements are the input to the theory rather than the output.

²⁷MacAskill 2014, pp. 149–157 and MacAskill *et al.* 2020a, pp. 141–147.

²⁸MacAskill 2014, p. 154.

²⁹MacAskill 2014, p. 136.

intertheoretic comparisons of value for a new problem of the arbitrariness of our distribution of credence between theories with same cardinal structure.

To see why, note first that, on the universal-scale view, there could be two versions of utilitarianism, call them Utilitarianism 1 and Utilitarianism 2, such that Utilitarianism 2 is an amplified variant of Utilitarianism 1. The problem is that there seems to be no reason to have any more credence in one of these versions of utilitarianism than the other. The standard arguments for utilitarianism, such as Harsanyi's social-aggregation theorem, don't support utilitarianism with any specific amplitude of moral value.³⁰ So these arguments cannot give us any reason to adopt any specific distribution of credence between Utilitarianism 1, Utilitarianism 2, and other versions with the same cardinal structure. This is a general problem: Plausible arguments for moral theories in moral philosophy do not mention any specific cardinal amplitudes of moral value on some universal scale. And it's hard to see how there could ever be any plausible argument for one version of a moral theory rather than an amplified variant. The upshot is that, if the distribution of credence between moral theories that only differ in amplitude is arbitrary, then the expectations of moral value would still be arbitrary, even if we had non-arbitrary intertheoretic comparisons of value. And, if so, the prescriptions of the Principle of Maximizing Expected Moral Value would be arbitrary.

6. *Arbitrariness or certainty*

Nevertheless, even if intertheoretical comparisons of value are arbitrary, following the Principle of Maximizing Expected Moral Value still guarantees that one satisfies the Weak Principle of Theory-Conditional Plan Dominance and Decision-Tree Separability. Or, more precisely, it does so as long as one relies on the same intertheoretic comparisons throughout. So, if you follow the Principle of Maximizing Expected Moral Value, it's guaranteed that you satisfy the Weak Principle of Theory-Conditional Plan Dominance as long as there is no change in your exchange rates between units of moral value from different moral theories. Even if the intertheoretic comparisons are arbitrary, they do impose a certain structure on our choices when they are combined with the Principle of Maximizing Expected Moral Value. And this imposed structure helps us avoid violations of the Weak Principle of Theory-Conditional Plan Dominance and Decision-Tree Separability. So, even if its recommendations are to some extent arbitrary, the Principle of Maximizing Expected Moral Value might still be our best approach to moral uncertainty.

³⁰Harsanyi 1955.

A remaining possibility, of course, is that rationality requires moral certainty. It may simply be irrational to have positive credence in more than one moral theory. Moral certainty would remove the need for arbitrariness.³¹ Granted, this upshot isn't very helpful if you have moral uncertainty, since it doesn't tell you what moral theory to be certain of. For that, we need to do moral philosophy.³²

Department of Philosophy, Linguistics, Theory of Science
University of Gothenburg, Sweden
Department of Philosophy
University of York, United Kingdom
Institute for Futures Studies, Sweden

REFERENCES

- Ahmed, A. 2017. 'Exploiting Cyclic Preference'. *Mind*, 126(504), 975–1022.
- Balasko, Y. and Crès, H. 1997. 'The Probability of Condorcet Cycles and Super Majority Rules'. *Journal of Economic Theory*, 75(2), 237–270.
- Broome, J. 2012. *Climate Matters: Ethics in a Warming World*. Norton: New York.
- de Borda, M. 1995. 'On Elections by Ballot'. In *Classics of Social Choice*, Eds. McLean, I. and Urken, A.B., University of Michigan Press: Ann Arbor, pp. 83–89.
- Fisher, R.A. 1918. 'The Correlation between Relatives on the Supposition of Mendelian Inheritance'. *Transactions of The Royal Society of Edinburgh*, 52(2), 399–433.
- Gracely, E.J. 1996. 'On the Noncomparability of Judgments Made by Different Ethical Theories'. *Metaphilosophy*, 27(3), 327–332.
- Gustafsson, J.E. 2021. 'The Sequential Dominance Argument for the Independence Axiom of Expected Utility Theory'. *Philosophy and Phenomenological Research*, 103(1), 21–39.
- Gustafsson, J.E. and Rabinowicz, W. 2020. 'A Simpler, More Compelling Money Pump with Foresight'. *The Journal of Philosophy*, 117(10), 578–589.
- Gustafsson, J.E. and Torpman, O. 2014. 'In Defence of My Favourite Theory'. *Pacific Philosophical Quarterly*, 95(2), 159–174.
- Harsanyi, J.C. 1955. 'Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility'. *The Journal of Political Economy*, 63(4), 309–321.
- Hedden, B. 2016. 'Does MITE Make Right? On Decision-Making under Normative Uncertainty'. In *Oxford Studies in Metaethics*, Ed. Shafer-Landau, R., Vol. 11, Oxford University Press: Oxford, pp. 102–128.
- Hudson, J.L. 1989. 'Subjectivization in Ethics'. *American Philosophical Quarterly*, 26(3), 221–229.
- Lewis, D. 1976. 'Probabilities of Conditionals and Conditional Probabilities', *The Philosophical Review*, 85(3), pp. 297–315.

³¹Note that, to avoid these problems with arbitrariness, you don't need to be certain of the *correct* moral theory.

³²I wish to thank Gustav Alexandrie, Gustaf Arrhenius, Gunnar Björnsson, Krister Bykvist, Hilary Greaves, Anders Herlitz, William MacAskill, Erik Malmqvist, Julia Mosquera, Wlodek Rabinowicz, Daniel Ramöller, H. Orri Stefánsson, and Christian Tarsney for valuable comments. Financial support from the Swedish Foundation for Humanities and Social Sciences is gratefully acknowledged.

- Lockhart, T. 1992. 'Professions, Confidentiality, and Moral Uncertainty'. *Professional Ethics*, 1(3–4), 33–52.
- Lockhart, T. 2000. *Moral Uncertainty and Its Consequences*. Oxford University Press: New York.
- MacAskill, W. 2014. 'Normative Uncertainty'. Ph.D. thesis, University of Oxford.
- MacAskill, W. 2016. 'Normative Uncertainty as a Voting Problem'. *Mind*, 125(500), 967–1004.
- MacAskill, W., Bykvist, K. and Ord, T. 2020a. *Moral Uncertainty*. Oxford University Press: Oxford.
- MacAskill, W., Cotton-Barratt, O. and Ord, T. 2020b. 'Statistical Normalization Methods in Interpersonal and Intertheoretic Comparisons'. *The Journal of Philosophy*, 117(2), 61–95.
- MacAskill, W. and Ord, T. 2020. 'Why Maximize Expected Choice-Worthiness?' *Noûs*, 54(2), 327–353.
- McClellenn, E.F. 1990. *Rationality and Dynamic Choice: Foundational Explorations*. Cambridge University Press: Cambridge.
- Pivato, M. 2022. 'Review of William MacAskill, Krister Bykvist, and Toby Ord, *Moral Uncertainty*'. *Economics and Philosophy*, 38(1), 152–169.
- Podgorski, A. 2020. 'Normative Uncertainty and the Dependence Problem'. *Mind*, 129(513), 43–70.
- Riedener, S. 2020. 'An Axiomatic Approach to Axiological Uncertainty'. *Philosophical Studies*, 177(2), 483–504.
- Roberts, F.S. 1979. *Measurement Theory with Applications to Decisionmaking, Utility, and the Social Sciences Reading*. Addison-Wesley: Reading, MA.
- Ross, J. 2006. 'Rejecting Ethical Deflationism'. *Ethics*, 116(4), 742–768.
- Sepielli, A. 2010. 'Along an Imperfectly-Lighted Path'. Ph.D. thesis, Rutgers. <https://doi.org/10.7282/T3B56JWG>
- Sepielli, A. 2013. 'Moral Uncertainty and the Principle of Equity among Moral Theories'. *Philosophy and Phenomenal Research*, 86(3), 580–589.
- Tarsney, C. 2017. 'Rationality and Moral Risk: A Moderate Defense of Hedging'. Ph.D. thesis, University of Maryland. <https://doi.org/10.13016/M2K931684>