# Pick the sugar

Seamus Bradley*

This paper presents a decision problem called the holiday puzzle. The decision problem is one that involves incommensurable goods and sequences of choices. This puzzle points to a tension between three prima facie plausible, but jointly incompatible claims. I present a way out of the trilemma which demonstrates that it is possible for agents to have incomplete preferences and to be dynamically rational. The solution also suggests that the relationship between preference and rational permission is more subtle than standardly assumed.

## The holiday puzzle

You are faced with a choice between two holidays: a scuba diving holiday (Scuba); or a safari holiday (Safari). The virtues of these two holidays you find impossible to compare. The pleasure derived from swimming with myriad colourful reef fish is incomparable with the pleasure of witnessing a majestic pride of lions in the morning sun. They are pleasures of very different sorts, and you don't judge either better than the other, nor do you consider them equally good.

Your preferences are such that you prefer more money to less, other things being equal, so you prefer a scuba diving holiday plus $50 spending money (Scuba+Money) to the original scuba holiday (with no money). Call this option the "sweetened" option. Likewise, you prefer Safari+Money to Safari. However, the $50 is not sufficient to break you out of the incommensurability: you find Scuba+Money to be incommensurable with Safari (with no money). And likewise for Safari+Money and Scuba.

This means that in a choice between Scuba and Safari, choosing either option is permissible. In a choice between Scuba, Safari, Scuba+Money and Safari+Money, either of Scuba+Money or Safari+Money is permissible. The other options are not permissible because there is something that is preferred to them. That is, Safari is not permissible because Safari+Money is preferred to Safari. In short, don't choose $\phi$ if you could choose something better, and if you can't choose something (determinately) better, then don't rule it out. If all

*seamusbradley.net

four options are on the table like this (as depicted in figure 1) then only the two options *with* the money are permissible.

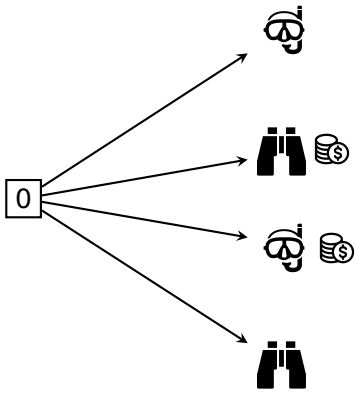


Figure 1: The "flat" holiday puzzle

Now let's imagine a different decision problem, due to Hedden (2015a). Call it "the holiday puzzle". On Monday, you will be confronted with two boxes; box A contains the safari holiday tickets and box B the scuba holiday tickets (and you know this). You will have the choice to put $50 in one of the two boxes. On Tuesday, you will choose which box to take (and you know all this on Monday, and on Tuesday you will remember what you did on Monday). This decision problem is depicted in figure 2. How ought you choose? On Monday, it seems like there's no reason to prefer putting the money in box A over box B or vice versa, so arguably on Monday either option is permissible.

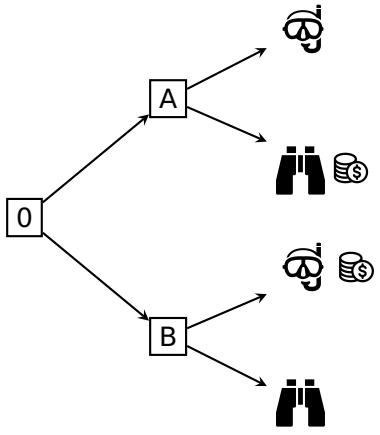


Figure 2: The holiday puzzle

However, on Monday it seems like you ought not put the money in one box and

then take the *other* box come Tuesday: that would be giving up on free money! Let's imagine that you put the money in box A. You come back on Tuesday, remembering that you put the money in box A. But consider the choice that confronts you on Tuesday:

- Box A: Safari+Money
- Box B: Scuba

By hypothesis, you have no preference between these options: the $50 didn't resolve the impasse. Given this, it seems like selecting either option should be permissible, in particular, selecting box B is permissible. This choice goes against what you thought on Monday: that taking the box you didn't put the money in would be irrational. And indeed, in this sequence of actions – money in box A, take box B – you end up with strictly less than you could have had, had you put the money in box B on Monday!

So there's two related problems with taking box B on Tuesday. First, putting the money in one box and then taking the other doesn't look like the sort of thing a rational agent should be permitted to do: it's just intuitively wrong to forgo free money. Second, it is "dynamically incoherent", in that you end up doing something on Tuesday you wouldn't endorse from the perspective of your Monday preferences. This disconnection between what choices you endorse your future self taking and the choices your future self actually makes is a distinctively diachronic form of incoherence. Put a little more carefully, the problem is that on Monday you prefer that you act a certain way on Tuesday (namely, taking whichever box Monday-you put the money in); but on Tuesday, you no longer have that preference

This is, in essence, the problem I seek to solve in this paper. Hedden (2015a) introduced this version of the puzzle, although related puzzles have a venerable history, as we'll see later. First we're going to diagnose the problem and demonstrate that problems structurally analogous to the holiday puzzle affect a number of other areas. Then, I will draw out what desiderata we might want a solution to the puzzle to have by exploring the debate surrounding one popular escape route for puzzles like this. I will then turn to presenting my own solution which builds on and elaborates a solution suggested by Ruth Chang.

## Diagnosing the holiday puzzle

The holiday puzzle is not just a problem for those struggling to choose between exciting holiday options. Structurally similar problems arise in a number of circumstances, as we will now see. The goal of this section is to pick out those features that cause the problem. Note that preference relations have the property that they confer a deontic status on their relata: the options that are at the top of the preference ranking are those it's permissible to choose, and those

lower down the ranking are impermissible as choices.[1] Other relations have this property too. For example, Broome (1991) focuses on a relation of objective betterness, and that which is at the top of that ranking is The Good. Or consider the widespread view that you ought to do what you have most reason to do (Portmore 2019): those options you have most reason to do are the ones that are top of the "more reason to $\phi$ than $\psi$" ranking. In all of these cases – preference, betterness, more reason – some kind of normative or deontic status – permissible choice, the good, what you ought to do – is conferred on an option according to its place in the relation. I shall call these kinds of relations "relations that confer deontic status". In what follows, I will mainly talk in terms of preference, but similar arguments should go through for other relations conferring deontic status.

Our earlier analysis of how to choose among the options in the holiday puzzle relies on something like the following principle linking your preferences to your rationally permitted choices.

**Maximality is sufficient for permission** If there is no option that is strictly preferred to option $\phi$, then $\phi$ is permissible.

This principle seems to be fairly commonly accepted, call it **MSP**. We will discuss this principle at length in a later section.

Note also that a preference relation can be incomplete. A relation such as a preference relation can hold between two options $\phi$ and $\psi$ in three ways: $\psi$ is strictly preferred to $\phi$, $\phi$ is strictly preferred to $\psi$, or $\phi$ and $\psi$ are considered indifferent. If one (and only one) of these conditions is satisfied for every pair of options $\phi$ and $\psi$ then the relation is said to satisfy "trichotomy", "completeness", "connectedness", "totality" (all these terms can mean the same).

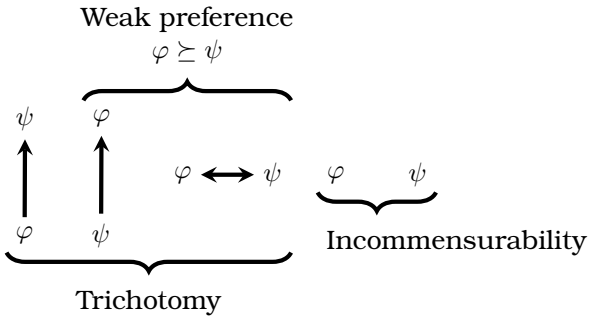


Figure 3: Trichotomy

If this property does not hold, that is, if there is a pair of options $\phi$ and $\psi$ such that none of the three condtions of trichotomy is satisfied, then the relation is

[1] We'll see later that we have to be a little more careful about this connection, but this will do as a first approximation.

incomplete. Call any such pair of options "incommensurable."[2]

There are a number of reasons why a relation of preference might be incomplete. It might be that the goods in question are just good in different ways that cannot be perfectly compared; goods might be found incommensurable because of imperfect introspection or limited access to your own preferences; the goods in question might somehow be indeterminate and such indeterminacy can affect your attitudes toward them (Williams 2014, 2017); or the goods might be bets on particular states of the world obtaining where you have imprecise credence about those states (Bradley 2019). This list is not exhaustive (the authors in Chang (1997) canvas a number of options). Other relations conferring deontic status, it seems to me, can also be incomplete for at least some of these reasons. For example, Broome grants that the objective betterness relation could be incomplete (Broome 1991, 137), although he doesn't discuss the possibility further.

If there are two such incommensurable goods $\phi$ and $\psi$, and the space of possibilities is suitably rich, there is a third good $\phi^+$ such that your preferences display the feature of being "insensitive to sweeteners". By this I mean that the three goods, $\phi$, $\phi^+$ and $\psi$ are such that $\phi^+$ is strictly preferred to $\phi$, but no other relation of preference holds between the goods.[3] Safari+Money is an example of a sweetened option. Insensitivity to sweeteners is a distinctive property of incommensurable goods.

So we have a relation that confers deontic status that is insensitive to sweeteners, and we have a principle linking the relata of that relation to deontic status. With these components, we can construct a decision problem analogous to the holiday puzzle by offering you a first choice of whether to sweeten $\phi$ or $\psi$, and then offer you a choice between your chosen sweetened option and the unsweetened version of the other good. And by hypothesis, you will have no preference for the sweetened option once you arrive at the second choice point (see Figure 4). So any instance of incompleteness of a relation that confers deontic status can be used to construct such a decision problem.

This leads us to a trilemma.

- A relation conferring deontic status can be incomplete
- Maximality is sufficient for permission
- Selecting the unsweetened option at the second choice point is impermissible

[2]Before continuing, a word on terminology. Chang (2002) makes a distinction between parity and incommensurablity. Parity is some sort of judgment that despite the lack of preference between the options they are somehow "on a par"; this is to be contrasted with genuine incommensurability which is an inability or unwillingness to make any comparison between the options. This is a distinction I won't make here and will talk about incommensurability throughout. The important structural feature of the cases I explore here is that they are cases appropriately modelled by a rational agent's having an incomplete relation represent their preference.

[3]Symmetrically, there is also some $\psi^+$ strictly preferred to $\psi$ that stands in no preference relation to $\phi$.

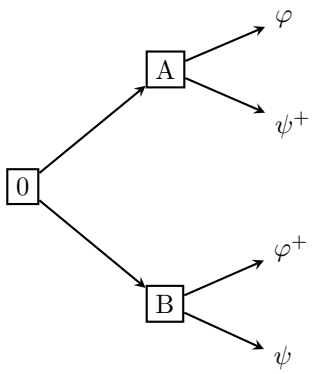


Figure 4: The generalised holiday puzzle

Or, in terms of the original holiday puzzle:

- Preference can be incomplete
- Maximality is sufficient for permission
- Selecting box B on Tuesday is impermissible

Each of these premises seems attractive. The problem, as discussed above, is that the first two of these seem to entail something incompatible with the third.

## Solving the holiday puzzle

We have three incompatible claims, and so (at least) one of them must be denied. Apart from Hedden's introduction of the puzzle (Hedden 2015a, 2015b), there hasn't been much explicit discussion of the holiday puzzle,[4] but many discussions of rational choice can be interpreted as suggesting which claim of the trilemma to deny. In what follows, I will take the liberty of writing as if the various authors I discuss were directly responding to the holiday puzzle, whereas many were actually responding to other similar puzzles.

Hedden himself seems to favour denying that selecting box B is irrational (as does Doody (2019a)) and one might also interpret Teddy Seidenfeld (1988, 2004) as going this way (but for different reasons). Alternatively we might side with Peterson (2007) and Elga (2010) in denying that preferences can be incomplete. These are not routes I will explore further in this paper. I think preference can be incomplete, and I think it is intuitively irrational to take the unsweetened option in the holiday puzzle, and so I think it is worth searching for a theory of decision that accommodates these features, and thus, we must look for a way to deny **MSP**.[5]

[4] Doody (2019a) is one paper that does so.

[5] One potential decision theory in the area is proposed by Caspar Hare (2010, 2013). I won't discuss this theory here because in Hare's example, you don't know which holiday is in

In the remainder of this section I explore a variety of methods for denying that Maximality is Sufficient for Permission by offering a theory of decision-making that doesn't make it true. The methods I explore can be broadly described as "resolute" decision methods. I discuss several criticisms of such resolute methods. My aim in doing so is to understand what desiderata critics of resolute choice are appealing to in order to reject it. In later sections, I will then show that my alternative decision theory satisfies these desiderata (suitably qualified) and that it is thus both distinct from extant resolute methods, and acceptable even to critics of resolute methods.

The basic idea of the resolute response is this: on Monday, as well as putting the sweetener in box A, you make a plan to take box A on Tuesday. Following through on your plan on Tuesday then guarantees that you choose in a dynamically coherent fashion. Depending on how we interpret this idea, we might understand this approach as denying that your preference can be incomplete *in the holiday puzzle on Tuesday* (while granting that preference can be incomplete in general). Alternatively, it could be interpreted as denying that maximality is sufficient for permission. This planning approach to choice has been discussed at length by, for example, Machina (1989); Bratman (1992); (1999); McClennen (1990). I'm going to lump together a number of views that one might want to distinguish from each other, because I think the same collection of criticisms apply to them all.

Here's one reading of how planning works. The plan you make on Monday somehow directly affects what is permissible on Tuesday (in a way that is not mediated by your Tuesday preferences).[6] So this solution denies that an option's being maximal entails its being rationally permissible: an option could be maximal, but ruled out by a plan you made.

Elga (2010) argues against this planning view as follows: it's problematic that the rational requirements on an agent are different depending on whether Tuesday's choice is part of a sequential decision problem or not, even if the decision facing the agent is the same. According to the advocate of planning, if you are simply offered the choice between Scuba and Safari+Money, you can choose however you like, but if you arrive at that choice point having put the money in box A (the safari box) and planned to then take that box, the rational constraints on choice are different, even though the actual decision now facing you is the same. Elga suggests that such a difference in rational requirement that is not grounded in a difference in your current (i.e. Tuesday) preferences is an unwelcome feature of a decision theory. He seems to be appealing to a premise like the following:

**Separability** Your past decisions are irrelevant to rational choice

which box, and the sweetener is allocated at random, rather than as a consequence of a choice, and so it's unclear to me whether his approach says anything about the holiday puzzle per se. For further discussion of Hare's example, see Bales, Cohen, and Handfield (2014); Schoenfield (2014); Doody (2019b).

[6]This caveat serves to emphasise the way that this view differs from an alternative we shall meet shortly. I read McClennen (1990) as seeing resolute choice this way.

In other words, how you face a decision problem at a choice node should be the same as if you were just facing the sub-tree that starts at that node. The name "separability" is used by McClennen (1990) although he doesn't endorse the principle; Buchak (2013) uses the term "Only Future" for the same concept. This rules out your earlier plans having an influence on what's permissible.

One might think that Elga's argument here begs the question, since one might deny that Scuba is the same option as Scuba-having-put-the-money-in-box-A. That is, if we individuate options finely enough that past choice affects what options are available, then Elga's argument doesn't get going. I don't think this counterargument holds much water, since such finely individuated options are different only in terms of their past histories and not in terms of any facts about their future outcomes. Thus, if we take Separability to be true, there's no justification for having your preference treat Scuba and Scuba-having-put-the-money-in-box-A differently. In short, such fine individuation is ruled out by Broome's "Individuation by Justifiers" principle (Broome 1991, 103).

It's worth stopping here for a minute to discuss motivations for Separability. One motivation for such a view is a modest kind of "Internalism". For example Hedden (2015b) argues for "Time-Slice Rationality" which is the view that all principles of rationality are synchronic. This obviously entails Separability. And he argues for this view on the basis that what it's rational to do should supervene on your (current) mental state. So, in particular, your past plans don't have an impact on how you ought to choose unless they have an impact on your current mental state.

Another strand of criticism against the planning view is encapsulated in this quote:

> What would it mean for an agent to choose against her preferences in order to fulfill a previously-selected plan? That would seem to defy the very notion of preference. Of course, an agent may place considerable importance on honouring previous commitments. Any such integrity concerns, however, should arguably be reflected in the specification of outcomes and thus in the agent's preferences at the time in question. (Steele and Stefánsson 2020)

The criticism is that what it's rational to choose should be determined by your preferences, not by your plans.

**Preference Determines Rational Choice** What it is rational to choose supervenes on your preferences

This rules out the possibility that your plan (which is not a preference) can influence your decision in a way that is not mediated through a change of preference. This again can be motivated by a kind of internalism. It is specifically the "preferences" part of your mental state that determines what's rational. That is not to say that preferences are basic: you could have the view that your preferences are determined by your beliefs and desires, but your rational choices

are still mediated by your derived preferences. So a plan that does not somehow affect how you prefer cannot influence what it is rational to choose.

Another argument against the view that planning can solve the holiday puzzle is the problem of unanticipated trades. Peterson (2007) criticises the suggestion that planning might be a way out by suggesting that you might make the Monday decision without anticipating the future decisions (and thus without having made a plan to act appropriately). Taking box B is irrational regardless of whether you knew you would confront that (Tuesday) decision when you put the money in box A on Monday, according to Peterson. I don't put too much weight on this line of argument, though, since it isn't clear to me when we should require agents to be dynamically coherent if they're ignorant of what options might later be available to them. Nevertheless, we can spell out this desideratum as follows:

**Accommodate Unanticipated Decisions** An agent should be dynamically coherent regardless of whether all future decisions are anticipated at earlier choice points

An alternative interpretation of the basic planning idea is that the plan you make on Monday influences what is permissible on Tuesday *by influencing your Tuesday preferences.*[7] This reading of the planning approach denies that your preferences are incomplete on Tuesday (since the plan makes you prefer the sweetened option). While this view is compatible with Preference Determines Rational Choice, it still falls foul of the other two criticisms we discussed. One response to this line of argument would be to defend the view that $\phi$ and $\phi$-having-planned-to-$\phi$ are distinct options (and that it is legitimate to justify a difference in preference on the basis of having made a plan to act a certain way). This amounts to giving an argument as to how plans can have (or ought to have) some motivational force or why you ought to follow through on your plans. I will return to this point later.

A related collection of views is what we might term "mind-making" and "identification" views (Williams 2014; Moss 2015b, 2015a). To understand these views, we need a short tangent into the structure of preferences. Often your preferences are not captured simply by an incomplete relation, but there is some set of (complete) relations that collectively capture your attitudes to options, call them the set of admissible completions. For example, if the preference relation is determined by a utility function and a set of probability functions,[8] then, each probability function in the credal set determines a complete order of the gambles (ordering them by their expectation). This set of relations collectively represent your preference. If we consider the relation "$\phi$ is preferred to $\psi$ according to all probabilities in the credal set", then this yields a transitive but possibly incomplete relation. This leaves open some questions about how you ought to choose (Bradley 2015), but at the very least if $\phi$ is strictly preferred to $\psi$ on each of those orderings, then it should not be the case that the aggregate preference

[7] I interpret Buchak (2013) as understanding planning in this way.

[8] Or, more generally, a set of probability, utility pairs.

ranks $\psi$ above $\phi$. Or consider incompleteness of preference due to indeterminacy. For every sharpening or precisification of the indeterminate facts, there is a complete ordering over that sharpened set of acts. And the "super-preference"[9] – the preference analogue of supertruth – is incomplete.

Back to mind making. Moss proposes that decision-making involves "identifying with" one of the admissible completions of your preference and selecting something optimal by the lights of that completion.[10] Since the rational completions of your preference relation are dynamically coherent,[11] making further decisions on the basis of this same privileged completed relation guarantees that you take the sweetened option in the holiday puzzle. Ruth Chang's "hybrid voluntarism" view (about which more later) could also be seen as a kind of identification view (Chang 2013). Williams (2014) focuses on the case of decision-making under indeterminacy. Among the views that Williams canvasses (and the one he seems to prefer) is the view that making a decision affects your epistemic state: choosing one way makes you confident that having chosen that way was maximal. Choosing $\phi$ means removing all admissible completions that don't make $\phi$ optimal. That is, putting the money in box A makes you prefer Safari+Money to Scuba. This is not a plan, but some other kind of influence on your mental state. Mind-making and identification views differ on whether it is the choice that causes the change in mental state (mind-making) or it is the change in mental state that determines the choice (identification). With that caveat in mind, I think we can treat them together.

Identification and mind-making views don't seem to allow it to be rational to select something determinately second best, something which we might at least want to leave open as a possibility (Bales 2018). In order to be permissible on these views, an option has to be optimal according to at least one admissible completion of your preference relation. There are, however, cases where an option is not optimal on any sharpening, but still somehow a good compromise option. Such determinately second best options are never permissible on mind-making views of rational decision. Moss criticises Williams for not accommodating "hedging", but the "hedging" that Moss' theory allows is making obligatory an option that is optimal on some sharpenings but not others. The view she advocates does not permit making permissible an option that is second best on all sharpenings.[12] Moss (2015b) discusses a case (pp.20–4) of a determinately second best gamble, but her solution involves requiring the agent be risk-averse, which seems to be a red herring. An agent could be risk-neutral but ambiguity averse enough to select determinately second best options: Moss' view can't

[9] $\phi$ is super-preferred to $\psi$ iff $\phi$ is preferred to $\psi$ on every sharpening. Note that some subtlety is required when it comes to "weak" super-preference and accommodating indifference, but these difficulties needn't detain us here.

[10] Moss is writing in the context of incommensurability due to imprecise credences, I don't think it is a stretch to generalise Moss' view to other cases of incomplete preference.

[11] If the admissible completions of your preference aren't dynamically coherent, then *that's* the problem with your preferences, not the incompleteness.

[12] In the context of imprecise credence, permitting determinately second best options is what distinguishes the Non-domination decision rule from E-admissibility.

accommodate this pattern of attitudes. So another desirable feature of a decision theory is the following:

**Allow Determinately Second Best** An option that is not optimal, but "near optimal" on all admissible completions can be permissible

It might appear that such "mind-making" views are incompatible with Separability, but, properly understood, they are not. According to the mind-making view, an intrinsic duplicate of your Tuesday self (who had not made the Monday choice) would choose as you do, since their mind would also be made up. In a sense, I think this highlights another undesirable feature of such views: they require that choices you make have an undue influence on your mental state. Of course, everyone thinks that making a choice has *some* influence on your mental state: you come to believe that that was the choice you made, for example. But the kind of change of mental state required to have mind-making suceed in the holiday puzzle seems too much. So, other things being equal, a theory of decision that solves the holiday puzzle while making more minimal commitments as regards how your attitudes change in response to decision is preferable.

**Minimal Attitude Change** Other things being equal, a decision theory is better, the more minimal are the changes in attitude entailed by choosing a particular way

I think I have done enough to cast some doubt on extant planning and mind-making theories of decision. I don't need my reader to find all of the above criticisms compelling. In a sense, my theory is the "minimal" theory that solves the holiday puzzle, and it is compatible with any more full-bodied version of planning or mind-making.

## Rational silence

My approach to solving the holiday puzzle takes its inspiration from the work of Ruth Chang. Chang is talking about a different decision problem, but we can, I think, paraphrase her as follows:

> The rational permissibility of choosing either of two items on a par, then, must be constrained by one's other choices. If one [puts the money in box A on Monday], one is thereby rationally prohibited from choosing [box B on Tuesday]. This is true even though there is a sense in which because [Scuba and Safari+Money] are on a par, it is rationally permissible to choose either. This is the sense in which if one had not already chosen [to put the money in box A on Monday], it would have been rationally permissible to choose [Scuba]. (Chang 2005, 347)

She continues by saying that the sense in which it's irrational to choose box B on Tuesday "depends on understanding the rationality of choice against a background of other choices" (p.347). Chang points to, but doesn't resolve a

tension between two ways of understanding permission here. On the one hand, there's a sense in which it's permissible to choose either of the incommensurable goods. On the other hand, there's a different sense in which it is *not* permissible to do so, given the background of other choices. In the remainder of this paper, we draw out these two senses, and reconcile them into a cohesive and plausible theory of the relationship between preference and rational permission.

I propose that we solve the holiday puzzle by denying **MSP**; by denying that an option's being maximal suffices for rational permission to select it. But there was some intuitive appeal to **MSP** so, why might one endorse such a principle? Peterson (2007) defends a principle similar to **MSP** but for a relation of "betterness" rather than preference. He takes the betterness relation to capture everything there is to say about value, and defends an analogue of **MSP** by arguing that value should be choice-guiding. Translated into the terms we are using here, I think Peterson would want to say that Preference Determines Rational Choice and that **MSP** captures an important aspect of *how* preference does that job. I broadly agree that Preference Determines Rational Choice, but I don't think **MSP** is part of explaining how that works. So we need to explore in more detail how we might link preference and rational choice.

The strategy for the remainder of this section is to show that the basic commitments that apparently made **MSP** plausible only warrant a weaker principle, one that is not inconsistent with the other two premises of the trilemma. This lays the groundwork for a plausible satisfying theory of rational decision with incommensurable goods. In order to make progress, we shall need a number of definitions. First, let's outline several properties that an option $\phi$ might have with respect to some relation (see Figure 5):

**Strict Optimality** $\phi$ is strictly preferred to every other option available
**Optimality** $\phi$ is weakly preferred to every other option available
**Maximality** There is no available option $\psi$ such that $\psi$ is strictly preferred to $\phi$

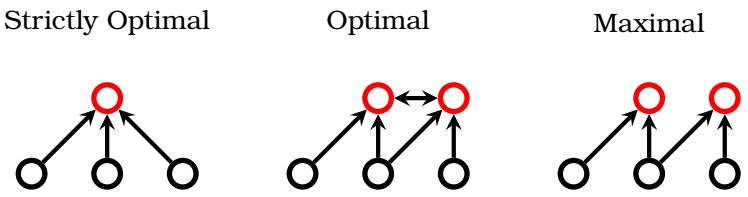


Figure 5: Optimality and Maximality

So long as your preference relation is transitive, there are some tight connections between these properties (Suzumura 1983). Strict optimality entails optimality, and optimality entails maximality, but neither entailment holds the other way, unless the relation is complete; in this case, optimality and maximality are equivalent. If there is an option that is optimal, then all other maximal options

are also optimal For any collection of available options, there is always at least one maximal option.

If we're operating in a context where the incomplete relation is something like "super-preference", determined by the "intersection" of a number of "admissible completions", then there is a further property that falls between Optimality and Maximality in terms of strength. This is the property of being optimal with respect to some admissible completion (call this "optimal on a completion"). This is obviously strictly weaker than Optimality, but it is also strictly stronger than Maximality.

Given these properties, we can explore some relationships between preference and rational permission. As a first pass, consider:

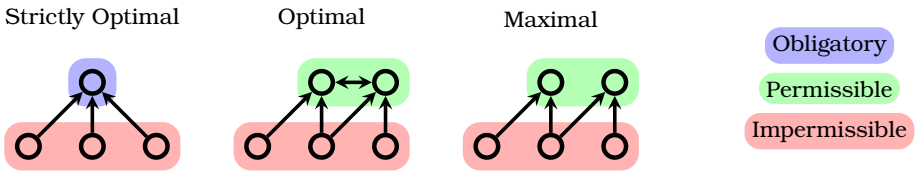


Figure 6: Preference–Choice links

**Maximality is equivalent to permission** $\phi$ is maximal if and only if $\phi$ is rationally permissible

This principle has, as logical consequences, the following principles:[13]

**Strict optimality is sufficient for obligation** If $\phi$ is strictly optimal then $\phi$ is obligatory

**Optimality is sufficient for permission** If $\phi$ is optimal then $\phi$ is permissible

**Maximality is necessary for permission** If $\phi$ is not maximal then $\phi$ is impermissible

**Maximality is sufficient for permission** If $\phi$ is maximal then $\phi$ is permissible

Each of these principles tells us something about the relationship between preference and rational permission to choose. Each does work in cashing out what we mean by Preference Determines Rational Choice. I suggest that the only one of these principles that we actually need to deny is the last. That is, for optimal options, things are as they were before. The unorthodox or novel aspect of this approach to decision only appears when the preference relation is incomplete.

In the context of admissible completions, we have the further principles that "optimality on a completion is sufficient for permission" and "optimality on a completion is necessary for permission". The conjunction of these principles

[13] Assuming that the relation is transitive and that if an option is the only permissible choice then it is obligatory.

is what Levi calls "E-admissibility" in the context of imprecise credence. If we want to Allow Determinately Second Best, then we will have to deny that "Optimality on a Completion is Necessary for Permission". We'll see later that taking the unsweetened option can be optimal on a completion, so replacing **MSP** with "Optimality on a Completion is Sufficient for Permission" will not solve the trilemma. So we will also need to deny this.

This still leaves open some questions about rational permission. If an option is maximal but not optimal, is it rationally permissible to choose it? The preference-permission connections we have endorsed so far are simply silent on this question. I take up how to resolve this question in the next section.

Summing up, **MSP** is false, but **MSP** followed from a more general principle, and most of the other consequences of that general principle can still hold. The basic idea that your preferences determine what it is rationally permissible to choose is *almost* right. Most (but perhaps not all) facts about rational permission are determined by your preferences, and determined in a way compatible with Separability and our other desiderata.

In a sense, we have resolved the tension in the trilemma. We have made room for the possibility that there might be some way to favour one maximal option over another. Maximality is not sufficient for permission, but the other two premises of the trilemma are true. Mission accomplished? Not really. We would like to say something positive about how non-optimal maximal options acquire deontic status, and ideally we would like to have some procedure for conferring those statuses that makes selecting the sweetened option the favoured option in the holiday puzzle. This is our task for the next sections.

# Picking

So an option's being maximal is not sufficient for it to be permissible: this leaves open the question of how to go about selecting among the maximal options. Before we tackle the difficult question of rational permissibility of selecting among (merely) maximal options, let's address the easier question of rational permissibility for optimal options. Consider Buridan's ass, stuck between two perfectly similar bales of hay, the ass has no reason to choose one bale rather than the other. Selecting either bale of hay is optimal, but since the ass is indifferent between them, it is not *obliged* to select either one. The right approach is not to die of starvation forever trapped with an impossible choice, but rather to *pick* one of the bales of hay. The ass needn't have a *reason* to prefer one to the other, but picking one is better than starving with the decision unmade. This important distinction – between picking and choosing – is due to Ullmann-Margalit and Morgenbesser (1977). We often have to select among items we have no reason to choose between, that selection can't be a choice (since choice implies some reason to choose one way rather than the other) and yet, there are considerations that favour picking one item over the other. Selecting *this* soup can rather than *that*

one because doing so makes the display of cans symmetrical hardly counts as a reason to choose, and yet it seems a not unreasonable consideration to determine your pick given that some kind of selection is required.

To make a selection between several optimal options, you need to employ some form of procedure for picking one over the others. This procedure is typically understood to be arational, that is, without rationality. That is, (because optimality is sufficient for permission) any selection of an optimal option would be rationally permissible, but some procedure is still needed to make the selection.

Let's turn now to cases of maximal options. It is not the case that any maximal option is permissible: the deontic status of maximal options is up for grabs, it is not decided by the principles governing choice. Some form of procedure for picking is also required in these cases. I propose that one important difference between the case of picking among optimal options and picking among maximal options is that, for the case of non-optimal maximal options, some procedures for picking are more rational than others. That is, picking among maximal options is not arational, but a substantive element of rationality remains to be uncovered in determining reasonable procedures for picking. In the holiday puzzle case, for example, picking the sweetened option on Tuesday is rational whereas picking the unsweetened option is not. Why not? Because your rational picking procedure favours the sweetened option.

There remains a question as to exactly what deontic status a picking procedure confers on the option picked. If your picking procedure picks out $\phi$, what deontic status does $\phi$ have? Is $\phi$ obligatory? Or is there some form of graded concept of rationality where the picked option is "more rational", or "recommended"?[14] Perhaps picking confers a kind of "derivative obligation" status on the option, in the same way that you're not obliged to charge into the burning building, but if you do, you acquire the derived obligation to help the people trapped inside (Kagan 1991, 16). For the purposes of this paper, I will just talk in terms of some generic status of "favoured option" that is conferred on a picked option, where this can be read as a synonym for whichever deontic status you prefer.

What exactly is the process of *picking* one out of several maximal options? Again, I take my lead from Ruth Chang. Chang has argued that, when faced with a choice between incommensurable goods, you can, simply by willing, cause there to be a reason to select one way rather than the other (Chang 2017). She calls these "voluntarist reasons" or "willed reasons" to contrast them with "given reasons". Your given reasons, in this sense, are those reasons that do not depend on an exercise of your will, and will include, for example, reasons due to your preferences. Given reasons are reasons to *choose*, whereas willed reasons give you a reason merely to *pick*. Willed reasons are weak in the following sense:

> Only when your given reasons 'run out', that is, when they fail to determine what you have most reason to do, can you create a

[14] On the idea of a graded concept of (ir)rationality, see Schervish, Seidenfeld, and Kadane (1997), Staffel (2020).

> voluntarist reason in favour of one alternative over the other. The existence of your voluntarist reasons depends on your given reasons running out. (Chang 2013, 104)

So you can will there to be a reason to pick one way rather than the other, and such acts of the will are subject to rational evaluation when you are willing a reason to pick one merely maximal option rather than another.[15] But how do you determine how you should will?

## Commitment

So, how and why does a rational picking procedure favour selecting the sweetened option on Tuesday? There's two parts to my answer. First, selecting to put the money in box A on Monday comes with certain commitments. Commitments to pick in a dynamically consistent fashion at later choice points. And second, at that later choice point (on Tuesday), the picking procedure tells you to honour those earlier commitments (to the best of your ability). That is, I think that we should understand the procedure for picking as involving not a plan, but a commitment.

So what is a commitment? Think of a commitment to pick along the lines of commitment in the context of assertion, belief or acceptance. To accept $P$ is to be committed to accepting $P$'s logical consequences.[16] Or to assert $P$ in conversation is, under normal circumstances, to commit to the logical consequences of $P$. I don't need to spell out what all those consequences are in order to be committed to them, nor do I even need to know what they are.[17] And being committed to those consequences is not some further act of will, it is simply part of what it is to accept or to assert $P$. This distinguishes commitments from plans: plans are conscious, commitments can be implicit. Note that this is quite different from Ruth Chang's use of the term "commitment" in this context (Chang 2017), where for her, a commitment is an explicit additional act of will – like committing to your marriage vows – that affects how you choose through affecting how you value. Chang's commitments are conscious and separate acts of will, whereas my use of the term "commitment" denotes an automatic and implicit consequence of choice.

And that's it. That's all there is to solving the puzzle. The solution to the

[15] This approach to picking works for preference, and also, arguably for the "more reason to $\phi$ than $\psi$ relation, but it's unclear who would be doing the willing in the case of the objective betterness relation. How exactly picking (and commitment) work for the objective betterness relation is something I won't explore further in this paper.

[16] Kyburg (1983), section 4 puts things in these terms, and the view is also close to what Stalnaker (1991) calls "implicit belief".

[17] One might suggest that it's consistent with my having asserted $P$ that, on being apprised of some consequence of $P$, I withdraw my commitment to $P$ instead of believing the consequence. Analogously, perhaps on Tuesday, instead of being committed to pick box A, I can regret placing the money in box A on Monday, and pick box B. I won't follow this line of thought further here.

sequential choice puzzles for incommensurable goods boils down to "it's rationally favoured to pick so as to be dynamically coherent". I call this view Dynamically Coherent Rational Picking (DCRP), and together with the rational silence afforded by denying **MSP** we have a response to the holiday puzzle that not only refrains from making box B permissible, but gives some element of positive favouring to selecting box A.

We can now understand the two senses of rational permission that Chang mentioned in the passage quoted in a previous section. In one sense, it's "rationally permissible" to select the unsweetened option in the holiday puzzle, in the sense that that option is maximal on Tuesday. The scare quotes are necessary, because I don't think this sense of "rationally permissible" is really rational permissibility.[18] On the other hand, there's a sense in which it's not permissible to do so: the sense in which there's a rational procedure for picking among the maximal options that doesn't sanction doing so (because of a previous commitment).

We can now add more detail to Chang's suggestion that the rationality of making a particular decision "depends on understanding the rationality of choice against a background of other choices". It's not past selections, *per se* that make up the background of later selections, it's the commitments you make in selecting one way rather than the other that influence the later pickings. Commitments, that is, to pick in a dynamically coherent way at later choice points, if that is possible.

At this stage, one might wonder why you should follow through on your commitments at the later time. That is, why should you honour the commitment you made on Monday when making the decision on Tuesday? Note this is similar to the problem often posed for theories of rational choice that rely on planning: why should you later enact a plan you made earlier? There's really two questions here: one, why should you (generically) follow the procedure to pick in a dynamically coherent way; and two, why should you pick in accordance with that procedure on this particular occasion. I'll discuss the first question in a later section, so I'll say a few things about the second question here. So the question is why should a mere commitment have motivational force on this particular Tuesday? My answer is just to point out that, on Tuesday, you have to pick some way, so why not let the picking procedure honour past commitments? You don't take the commitment to have any particularly strong motivational force on Tuesday, but it's a tie-breaker that, if adopted as a general policy, has some benefits. Compare: there's no particularly strong reason for me to exercise on this particular Tuesday, but a policy of exercising regularly has benefits (Tenenbaum 2020). Since we're specifically dealing with cases of decisions among incommensurable goods, in the cases at issue there are no coutervailing reasons to act against the policy. (This contrasts with the role planning is supposed to play: planning to act is

[18]That's not to say that I am committed to thinking that such options are rationally impermissible: recall that the kind of deontic favouring and disfavouring that comes with being picked or not picked is a free parameter in the theory as it stands.

supposed to give you a reason to persevere with acting that way contrary to your later preference against acting in the way you planned). Note also that, for all I've said, picking could be quite weak: recall that the picked option receives some sort of "deontically favoured" status, but I've been noncommittal about what that status is. And if we take a view on which that favouring is merely a lukewarm endorsement, then there might be other factors that overrule it.[19]

Note that this theory of commitment and rational picking is consistent with Separability, Preference Determines Rational Choice, Accommodate Unanticipated Decisions, Allow Determinately Second Best and Minimal Attitude Change, if these principles are interpreted to apply to choosing but not to picking. Consistency with these principles demonstrates that this approach to rational sequential decision-making is distinct from extant resolute theories of choice such as the planning approach or mind-making.

What we have here is an explicitly two-layered theory of rational decision. On the deeper, weightier layer, we have more or less the standard theory of rational choice: choice determined by your immediate preferences. This layer of *choice* could be purely internalist and synchronic in a way consistent with Time-Slice Rationality. This layer of decision does not settle all questions of the deontic status of options in the case of incomplete preference relations (because we denied **MSP**). So we supplement this basic layer of decision-making with a second tier: the layer of picking. Unlike standard views of picking, we allow that procedures for picking can be rationally assessed, at least when it comes to picking among non-optimal maximal options. I proposed a particular procedure for picking – DCRP – that seems to do fairly well in the puzzle cases we started with. The formal groundwork for a theory of two-tiered choice rules has been laid (Helzner 2013).[20] Helzner's paper is inspired by Isaac Levi's work on decision-making (1980, 1986), and we can see something of the same structured, layered approach to decision-making in his thought. This picking layer has to be diachronic, but then, since the problem we are trying to overcome – dynamic incoherence in the holiday puzzle – is distinctively diachronic, it is hardly surprising that the solution is too.

[19]This leaves open the question of exactly when you are and when you are not bound by a commitment. Perhaps commitments "decay" over time, or can be overruled by a "change of heart" that comes from reassessing the options in a truly weighty clash of values as in the case of Sartre's student that Moss (2015a) discusses. I don't have strong views here, so I will simply note that I believe my theory has the resources to accommodate the "contours of our normative judgements" (Moss (2015a), p.667). For example, an agent who systematically and strategically acts against DCRP can't really be said to be conforming to it, whereas one-off violations of the policy might be permitted in some circumstances.

[20]The basic idea behind how DCRP fits into Helzner's theory is that you condition the choice function on your path through the decision tree up to the current choice node.

## Dynamic rationality is always maximal

Let's return to the question we set aside in the last section: why is the particular procedure for picking that I propose the correct one to endorse? My argument in favour of DCRP boils down to "Why Aincha Rich?": such a procedure allows a rational agent with incomplete preferences to remain dynamically coherent, and to consistently pick in the intuitively favoured manner in decision problems like the holiday puzzle. The approach I have suggested has the advantage that, in cases involving incommensurable goods, an agent will never end up with something that they strictly disprefer to something else they could have had, unless their current preferences change or were irrational to begin with.[21] Are there other systematic methods of determining how to pick that are worthy of further study? I certainly don't want to rule that out. But this method does at least have the advantage of allowing you to never reject free money in decision problems like the holiday puzzle.

Consider an agent faced with a sequential choice among incommensurable goods. Recall the idea of an "admissible completion" from Section 3. We might call these admissible completions "avatars" (borrowing a term from Bradley (2009)), which are ways the agent could be maximally opinionated consistent with her current opinions. For simplicity, assume that each avatar satisfies the principles of dynamic rationality that McClennen (1990) lists in Chapter 7, in particular, an avatar takes a sequence of choices to be rational if and only if she takes each choice in the sequence to be rational at the time of choice. As we've seen from the holiday puzzle, this isn't sufficient for the agent herself to be dynamically rational. This is a strong assumption that could almost certainly be weakened.

Let's talk through the holiday puzzle from this perspective. Some avatars prefer Scuba to Safari, and others prefer Safari to Scuba: that's what it is to find those options incommensurable. All of them prefer Scuba+Money to Scuba, and Safari+Money to Safari: that's what it is to determinately prefer more money to less. And finally, some avatars prefer Safari+Money to Scuba and some have the opposite preference, and likewise for Scuba+Money and Safari. Leaving aside the possibility of indifferences, there are basically four kinds of avatar. Letting W,X,Y, and Z stand for Scuba+Money, Safari+Money, Scuba and Safari respectively, we can label these four kinds of avatar WXYZ, WYXZ, XWZY, XZWY.[22] So, WXYZ prefers Scuba+Money to Safari+Money to Scuba to Safari.

The options on Monday are to put the money in box A (yielding a choice between X or Y on Tuesday) or put the money in box B (yielding a choice between W or Z on Tuesday), and there is no consensus among the avatars which is the better, so either option is permissible. On Monday, every avatar agrees that

[21] I offer no solution to other examples of what Hedden (2015a) calls "diachronic tragedy" that exploit, for example, infinite state spaces or infinite utilities.

[22] For the purposes of illustration, we're assuming here that if W is preferred to X, then Y is preferred to Z, and vice versa. This assumption that money and holiday values are "independent" isn't really necessary, but it simplifies the presentation.

putting the money in box A and then taking box B on Tuesday is impermissible: every avatar agrees that SCUBA (i.e. Y) is not permissible because every avatar prefers SCUBA+MONEY to SCUBA. There is, however, disagreement about what the best option is (whether it is W or X). As before, let's say the agent opts to put the money in box A. Now, on Tuesday, the avatars are split on which is the better option out of X and Y. And now some avatars are advocating for choosing Y, even though on Monday everyone agreed that choosing Y was impermissible! This is because those who now prefer Y (e.g. WYXZ) are those who wanted to put the money in box B on Monday. The second best option for them (Y) is the best option that remains available. So that's how consensus among avatars can disappear in a sequence of choices.

On Tuesday, there are still some avatars that prefer to select X over Y, and so it is at least a possibility that the agent can pick that way. That is, X and Y are incommensurable, so it is open for a previous commitment to favour picking one over the other. A commitment to pick X over Y on Tuesday (which is a consequence of the way the agent chose on Monday) can break the avatars out of their impasse and yield a dynamically rational sequence of choices for the agent.

Note that in this example, there are no determinately second best options: every maximal option is optimal for some avatar. In this case, it is always possible for an agent whose preferences don't change to commit to pick a dynamically rational sequence of options and then follow through. If a sequence of options is at least indeterminately permissible (in the sense of Bales (2018), i.e. permissible according to some of the avatars), then each act in that sequence is at least indeterminately permissible for the agent. This is clear since if some avatar takes a sequence of choices to be permissible, she considers each member of that sequence permissible, and so each member of that sequence is at least indeterminately permissible for the agent. If an option in a sequence is at least indeterminately permissible then it is possible for a consideration that favours picking it to tip the balance in that option's favour. And thus, if an agent wants to bring about a sequence that is at least indeterminately permissible, she can achieve that through committing to pick that way. Such an agent will never end up selecting each of a sequence of choices that is determinately impermissible. Thus DCRP allows an agent whose preferences don't change to remain dynamically coherent in this sort of case. A more formal and more general version of this result will have to wait for future work.

# Conclusion

The possibility that your preference relation could be incomplete is often granted but then ignored. This is a mistake since presuming the relation to be complete obscures some important subtleties in how preference connects to rational decision. Sometimes it is suggested that having incomplete preferences would permit you to make bad choices in sequential decision problems. Given the correct view of how preference connects to rational permission, this is incorrect. If we deny that

maximality is sufficient for permission, and use a dynamically coherent rational picking procedure, then no bad consequences befall those whose preferences are incomplete.

# References


Bales, Adam. 2018. "Indeterminate Permissibility and Choiceworthy Options." *Philosophical Studies* 175: 1693–1702.

Bales, Adam, Daniel Cohen, and Toby Handfield. 2014. "Decision Theory for Agents with Incomplete Preferences." *Australasian Journal of Philosophy* 92 (3): 452–70.

Bradley, Richard. 2009. "Revising Incomplete Attitudes." *Synthese* 171: 235–56.

Bradley, Seamus. 2015. "How to Choose Among Choice Functions." In *ISIPTA 2015 Proceedings*, 57–66.

———. 2019. "Imprecise Probabilities." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. https://plato.stanford.edu/archives/spr2019/entries/imprecise-probabilities/.

Bratman, Michael. 1992. "Planning and the Stability of Intention." *Minds and Machines* 2: 1–16.

———. 1999. *Faces of Intention: Selected Essays on Intention and Agency*. Cambridge University Press.

Broome, John. 1991. *Weighing Goods*. Blackwell.

Buchak, Lara. 2013. *Risk and Rationality*. Oxford University Press.

Chang, Ruth, ed. 1997. *Incommensurability, Incomparability and Practical Reason*. Harvard University Press.

———. 2002. "The Possibility of Parity." *Ethics* 112: 659–88.

———. 2005. "Parity, Interval Value and Choice." *Ethics* 115: 331–50.

———. 2013. "Commitments, Reasons and the Will." *Oxford Studies in Metaethics* 8: 74–113.

———. 2017. "Hard Choices." *Journal of the American Philosophical Association* 92: 586–620.

Doody, Ryan. 2019a. "If There Are No Diachronic Norms of Rationality, Why Does It Seem Like There Are?" *Res Philosophica* 96 (2): 141–73.

———. 2019b. "Parity, Prospects and Predominance." *Philosophical Studies* 176: 1077–95.

Elga, Adam. 2010. "Subjective Probabilities Should Be Sharp." *Philosophers' Imprint* 10.

Hare, Caspar. 2010. "Take the Sugar." *Analysis* 70 (2): 237–47.

———. 2013. *The Limits of Kindness*. Oxford University Press.

Hedden, Brian. 2015a. "Options and Diachronic Tragedy." *Philosophy and Phenomenological Research* XC: 423–51.

———. 2015b. *Reasons Without Persons*. Oxford University Press.

Helzner, Jeffrey. 2013. "Rationalizing Two-Tiered Choice Functions Through Conditional Choice." *Synthese* 190: 929–51.

Kagan, Shelley. 1991. *The Limits of Morality*. Oxford University Press.

Kyburg, Henry E. 1983. "Rational Belief." *The Brain and Behavioural Sciences* 6: 231–73.

Levi, Isaac. 1980. *The Enterprise of Knowledge*. The MIT Press.

———. 1986. *Hard Choices: Decision Making Under Unresolved Conflict*. Cambridge University Press.

Machina, Mark J. 1989. "Dynamic Consistency and Non-Expected Utility Models of Choice Under Uncertainty." *Journal of Economic Literature* 27: 1622–68.

McClennen, Edward F. 1990. *Rationality and Dynamic Choice: Foundational Explorations*. Cambridge University Press.

Moss, Sarah. 2015a. "Credal Dilemmas." *Noûs* 49: 665–83.

———. 2015b. "Time-Slice Epistemology and Action Under Indeterminacy." *Oxford Studies in Epistemology*, 172–94.

Peterson, Martin. 2007. "Parity, Clumpiness and Rational Choice." *Utilitas* 19: 505–13.

Portmore, Douglas W. 2019. *Opting for the Best: Oughts and Options*. Oxford University Press.

Schervish, Mark J., Teddy Seidenfeld, and Joseph B. Kadane. 1997. "Two Measures of Incoherence: How Not to Gamble If You Must." Department of Statistics, Carnegie Mellon University.

Schoenfield, Miriam. 2014. "Decision Making in the Face of Parity." *Philosophical Perspectives* 28: 263–77.

Seidenfeld, Teddy. 1988. "Decision Theory Without 'Independence' or Without 'Ordering'. What's the Difference?" *Economics and Philosophy*, 267–90.

———. 2004. "A Contrast Between Two Decision Rules for Use with (Convex) Sets of Probabilities: Gamma-Maximin Versus $E$-Admissibility." *Synthese* 140: 69–88.

Staffel, Julia. 2020. *Unsettled Thoughts*. Oxford University Press.

Stalnaker, Robert. 1991. "The Problem of Logical Omniscience, I." *Synthese* 89: 425–40.

Steele, Katie, and H. Orri Stefánsson. 2020. "Decision Theory." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. https://plato.stanford.edu/archives/win2020/entries/decision-theory/; Metaphysics Research Lab, Stanford University.

Suzumura, Kotaro. 1983. *Rational Choice, Collective Decisions and Social Welfare*. Cambridge University Press.

Tenenbaum, Segio. 2020. *Rational Powers in Action*. Oxford University Press.

Ullmann-Margalit, Edna, and Sidney Morgenbesser. 1977. "Picking and Choosing." *Social Research* 44 (4): 757–85.

Williams, J. R. G. 2014. "Decision-Making Under Indeterminacy." *Philosophers' Imprint* 14: 1–34.

———. 2017. "Indeterminate Oughts." *Ethics* 127 (3): 645–73.