Is Evaluative Compositionality a Requirement of Rationality?

Nicholas J. J. Smith
The University of Sydney
njjsmith@sydney.edu.au

This paper presents a new solution to the problems for orthodox decision theory posed by the Pasadena game and its relatives. I argue that a key question raised by consideration of these gambles is whether evaluative compositionality (as I term it) is a requirement of rationality: is the value that an ideally rational agent places on a gamble determined by the values that she places on its possible outcomes, together with their mode of composition into the gamble (i.e. the probabilities assigned to them)? The paper first outlines a certain simple response to the Pasadena game and identifies two problems with this response, the second of which is that it leads to a wholesale violation of evaluative compositionality. I then argue that rationality does not require decision makers to factor in outcomes of arbitrarily low probability. A method for making decisions which flows from this basic idea is then developed, and it is shown that this decision method (Truncation) leads to a limited — as opposed to wholesale — violation of evaluative compositionality. The paper then argues that the truncation method yields solutions to the problems posed by the Pasadena game and its relatives that are both attractive in themselves and superior to those yielded by alternative proposals in the literature.

1. Introduction

Nover and Hájek (2004) present a new type of gamble — the Pasadena game — which, they claim, presents a 'headache', a 'serious problem', a 'paradox' for orthodox decision theory (p. 248, p. 237, p. 245). In response, some authors propose specific additions or modifications to orthodox decision theory to accommodate the Pasadena game and its ilk (Colyvan 2006, 2008; Easwaran MS, 2008; Sprenger and Heesen MS) while others argue that really there is no problem and that orthodox decision theory is fine as is (Fine 2008).[1]

The importance of a puzzle is measured not by the quantity of responses it sparks but by the depth of the issues that discussion of

[1] Baker (2007) has a foot in both camps: he argues that the Pasadena game does not pose a headache for orthodox decision theory but then proposes a new gamble — the Alternating St Petersburg Game — which, he argues, does.

Mind, Vol. 123 . 490 . April 2014
doi:10.1093/mind/fzu072

the puzzle brings to light. I shall argue that consideration of the Pasadena game does indeed lead to a very deep issue: evaluative compositionality. The issue is whether the value which an ideally rational agent places on a bet is determined by the values which she places on the possible outcomes of the bet, together with their 'mode of composition'—the way they are combined into the bet: that is, the probabilities assigned to the outcomes by the bet.[2]

I shall argue that evaluative compositionality is not a requirement of rationality, and that this provides the key to seeing why the Pasadena game and its relatives do not ultimately threaten orthodox decision theory. This is not to say, however, that 'anything goes' when it comes to valuing a gamble: rather, compositionality fails in a specific, constrained way.

The paper proceeds as follows. Section 2 introduces the Pasadena game and the problem it apparently poses for orthodox decision theory. Section 3 outlines a certain simple response to the game, according to which it poses no problem at all. Sections 4 and 5 outline two problems with this response; the second problem is that the response leads to a wholesale violation of evaluative compositionality. Section 6 introduces the idea that rationality does not require decision makers to factor in outcomes of arbitrarily low probability. Section 7 spells out a method for making decisions which flows from this basic idea, section 8 shows that this method leads to a limited (as opposed to wholesale) violation of evaluative compositionality, and section 9 considers possible additions to the decision method. Section 10 then argues that the method yields solutions to the problems posed by the Pasadena game and its relatives that are both attractive in themselves and superior to those yielded by alternative proposals in the literature.[3]

2. The Pasadena game

A coin is tossed as many times as it takes for the coin to land Heads; that is, upon the first appearance of Heads, the game is over. The greater the number of tosses required, the greater the *magnitude*

[2] In this paper I follow established practice in the Pasadena literature in using the terms 'gamble, 'lottery', 'bet', and 'game' interchangeably. Note however that the Pasadena game is not a 'game' in the sense in which that term is used in game theory; cf. Fine 2008, p. 613.

[3] In this paper I follow the Pasadena literature in restricting attention to lotteries with countably many outcomes—either finitely many or countably infinitely (denumerably) many. Hence when I speak of 'infinite' lotteries I shall always mean lotteries where *countably* infinitely many outcomes are assigned non-zero probability.

(i.e. absolute value) of the payoff (prize) of the game: if Heads appears on toss *n*, the magnitude is $\$\frac{2^n}{n}$. But there is a twist: the *sign* of the payoff alternates from positive to negative according to whether Heads appears on an odd-numbered toss or an even-numbered toss. So if Heads appears on toss 1, the prize is (positive) \$2 (i.e. the person offering the game pays this amount to the player); if Heads appears on toss 2, the prize is *negative* \$2 (i.e. the player pays \$2 to the person offering the game); if Heads appears on toss 3, the prize is $\$\frac{8}{3}$; if Heads appears on toss 4, the prize is −\$4; and so on.

So far so good. One's problems begin if one wishes to value this game at its *expected value*. For supposing that we identify the value (utility) of a prize with its amount in dollars,[4] the game *has no* unique expected value. That is because the value of the infinite sum[5]

$$\sum_{n=1}^{\infty} \frac{1}{2^n}(-1)^{n-1}\frac{2^n}{n}$$

can be made to be whatever we like (negative infinity, positive infinity, or any negative or positive finite number) simply by varying the *order* of the terms in the series (for details see Nover and Hájek 2004, Sects 1–2). As standardly conceived in decision theory, however, a gamble is well defined once its possible outcomes together with their probabilities are specified: it is not in addition required that an *ordering* of the outcomes be given. Hence, the Pasadena game — in itself, without any additional specification of the order in which the possible outcomes are to be taken in the expected utility calculation — is a well-defined gamble which has no (unique) expected value.

3. Chocolate ice cream

So what? The Pasadena game has no expected value: Why should we think this is a *problem*? I shall begin by discussing one reason why it is

[4] This supposition makes for simplicity of presentation: it is not essential to generating the problem. Elsewhere in this paper I shall sometimes identify utility and monetary value without further comment — but as in the present case, nothing will hang on this except ease of presentation.

[5] The first part of the formula, $\frac{1}{2^n}$, is the probability that the coin will first land Heads on toss *n*. The remainder of the formula, $(-1)^{n-1}\frac{2^n}{n}$, is the value of the outcome of the bet in which the coin first lands Heads on toss *n*. Within this part of the formula, the second part, $\frac{2^n}{n}$, is the magnitude of the value of the outcome, while the first part, $(-1)^{n-1}$, captures the fact that the value is positive or negative according to whether *n* is odd or even.

not a problem — that is, a mistaken reason for thinking that it is a problem. Understanding why this line of thought is mistaken is essential to understanding the real problem posed by the Pasadena game, and the solution to this problem to be presented in this paper.

The following passages suggest the mistaken line of thought:

> once we see the differences that rearrangements can make, we realize that we can say absolutely nothing about the value of the Pasadena game. The game is apparently well defined, and yet decision theory cannot handle it. Something has to give — either the game itself, or decision theory. (Nover and Hájek 2004, p. 242)

> As decision theory stands, we cannot place the value of the Pasadena game on a cardinal scale that would allow numerical comparisons between it and the values of other prospects. There is no *total ordering* of prospects that includes the Pasadena game. (Hájek and Nover 2006, p. 713)

> the Pasadena game's expectation is undefined, so it seems that it cannot be valued at all. (Hájek and Nover 2008, p. 644)

The mistaken line of thought is that, given orthodox decision theory, the Pasadena game *cannot be assigned a value*; that, because it has no expected value, it is therefore impossible to assign it a value at all, within orthodox decision theory; that it is, according to orthodox decision theory, a singularity in value space — a place where one's utility function must be *undefined*.[6] This is a mistake for the following reason. If the only thing that decision theory tells us about lotteries is to value them at their expected values, then — given that the Pasadena game has no expected value — decision theory tells us nothing about how to value it. As Nover and Hájek (2004, p. 241) say, decision theory 'goes silent'. But this does not mean that we cannot value the Pasadena game, within orthodox decision theory: it means that we can value it *however we like*. Compare chocolate ice cream. Decision theory tells us nothing about how to value it.[7] This does not mean that one cannot

[6] I do not mean 'singularity' in the technical sense in which that term is used in mathematics and physics: I mean to use it as a term for an object in the domain of a partial function for which that function is undefined — that is, an object to which the function assigns no value. I am not aware of an existing simple term for the latter notion, nor of any English word apart from 'singularity' which has the right connotations — hence the decision to employ this term with a technical meaning distinct from the technical meaning it already has.

[7] More precisely, decision theory tells us nothing in absolute terms about how to value chocolate ice cream. It does tell us — via the basic preference axioms — such things as that if we place a higher value on chocolate ice cream than on strawberry ice cream and place a higher value on strawberry ice cream than on butterscotch ice cream then we should place a higher value on chocolate ice cream than on butterscotch ice cream.

assign any utility at all to receiving a bowl of chocolate ice cream, within orthodox decision theory: one can assign it whatever value one wants. The point is that where decision theory gives no advice as to how to value something, this does *not* make that thing a singularity in value space. Rather than being a point at which one's utility function is *undefined*, it is a point at which one's utility function is *unconstrained*. Things which decision theory does not tell us how to value are *freely* valuable, not *un*valuable.

One thing that can make this point hard to see is the idea — call it the 'value-setting' idea — that, within orthodox decision theory, one does not value a lottery at all *until* one has determined its expected value — and then one *sets* one's value for the lottery equal to its expected value.[8] It follows from the value-setting idea that if a lottery has no expected value, then one does not value it — at all. But the value-setting idea is not the right way to think about things. On the classical, orthodox approach to decision theory, derived from von Neumann and Morgenstern (1944), it is taken as a starting point that one has preferences over the set of all (simple) lotteries.[9] Assuming one's preferences meet some basic consistency constraints, it then turns out that one can be represented as assigning a numerical value (unique up to a positive linear transformation) to each lottery and that the value one assigns to a lottery is in fact equal to its expected value (given the values one assigns to its outcomes). So the value of the lottery is not *generated by* the expected value. Rather, there are two numbers — the value and the expected value — and it turns out to be a *fact* that they are always the *same* number.

This provides a very nice answer to the normative question as to why one should maximize expected utility. On the value-setting approach, the question arises: *Why* should I so set my value for the lottery? A standard answer appeals to the laws of large numbers to show that if one repeatedly buys a lottery at a price higher (lower) than its expected value, one will almost certainly make a loss (gain) in the long term. This answer is by no means watertight, however. For it is not at all clear that

[8] This idea seems to be widespread amongst that large group of philosophers who are familiar with the idea of maximizing expected utility but have not closely studied the technical foundations of decision theory.

[9] A *simple* lottery is one which assigns non-zero probability to only finitely many outcomes.

from the fact that there is a single best way of doing something given that one will be doing it many many times in a row, it follows that one should also do the thing this way given that one will be doing it only once. Thus, on the value-setting approach, the question remains open: Why, faced with a single lottery, should I value it at its expected value? The classical (von Neumann-Morgenstern) approach to decision theory, on the other hand, provides a definitive answer to this question. One should choose the lottery with the highest expected value because one values it the most! (The only reason this could fail to be the case is if one violates the consistency constraints on preferences — and one should not do that.) As Resnik (1987, p. 99) puts it: ‘In choosing an act whose expected utility is maximal an agent is simply doing what he wants to do!’

An analogy may be useful here. Suppose that you move house, and the movers give you a sheaf of adhesive dots, in various shades on a spectrum from yellow through orange to red, to stick on your possessions. The idea is that the more you care about an object, the redder a sticker you should put on it (i.e. the closer its sticker should be to the red end of the spectrum) and the less you care about an object, the yellower a sticker you should put on it (i.e. the closer its sticker should be to the yellow end of the spectrum). The rationale is that no movers are perfect: it is no use telling them that *everything* you own is supremely valuable, for they are bound to break *something*. This way, however, they can apportion their efforts appropriately: the redder the sticker on a given object, the harder they will try not to drop it. Great — so you moved a while ago, but you never got around to removing the stickers. Suddenly there is a knock at the door: the fire brigade is here to evacuate you. There has been an explosion at a paint factory in your neighbourhood and a plume of toxic smoke is spreading across the suburb. You have two minutes to grab your most valued possessions and leave. ‘Grab the things with the red stickers!’ you yell to your family … Later on, when the danger has passed and you are returning home, a fireman says to you: ‘Hey, I know it’s none of my business, but I just gotta ask. What’s so important about those little red stickers? Why do you care about them so much?’ So you explain … The key point is that you do not value the stickers themselves, nor do you value things highly *because* they have red stickers on them. The redness of its sticker is not a *reason* to value a thing: it is a *sign* that you *do* value it. The same can be said of expected utilities in the framework of classical decision theory. The idea is not that you *should* value a certain lottery more than another *because* it has the higher expected utility; the idea is that you *do* value it more. Greater expected utility is a *sign* of greater value, not a *reason* to value a lottery more highly.[10]

Let us now reconsider the three quotations at the beginning of this section. It is not clear that these passages were intended to endorse the view that the Pasadena game is, according to standard decision theory, a singularity in value space: there are other ways of reading the passages and there are other passages in these papers which suggest that their authors do not hold this view. In any case — whether or not Nover and Hájek so regard(ed) it — the crucial point is that it is incorrect to think that, on orthodox decision theory, the Pasadena game is a singularity in value space. It is not true that, within orthodox decision theory, we can say absolutely nothing about the value of the Pasadena game: rather, we can say whatever we like. It is not true that decision theory cannot handle the Pasadena game — that something has to give, either the game or the theory: decision theory handles the Pasadena game in just the way it handles chocolate ice cream. It is not true that, according to orthodox decision theory, we cannot place the value of the Pasadena game on a cardinal scale that would allow numerical comparisons between it and the values of other prospects: we can place it wherever we please. It is not true that the Pasadena game cannot be valued at all: it is *freely* valuable, not *un*valuable.

At this point, one might think that the problem posed by the Pasadena game has been resolved — that is, that there was really no problem at all. Orthodox decision theory accommodates the Pasadena game in just the way it accommodates chocolate ice cream: both are freely valuable (not unvaluable). Call this the 'chocolate ice cream response'. Something along these lines seems to be the response of Fine (2008) — although Fine actually goes further and *proves* that the Pasadena game can be assigned any real-numbered value at all without conflicting with the axioms of orthodox decision theory.

4. Problems with chocolate ice cream: dominance

The problems have *not* all been (dis)solved, however. There are two objections to the chocolate ice cream response. The first one was noted

[10] All that matters for the analogy is that *as a matter of fact*, your values are correlated with degrees of redness in one case, and with expected values in the other case: that is, your values are *in step with* redness/expected utility but this is not because your values *follow* redness/expected utility. The *mechanism* by which the correlation is set up in the first place is not important in this context — hence it is irrelevant that the mechanism is different in the two cases.

already by Nover and Hájek (2004).[11] After noting that decision theory goes silent on the Pasadena game (i.e. does not assign it an expected value), they continue:[12]

> It is an uncomfortable silence. For intuition tells us — indeed, yells at us — that we can make meaningful comparisons between the Pasadena game and other games. It is clearly worse than the St. Petersburg game, for starters. It is clearly worse than a neighbouring variant of the game — call it the *Altadena game* — in which every pay-off is raised by a dollar. (Notice that the Altadena game has all the problems of the Pasadena game.) And the Pasadena game is clearly better than a 'negative' St. Petersburg game, in which all the pay-offs of the St. Petersburg game are switched in sign. Yet expected utility theory can say none of this. (Nover and Hájek 2004, pp. 241–2)

The problem, then, is that allowing agents total freedom over how to value games such as Pasadena and Altadena conflicts with the clear intuitive judgement that one should value Altadena more highly.

At this point the proponent of the chocolate ice cream response might try to brush off the objection as follows. Classical decision theory does not tell us to value a bowl of chocolate ice cream over a bowl of toxic sludge scraped from the bottom of a vat in a fly spray factory, even though intuition tells us — indeed, yells at us — that the chocolate ice cream is the better option. It is, then, no objection to decision theory that it fails to back up all our intuitive judgements concerning what is better than what.

This response is mistaken. The thought that the Altadena game is better than the Pasadena game (etc.) is backed up by more than just (brute) intuition: as Colyvan (2006, pp. 698–9) points out, it is backed up by dominance reasoning. The two games have the same possible outcomes (Heads first on toss 1, Heads first on toss 2, …) and assign each outcome the same probability — but the *value* of each outcome is higher in the Altadena game than in the Pasadena game. Thus, whatever happens, you will be better off *by your own lights* having chosen the Altadena game than the Pasadena game: the Altadena game *dominates* the Pasadena game. The point of the toxic sludge case is that it is not decision theory's job to tell one how to value things in any

[11] Although not, of course, in these terms — that is, as a problem for the chocolate ice cream response — because Nover and Hájek (2004) did not draw the distinction made in the previous section between the Pasadena game being *freely* valuable and its being *un*valuable.

[12] The St Petersburg game is like the Pasadena game in that a coin is tossed until it lands Heads. Where n is the number of the toss on which the coin first lands Heads, the payoff of the game is $\$2^n$.

absolute sense. However it *is* within decision theory's remit to tell one how to value things *given* how one values other things (for example, decision theory *would* tell you to prefer chocolate ice cream to toxic sludge *if* you had already specified that you prefer chocolate ice cream to strawberry ice cream and strawberry ice cream to toxic sludge) — and it looks as though dominance should come in here. For in the Altadena versus Pasadena case, the point is that *by your own lights* you will be better off having chosen Altadena, no matter what outcome occurs: that is, *you* value the outcomes of the Altadena game more highly (and the two games give each outcome the same probability). So it seems that it would be *irrational* not to prefer Altadena. Hence decision theory, qua theory of ideal rational choice, *should* advise one to choose Altadena, rather than allowing one to rank these games however one pleases.

5. Problems with chocolate ice cream: evaluative compositionality

The second problem with the chocolate ice cream response is that it conflicts with the principle that evaluative compositionality is a requirement of rationality. This problem has not been brought into focus in the literature — indeed, the principle of evaluative compositionality has not been explicitly formulated.[13] Let me then explain the principle, outline the case for taking it to be a condition on being a rational agent that one conform to this principle, and then show how the chocolate ice cream response flouts the principle.

The principle of evaluative compositionality is this: the value which a rational agent places on a gamble is a function of the values which she places on the possible outcomes of the gamble, together with the probabilities assigned to those outcomes by the gamble. That is, the

[13] This is not to say that no author has said anything that touches on the problem. For example, Hájek and Nover (2008, pp. 660–1) mention the issue of the 'supervenience of the values of certain compound gambles on the values and probabilities of their constituents' and say that failures of such supervenience make them 'uneasy', and Fine (2008, p. 620) writes (cf. also p. 629): 'The standard axioms of utility theory so restrict allowable notions of preference that the associated utility function for certain complex gambles is determined by the utility function restricted to much simpler gambles. Linear utility theory is about "rationally" reducing the complexity of choices between gambles. By making certain simpler choices we can then delegate to the mathematics to say what our choices between more complex gambles should be. What makes [the Altadena, Pasadena and St Petersburg gambles] interesting is their resistance to being encompassed in this frame, while still allowing for the fundamental (subjective) binary preferences to include them.'

value of a gamble is a function of the values of its outcomes together with their *mode of composition* into the gamble (for the way that we combine possible outcomes into a gamble is precisely by specifying probabilities for each outcome: the assignment of probabilities is therefore the mode of composition of the outcomes and the gamble is the compound object formed from them).

Compare the principle of semantic compositionality (also known as Frege's principle): the meaning of a complex expression (typically a sentence) is a function of the meanings of its constituent expressions together with their mode of composition — that is, the way they are combined syntactically to form that particular complex expression. Two familiar manifestations of semantic compositionality are:

Supervenience: If two sentences with the same syntactic structure have different meanings, then some expression in one of them must have a different meaning from its counterpart in the other sentence.[14]

Transparency: If we substitute for an expression in a sentence another expression of the same syntactic category and with the same meaning, then the meaning of the whole sentence does not change.

Evaluative compositionality has analogous manifestations; in order to state them it will be helpful to introduce some terminology. Given a gamble X, let O_X be the set of possible outcomes of X. For any $o \in O_X$, let $X(o)$ be the probability assigned to o by X.[15] Say that two gambles X and Y are *isomorphic* if there exists a bijection $f : O_X \rightarrow O_Y$ such that for all $o \in O_X$, $X(o) = Y(f(o))$ (i.e. the probability assigned to o by gamble X is the same as the probability assigned to $f(o)$ by gamble Y); call such a bijection an *isomorphism* from X to Y. Given an isomorphism f from X to Y, call a pair (o, p) with $o \in O_X$ and $p \in O_Y$ a pair of

[14] By 'its counterpart' we mean the expression in the other sentence which occupies the same position as it does in the syntactic structure which the sentences share.

[15] Strictly speaking, a lottery X determines a probability measure, which assigns probabilities to *subsets* (not *elements*) of the sample space O_X. In a lottery with a countable sample space, however, we can restrict our focus to singleton sets: for by countable additivity, the probability assigned to any countable subset of the sample space is the sum of the probabilities assigned to the singleton sets of the members of that subset — and every subset of a countable set is countable. In this case it is convenient to write $X(o)$ in place of $X(\{o\})$. Thus, apparent talk ('$X(o)$') of the probability assigned to an *element* of the sample space — that is, to a possible outcome of the lottery — is shorthand for talk ('$X(\{o\})$') of the probability assigned to the singleton set containing just this outcome.

f-counterparts just in case $p = f(o)$. Now the two manifestations of evaluative compositionality can be stated as follows:

> **Supervenience:** If a rational agent values two gambles X and Y differently and there is an isomorphism f from X to Y, then there must be a pair of f-counterparts such that the agent places different values on the two elements of this pair.

> **Transparency:** If we substitute for an outcome of a lottery a different outcome with the same value for a given agent, then that agent's value for the lottery does not change.

Why might one think that satisfying evaluative compositionality is a condition on being a rational agent? Well, let us start by explaining what is wrong with a certain kind of apparent counterexample to the principle. Suppose you offer me the following two gambles:

(1) You will toss a fair coin: Heads I get $20; Tails I get nothing.
(2) You will balance an egg on the point of one of the spikes of the wrought iron fence surrounding the university and then wait for it to fall; if it falls inside I get $20; if it falls outside I get nothing.

These two gambles offer the very same outcomes — a fortiori, outcomes which I value equally — at (we may suppose for the sake of argument) the same probabilities. So, by the principle of evaluative compositionality, I ought to value the gambles equally. But I do not. I prefer the first gamble. The second gamble repulses me: what a waste of an egg! Surely this reaction of mine is not *irrational*.

Indeed it is not an irrational preference — but what this signals is that the situation has not been described correctly. The two gambles which I value differently have *different* outcomes, and I do *not* value these outcomes equally. The first gamble has outcomes of (a) $20 and (b) nothing (the status quo). The second gamble has outcomes of (a′) $20 *and a broken egg* and (b′) the status quo plus a broken egg. As I do not like eggs to be broken (unless we are making omelettes), I prefer the first lottery.

In light of this resolution of the apparent counterexample to evaluative compositionality — in which an apparent difference in value between two gambles with the *same* outcomes turned out really to involve gambles with *different* outcomes which are *not* equally valuable — it starts to seem very plausible that once we describe gambles properly, anything a rational agent likes or dislikes about a gamble *must* be reflected either in the values he assigns to its outcomes, or in the probabilities which the gamble assigns to those outcomes. What else could make one gamble rationally preferable to another? Once the outcomes of the gamble are specified accurately (unlike in the original presentation of the egg-on-the-fence gamble), all of what an agent cares about in relation to how the gamble might turn out is encapsulated in the values he places on the possible outcomes; so it seems that the only thing which could rationally affect his value for the gamble itself is how likely it renders each of the things he cares about — that is, each outcome. So if a rational agent knows how much he values each outcome, and knows how probable each outcome is, then what *else* could he possibly need to know, before he can value the lottery itself? What else could possibly make a *rational* difference?

We can make the point in terms of Supervenience. If an agent values two isomorphic gambles differently, even though she values counterpart outcomes of those gambles equally — that is, if she violates Supervenience — then we may ask *what it is* that is tipping her preference one way or the other. It just does not look as though there is room for any further rational input, beyond the values placed on the outcomes and the probabilities of achieving those outcomes. That is, it looks as though the tipping one way or the other must be irrational. We can also put the point in terms of Transparency. Suppose that we substitute an outcome of a gamble with a new outcome which the agent values equally — and her value for the gamble changes. Why did it change? Not because of how much she *valued* some possible outcome of the gamble, nor because the *likelihood* of getting what she values changed. So why then? The fact that her value for the gamble changed starts to look distinctly irrational — as though it changed not for any good reason, but simply on a whim, because the wind shifted, as it were. It looks as though gambles are transparent contexts: transparent to the values of their outcomes. That is, all that matters to the value of a gamble — apart from the probabilities it assigns to outcomes — are the *values* of the outcomes: other aspects of the outcomes are relevant *only* in so far as they have an impact on the values which an agent places on the outcomes. In other words, it looks as though evaluative compositionality *is* a requirement of rationality.

Well, not quite: there is a second possibility that has been obscured in the preceding discussion. Consider again the case where we substitute an outcome of a gamble with a new outcome which the agent values equally — and her value for the gamble changes. We said this change looks irrational — but that is true only given an extra assumption: that the original value was *rationally mandated*. In other words, what seems true is this: if it is mandatory for a rational agent to place a certain value on a gamble, then it is mandatory for her to place the same value on any gamble which differs from it only by the replacement of some outcome by a different outcome which she values equally. But evaluative compositionality does not follow from this: for it to follow, we need the additional assumption that for every gamble, there is a unique value that any rational agent is mandated to place on it (given the values that she places on its outcomes). Now suppose this assumption is false: then there need be nothing irrational about the example of Transparency violation considered above. Suppose that there is no particular value that a rational agent must place on the first gamble (given her values for the outcomes): there are many possible values, all of which are equally rationally permissible. Then it seems that there is nothing irrational about picking one of these values for the first gamble, and a different one for the second gamble. What *does* seem highly plausible is that if two gambles differ only by the replacement of an outcome by a different outcome which the agent values equally, then the *range* of rationally admissible values of the two gambles — where this range may include just one, rationally *mandated* value, or multiple rationally *admissible* values — must be the same:

> **Irrelevance of Further Factors (IFF):**
> Whether it is rationally permissible for an agent to place a certain value on a certain gamble is determined by the values she assigns to its outcomes and the probabilities assigned to those outcomes by the gamble.

There are, then, two options open. One is evaluative compositionality. The other is the view that some gambles *have no* unique rational value, no unique rational price (i.e. not even once one has specified one's values for the outcomes). On this view, it is not that something *else* — together with the values and probabilities of the outcomes — determines the (unique) rational price. Rather, there *is no* (unique) rational price. I have no argument against this view. I do think, however, that it is prima facie unattractive. That means, I take it, that in the absence of any cogent argument one way or the other, we are entitled to assume (provisionally — until arguments should be forthcoming) that every gamble has a unique rational price — and hence that decision theory should tell us what it is. Furthermore, I take it that part of the very idea of a *rational* price is that it is not a brute or primitive fact what the rational price is. Rather, the rational price is determined by other facts — such as the values and probabilities of outcomes — and can in principle be calculated given knowledge of those other facts:

> **Unique rational price (URP):**
> For any well-defined gamble and any rational assignment of values to its outcomes, there is a particular value which a rational agent who assigns those values to its outcomes must assign to the gamble. In short: every gamble has a unique rational price. Furthermore, this rational price is not brute: it is determined by other facts and can in principle be calculated given knowledge of those other facts.

Summing up: It follows from (IFF) and (URP) that evaluative compositionality is a requirement of rationality. (IFF) says that the values and probabilities of its outcomes fix the set of admissible values of a gamble; (URP) says that, given values for its outcomes, the set of admissible values for a gamble is one-membered. Given both claims, it follows that a rational agent is mandated to place a particular value on each gamble: a value that is determined by the values he places on the outcomes of the gamble, together with the probabilities assigned to those outcomes by the gamble. Once we set aside apparent counterexamples along the lines of the egg-on-the-fence example, (IFF) seems extremely plausible. (URP) is simply an article of faith: however, its prima facie plausibility means that we are entitled to maintain it — provisionally — *in the absence of* any good reason for rejecting it.

So much for the case for evaluative compositionality. Now recall the chocolate ice cream response to the Pasadena game: the game is in the same boat as chocolate ice cream; it is not unvaluable — a singularity in value space; it is simply freely valuable — a rational agent can place any value she likes on it. This response flouts evaluative compositionality: for according to this response, a rational agent's value for the Pasadena game is *not* determined by her values for its outcomes together with the probabilities assigned to those outcomes by the gamble. Two rational agents who place exactly the same values on the possible outcomes of the game are free to value it differently.

Let us take stock. We have noted two problems with the chocolate ice cream response: it violates the idea that it is a requirement of rationality that one prefer a dominant lottery to a dominated one; and it violates the idea that evaluative compositionality is a requirement of rationality.

Motivated by the *first* of these problems, Colyvan (2006) proposes adopting a plurality of decision rules, with dominance having an equal place alongside the principle of maximizing expected utility, while Easwaran (2007) and Colyvan (2008) propose unified extensions or generalizations of orthodox decision theory in which dominance — or something like it — plays the central role. Note that both kinds of approach put dominance in a position that is different from the one it occupies in classical decision theory, where it is simply a *consequence* of utility maximization (if one simple lottery dominates another, then the expected utility of the dominant lottery will be higher).

In the remainder of this paper I wish to pursue the *second* problem: the violation of evaluative compositionality. I shall argue that evaluative compositionality is not in fact a requirement of rationality. This is not to say, however, that anything goes when it comes to valuing lotteries. Rather, evaluative compositionality fails for a specific, principled reason and hence fails in a particular, controlled way. So I will not be defending the chocolate ice cream response: the view that we can value the Pasadena game *however we like*. I will, however, be defending the view that a rational agent is not required to value the game in one particular way (i.e. even once she has valued its outcomes in particular ways). One important upshot of the approach will be that it allows us to *retain* dominance. Furthermore, dominance will have the same role that it has in the classical theory: it will turn out to be a *consequence* of the fundamental constraints on rational decision making.

6. Rationally negligible probabilities

Normative theories of practical activities must make room for *tolerances*: allowable variations from the specified norm. In other words, they should not require arbitrary precision (i.e. arbitrarily high or fine precision): it should not be that any variation from the norm, *no matter how small*, is disallowed. Rather, there must be an allowable range such that variation from the norm within that range is ignored: any value within that range is just as good as the norm itself. For example, an engineering plan — which can be seen as a normative theory that constrains how a certain activity (constructing such and such) should be performed — will say (for example) that a certain component should have a certain diameter. However, it will also specify a tolerance: a range such that discrepancies from the norm within that range can be ignored. In order to be acceptable by the lights of the normative theory, the diameter needs to deviate from the norm by less than some specific finite amount (the tolerance for this measurement). The key point is that it is *not* the case that the deviation must be less than *every* positive amount: for that would amount to the requirement of *zero* deviation from the norm — that is, the requirement of infinite precision.

Decision making is a practical activity. Decision theory is the normative theory of this practical activity: the theory that tells us how this activity should be performed — how it ought to be done. So decision theory must make room for tolerances. More specifically, decision theory specifies that decision makers should ignore (i.e. not factor into their decision making) outcomes with *zero* probability. The way that it specifies this is by having such outcomes make no difference to the output when the machinery of decision theory — the expected utility calculation — is applied to them. The point now is that there must be a tolerance on this norm. From the point of view of the normative theory, ignoring outcomes with probability less than some finite threshold should be just as good as ignoring outcomes whose probability is *precisely* zero. That is, the requirement on ideal decision makers cannot be that they take into account *arbitrarily small* probabilities: that they ignore only outcomes whose probabilities are less than *every* positive threshold. Infinite precision cannot be required: rather, in any given context, there must be some finite tolerance — some positive threshold such that ignoring all outcomes whose probabilities lie below this threshold counts as satisfying the norm.

The foregoing idea can be captured in the following principle:

> **Rationally negligible probabilities (RNP):**
> For any lottery featuring in any decision problem faced by any agent, there is an $\epsilon > 0$ such that the agent need not consider outcomes of that lottery of probability less than ϵ in coming to a fully rational decision.

(RNP) asserts the existence of a positive number: ϵ. This is the tolerance: the amount of deviation from the norm (i.e. zero) which is permitted by the part of the normative theory which tells decision makers to ignore outcomes whose probability is zero.

Some clarificatory remarks are in order. First, note that in (RNP) the existential quantifier is inside the scope of the universal quantifiers. So we are *not* saying that there is some probability threshold such that no-one need ever — in any decision problem — consider outcomes — of any lottery — whose probability lies below this threshold. The latter idea is familiar from the literature on risk assessment, where it appears as the claim that 'society ought to ignore very small risks' (Shrader-Frechette 1985, p. 431). But (RNP) makes a quite different claim: that for any lottery featuring in any decision problem facing any agent, there is some ϵ such that the agent is rationally permitted to ignore all outcomes of that lottery whose probability is less than ϵ for purposes of making the required decision (i.e. it is consistent with her being fully rational that she ignore these outcomes). For another decision problem and/or another lottery and/or another decision maker, it might be a different ϵ. So the claim is simply that no decision problem requires any rational agent to consider *arbitrarily small* probabilities — to consider smaller and smaller probabilities *ad infinitum*. This is *not* to say that there is some probability so small that no decision maker need ever consider it. One way to see the difference between (RNP) and the claim that we may ignore small risks is to note that (RNP) exerts no constraint at all on finite lotteries: the condition is automatically met by setting ϵ less than the probability of the (equal-) least likely outcome. The proposal always to ignore very small risks, on the other hand, impacts finite lotteries: if the lottery has outcomes whose probabilities lie below the (fixed, universal) threshold, then they should be ignored.

Second, it is important to be clear about the rationale behind (RNP), and behind the more general idea that normative theories of practical activities must allow for tolerances. The rationale is *not* that tolerances must be allowed because in practice we cannot achieve infinite precision. For a start, it does not seem to be true that a machinist *cannot* make a part which is *precisely x* nanometres in diameter (or whatever). Whatever part she makes has some precise diameter — and it is possible that this diameter is *x* nanometres. In any case, the crucial point is that in practical contexts infinite precision never *matters*. So the point is not that arbitrary precision is unattainable: it is that no-one should *require* it. *That* is why a normative theory must not demand infinite precision. The point is particularly clear in the case of decision theory — for in many cases in decision theory, arbitrary precision *is* possible. For example, one *can* factor in *every* possible outcome of the St Petersburg game, even though there are infinitely many possible outcomes whose probabilities go arbitrarily low. One does so when one calculates the expected value of this game. This involves summing an infinite series — but one can (in this case) do that quite easily, using high-school mathematics. So the point is *not* that factoring in *every* outcome, no matter how small its probability, is practically impossible — that it involves performing a supertask, say.

Rather, the idea behind (RNP) is that in any actual context in which a decision is to be made, one never *needs* to be infinitely precise in this way — that it never *matters*. There is (for each decision problem, each lottery therein, and each agent) some threshold such that the agent would not be *irrational* if she simply ignored outcomes whose probabilities lie below that threshold. Hence decision theory, qua theory of ideally rational decision making, must not mandate that she factor in outcomes of *arbitrarily low* probability: that is, that she consider smaller and smaller probabilities *ad infinitum*.

If, at this point, someone is inclined to say that outcomes of *any* positive probability, no matter how small, *always* matter — that they must always be factored into a decision — then we should remind him of the utter vastness of infinity. Suppose that a round of the St Petersburg game is offered to the highest bidder, and we wish to decide how much to bid. We have factored in outcomes down to probability 10^{-n}, where n is the distance from the earth to the moon, measured in nanometres. Not enough, says our man, to make a fully rational decision. So we increase n to the distance from the earth to the sun, measured in femtometres. Not enough, says our man. So we increase n to the time, measured in yoctoseconds, taken for Nelson's column to be worn down to the ground, were it to be protected from all erosive forces save for those generated by a single pigeon alighting on it on the first day of each new century, starting on 1 January 2100. Still not enough! In fact, from our man's point of view, our effort is as pathetically inadequate as that of someone who considers outcomes down to probability $\frac{1}{4}$: both of us have an *infinite* number of further outcomes to consider. And no matter how many more outcomes we factor in — provided we factor in only finitely many — our man will say exactly the same thing. He will say that we still have exactly as far to go as we had at the beginning — infinitely far — and that we are *irrational* if we do not go all the way and consider *all* the remaining outcomes. But this seems to be demanding too much of rational agents. Not too much *effort* — we have already seen that calculating the expected value of the St Petersburg gamble is straightforward (it certainly does not involve performing a supertask) — just too much *precision*. It seems to demand far more precision than could possibly *matter*, for purposes of making a practical decision.

Note that when I say that beyond a certain finite point, further precision does not *matter*, I do not mean that factoring in further outcomes makes no *difference* to the value one calculates for the gamble. It might make a big difference. For example, suppose we have a gamble based on the same coin-tossing set-up as the St Petersburg gamble, but where the payoff for Heads on toss n increases much faster with n than it does in the St Petersburg case (e.g. rather than being 2^n, it is $2^{(2^n)}$ — or $2^{(2^{(2^n)})}$, …). Rather, what I mean is that beyond a certain point, whether or not one factors in further, more improbable outcomes makes no difference to the *acceptability* of the result for practical purposes — for making the decision at hand. That is, while it may make a difference to one's decision — to what one decides to do — it cannot (beyond a certain point) make a difference to the *rationality* of one's decision. Compare the engineering case. Lowering the tolerance on a particular measurement will, in general, lead to a *difference* in the finished product — in particular, to a difference in the diameter of the component in question. Beyond a certain point, however, it will not make any difference to the *acceptability* of the finished product for the purposes at hand: it will not make it any *better*. That is why the *normative* theory — the engineering plan — specifies *finite* tolerances. Specifying zero tolerances — infinite precision — would change the finished product, but would not improve it. Similarly, the idea in the case of decision theory is that beyond a certain point, factoring in outcomes of lower and lower probability ad infinitum does not make one's decision any *better*, any more rational. Specifying zero tolerance — infinite precision — on the norm that one ignore outcomes of probability zero will, in general, lead to *different* decisions being made — but (the idea goes) they will not be any more *rational* than those made by someone operating with some positive tolerance.

Third, I am not claiming that the view just outlined is intuitively *compelling*. I am claiming only that it is prima facie plausible: it is prima facie plausible that for any decision problem, lottery, and agent, there is *some* threshold such that the agent would not be *irrational* if she ignored outcomes of that lottery whose probabilities lie below that threshold for purposes of making that decision; it is prima facie plausible that *infinite precision* is not required in decision making any more than it is required in any other practical activity. (If one of the ignored outcomes ended up occurring, we would call the agent astronomically *unlucky* — not *irrational*.) This claim of prima facie plausibility is all I need in the context of this paper. Recall the dialectic. The case for evaluative compositionality depended on a principle — (URP) — for which the positive case was simply that it is prima facie plausible. We shall see that adopting (RNP) leads to a view which violates evaluative compositionality. But the debate between (URP) and (RNP) is not to be *decided* by a wrestle of intuitions.

Their prima facie plausibility gets *both* views onto the table. We then decide between them on the basis of consideration of the overall theoretical landscape. I shall argue (in Sect. 10) that adopting (RNP)—and consequently denying evaluative compositionality—leads to the most satisfying resolution of the problems posed by the Pasadena gamble and its relatives.

Before moving on, there is a question we should briefly discuss. It has been suggested to me on several occasions that if there is to be a tolerance on the norm 'ignore outcomes whose probability is zero', then there should also be a tolerance on the norm 'ignore outcomes whose utility is zero'. The problem with this suggestion is that the latter is *not* a norm of classical decision theory: it cannot be, because the notion of an outcome having 'zero utility' makes no sense in that theory. Of course it is true that, in the expected utility calculation, outcomes whose utility value is zero make no difference to the final result (just like outcomes whose probability is zero). But recall that utility is measured on an interval scale: utility values are unique only up to a positive linear transformation. An outcome whose utility value is zero according to one legitimate assignment of utilities will have a non-zero utility according to another, equally legitimate assignment. Only those statements which are stable in truth-value across all legitimate utility assignments are regarded as meaningful: so while it is meaningful to say that one outcome has *greater* utility than another, it is not meaningful to say that an outcome has *zero* utility.

An anonymous referee agreed with the point just made, but suggested that 'a much better tolerance principle [for utilities] is available … one might say that it is rationally permissible to ignore sufficiently small differences of utility. That is, substitution of an outcome with one utility for an outcome with extremely close utility should not affect the valuation of the [gamble].' Again, however, this proposal will not work: for the notion of two utilities being 'extremely close' makes no sense, given that utility is measured on an interval scale, and assuming that two utilities x and y are to be 'extremely close' if the difference between them is less than some margin c, namely $|x - y| < c$. For however small c is, there will always be a positive linear transformation f of the utilities such that $|f(x) - f(y)| > c$.

A different referee made a related suggestion: 'A far more natural understanding of the tolerance norm would be *tolerance about the exact values of expected utilities*. That is, rather than identifying the value of a gamble with a precise number, its expected utility, one could identify it with an interval around the expected utility. Then, when there is a small difference between the expected utilities of two gambles, tolerance might bid one to treat them as on a par.' For reasons similar to those just discussed, this proposal too will not work: in general, the property of two expected utilities being 'very close' (i.e. the difference between them is less than c) is not preserved under positive linear transformations of the utilities.

In general, the point is this. There is a norm of decision theory which says to ignore outcomes whose probability is zero. Because this norm mentions a specific probability value (zero), it is the kind of norm where it makes sense to impose a tolerance: zero plus or minus ϵ (which becomes zero plus ϵ, given that probabilities are all between 0 and 1).[16] When we move from probabilities to utilities, and expected utilities, the situation is different. Decision theory does not have a norm which says to treat outcomes with utility k—or gambles with expected utility k, or outcomes whose utilities differ by k or less, or gambles whose expected utilities differ by k or less—in such and such a way. It cannot have such norms, for the reasons discussed: these norms would not be meaningful, because utilities are measured not absolutely, but on an interval scale. And the norms that it does have that relate to utilities and expected utilities—pick the option with the higher utility, choose the gamble with the greater expected utility—are not the kinds where it makes sense to impose tolerances. A tolerance, in the specific sense under discussion in this paper, is an allowance that a specific number (given in the original statement of the norm) can be treated as an interval (i.e. any value in the interval is as good as that single value, from the point of view of the norm). Where there are no specific numbers—only comparisons (greater than, less than, equal to)—there can be no tolerances in this sense. Of course, one might think that even in these cases there should be a less specific kind of tolerance—that is, a general permissiveness of departures from the stated norm. But when we think about it, this is not plausible at all. Compare an engineering plan which specifies simply that one part should be *longer* than another. This norm is *already* highly permissive (i.e. *many* specific outcomes are compatible with it): there is, in general, no reason why it should be loosened to 'longer—or exactly the same' or 'longer—or exactly the same—or even a bit shorter'.

[16] This is not to say we *should* impose a tolerance: just that it would *make sense* to do so. I have already said why I think we should impose one: the reason was not simply 'because we can'.

7. Truncate and maximize

The way that a decision maker ignores — does not consider — an outcome when coming to a decision is by treating it as having probability zero (for as we have mentioned, outcomes with zero probability pass through the decision-theoretic machinery — in particular, the expected utility calculation — without making any impact on the output). (RNP) says that for any lottery featuring in any decision problem faced by any agent, there is an $\epsilon > 0$ such that the agent need not consider outcomes of that lottery of probability less than ϵ in coming to a fully rational decision. That is, the decision maker may treat all outcomes whose probability is less than ϵ as having probability zero.

What does this mean in practice? What does it mean to 'treat all outcomes whose probability is less than ϵ as having probability zero'? Well, it does *not* mean coming to believe that the probability of these outcomes really is zero. That would be *epistemically* irrational. (It would be analogous to an engineer believing that the diameter of the component he has manufactured *really is* 100 millimetres, when in fact it is not, but is within the specified tolerance of this norm.) Rather, what it means is *treating* the probabilities as zero *for purposes of making the decision at hand* — and what *this* means is as follows. Suppose a decision problem involves a lottery L_1 which assigns infinitely many non-zero probabilities. To ignore the outcomes of L_1 which have probability less than ϵ is to treat L_1 as interchangeable (in the context of this decision problem) with a lottery L_2 which really does assign probability zero to these outcomes. So the agent runs through her decision-making process (whatever that is) with L_2 in place of L_1, and then takes the result of this process — which is a verdict about the substitute gamble L_2 — and applies it to the original gamble L_1.

We need to say more about the nature of the substitute gamble L_2. It assigns probability zero to outcomes to which L_1 assigns probability less than ϵ. This cannot, however, be the *only* difference between L_1 and L_2. L_2 cannot assign the very same probabilities as L_1 to the remaining outcomes — for then the sum of all the probabilities assigned to outcomes by L_2 would be less than 1, and so L_2 would not be a well-defined gamble. L_2 is therefore to be the closest gamble to L_1 which assigns probability zero to outcomes to which L_1 assigns probability less than ϵ — where 'closest' is cashed out this way: the probabilities assigned by L_2 are obtained from those assigned by L_1 by conditionalizing on the supposition that some outcome of probability greater than or equal to ϵ occurs. We can make this precise as follows.

Where L is any lottery, let $\hat{L}$ be the (equal-) highest probability assigned to any outcome by L.[17] For any real number x, let L^x be the set of outcomes of L which are assigned probabilities greater than or equal to x; that is, $L^x = \{o \in O_L : L(o) \geq x\}$. Now where L is any lottery and ϵ is any real number with $0 < \epsilon \leq \hat{L}$, we define the *ε-truncation* L/ϵ of L to be that lottery which has the same set of possible outcomes as L and assigns probabilities to these outcomes as follows:[18]

$$L/\epsilon(o) = L(o/L^\epsilon) = \frac{L(o \cap L^\epsilon)}{L(L^\epsilon)}$$

The ϵ-truncation L/ϵ will be our substitute gamble for L. [19]

[17] There must be such a number: the probabilities assigned by L cannot increase indefinitely without reaching a maximal value, for if they did then the sum of all of them would exceed 1.

[18] The middle term, $L(o/L^\epsilon)$, represents the conditional probability that outcome o occurs, given that some outcome in the set L^ϵ occurs; the right-hand term, $\frac{L(o \cap L^\epsilon)}{L(L^\epsilon)}$, represents the standard definition of this conditional probability as a ratio of unconditional probabilities. Because $\epsilon \leq \hat{L}$, $L(L^\epsilon) \neq 0$ and so these probabilities are well defined. Recall (n. 15) that $L(o)$ is shorthand for $L(\{o\})$; when the shorthand is unpacked, $L(o \cap L^\epsilon)$ (which superficially makes no sense, because it talks of the intersection of an element of the sample space with a subset of the sample space) becomes $L(\{o\} \cap L^\epsilon)$ (which talks of the intersection of two subsets of the sample space and so makes sense).

[19] There are other possible choices for the substitute gamble. In an earlier version of the paper, I took the substitute to be the gamble obtained by setting all probabilities less than ϵ to 0, and then spreading the leftover probability (i.e. the sum of all the probabilities less than ϵ) evenly across the outcomes with non-zero probability:

$$L\backslash\epsilon(o) = \begin{cases} 0 & \text{if } L(o) < \epsilon \\ L(o) + \frac{1}{|L^\epsilon|}\left[1 - \sum\limits_{x \in L^\epsilon} L(x)\right] & \text{if } L(o) \geq \epsilon \end{cases}$$

(Note that this way of truncating L at ϵ is symbolized using a backslash, i.e. $L\backslash\epsilon$, to distinguish it from the way presented in the text, which is symbolized using a forward slash, i.e. L/ϵ. Note also that $|L^\epsilon|$ is the cardinality of L^ϵ, i.e. the number of outcomes of L which are assigned probabilities greater than or equal to ϵ. Because $\epsilon \leq \hat{L}$, $|L^\epsilon|$ is non-zero; by reasoning similar to that in n. 17, $|L^\epsilon|$ is finite.) The choice of substitute gamble in the text (i.e. L/ϵ) has the advantage that it maintains ratios of (non-truncated) probabilities: if L says that outcome x is n times as probable as outcome y, then so does L/ϵ (provided it assigns x and y non-zero probabilities). The choice of substitute gamble in this footnote (i.e. $L\backslash\epsilon$) has the advantage that it minimizes the maximum shift of (non-truncated) probabilities: setting aside outcomes whose probabilities are shifted to zero in moving from L to the substitute gamble, any other choice of substitute gamble will involve shifting the probability of *some* outcome more than the probability of *any* outcome is shifted in moving from L to $L\backslash\epsilon$. Perhaps there are also other reasonable choices of substitute gamble—that is, other ways of cashing out 'closest' when we say that the substitute gamble L_2 is to be the closest gamble to the original gamble L_1 which assigns probability zero to outcomes to which L_1 assigns probability less than ϵ. In any case, the existence of multiple options here does not affect the conclusions of this paper. I shall argue that purely rational considerations do not fix *where* to truncate a given infinite gamble—i.e. they do not fix a uniquely correct choice of ϵ—and so evaluative compositionality fails. If, even given a choice of ϵ, purely rational

We are now in a position to propose a method for handling decision problems in light of (RNP). For any agent facing any decision problem and any lottery therein, (RNP) says that there is some $\epsilon > 0$ such that the decision maker need not consider outcomes of that lottery of probability less than ϵ in coming to a fully rational decision. Call a probability *rationally negligible* (with respect to a particular decision problem and a particular lottery which features in it and a particular decision maker) if the decision maker need not (in the context of that decision problem) consider outcomes (of that lottery) which have that probability, in order to make a fully rational decision. The decision method which flows naturally from (RNP) is this:

> **Truncation:** When you are faced with a decision problem *d* that involves a lottery *L*, pick a probability ϵ that is rationally negligible with respect to *d*, *L*, and yourself — by (RNP) we know there is at least one such — and then set your value for *L* to your value for L/ϵ.

The beauty of Truncation as a decision method is twofold:

First, it makes *no* impact on simple lotteries. (RNP) is *trivially* true for such lotteries. (RNP) says that agents need not consider further and further possible outcomes whose probabilities get smaller and smaller *without end*. In a simple gamble, an agent is never required to do this: and so (RNP) is automatically satisfied. Of course (RNP) is *compatible* with the view that it is rational to ignore some outcomes of a simple lottery which have non-zero probability — that is, to treat them as having probability zero: but it does not *imply* this view.[20] All that follows from (RNP) is that in any lottery — simple or infinite — there is *some* positive ϵ such that a rational decision maker may ignore outcomes of that lottery of probability less than ϵ. In a *simple* lottery, this is *already* true: a rational decision maker may ignore outcomes of probability zero, and in a simple lottery, there will be a *positive* ε such that every outcome which has a probability less than ϵ has probability zero — and hence may be ignored. Thus, (RNP) exerts no constraint on simple lotteries: what it implies about them is already true of them in any case.

---

considerations furthermore do not fix *how* to truncate *L* — e.g. they do not fix that one should consider L/ϵ rather than $L\backslash\epsilon$ — this will be grist for my mill. For further discussion of this point, see n. 21.

[20] To be clear: the present paper is not committed one way or the other with regards to this view.

We can make the same point in a slightly different way. A rational decision maker ignores outcomes which have probability zero. Where a simple lottery is involved, there will be a lowest positive probability assigned to any outcome by this lottery. Pick an ϵ less than this lowest probability, but greater than zero. Then of course a rational agent is permitted to ignore outcomes whose probability is less than ϵ when coming to a decision with regards to this lottery: for those are just the outcomes whose probability is zero. Now if one truncates the simple lottery at ϵ, one ends up with that lottery itself. So the truncation method yields the advice to value the simple lottery at the value one places on that lottery itself. What value is *that*? Well, it is a *simple* lottery — so, given classical decision theory, one values it at its expected value.

Second, the truncation method reduces *all* decision problems to problems involving only simple gambles. For note that whether or not L is simple, L/ϵ must be simple. So when we apply the truncation method, we value L as we value L/ϵ — and because L/ϵ is simple, we value it at its expected value (assuming orthodox decision theory, in which the expected utility property holds for simple lotteries: the value of a simple gamble is its expected value, which is always defined). So Truncation brings all decision problems — whether or not they involve non-simple gambles — within the scope of orthodox decision theory, which handles simple gambles nicely.

Nover and Hájek (2004, pp. 246–7) consider, as a response to the Pasadena game, the proposal to 'restrict decision theory to finite state spaces'. They argue against any such restriction. I *agree* that such a restriction is unwarranted — and it is not what I am proposing here. Infinite gambles can be well defined and when they are they *can* be assigned values. The *method* I propose for assigning a value to an infinite gamble L is to pick a rationally negligible ϵ and then find a gamble which is like L except that it assigns probability zero to every outcome to which L assigns probability less than ϵ. The latter gamble will always be simple — hence orthodox decision theory tells us how to value it (namely, at its expected value). We then place the *same* value on the original infinite gamble L. So infinite gambles are not in any way ruled out as ill-formed or unvaluable. Rather, they are assigned values *via* finding the values of associated simple gambles.

A referee suggested that there is something arbitrary about the truncation method: that, given an infinite gamble, there are arbitrarily many ways to eliminate low-probability outcomes and leave behind a simple gamble. In particular, rather than eliminating each outcome whose probability is less than ϵ, one could eliminate an infinite set of outcomes whose total probability is ϵ — and of course there are, in general, infinitely many such sets (whereas there is a unique set of all outcomes whose probabilities are each less than ϵ). First of all, it is important to be clear about the objection here. My proposal is that we evaluate infinite gambles in three steps: (1) Pick an ϵ; (2) Zero out the probabilities less than ϵ; (3) Find a substitute gamble which is both well defined and has zeros in the right places. One could, in theory, direct a charge of arbitrariness at any of the three steps. Step 1: It is part of my view — to be discussed further below — that purely rational considerations do not mandate a particular choice of ϵ. Step 3: The issue of whether, given ϵ, there are multiple acceptable choices of substitute gamble was discussed in footnote 19. Step 2: This is the locus of the present objection, which is that it is arbitrary to zero out each probability less than ϵ, rather than picking some set of outcomes whose total probability is ϵ and then zeroing out the probability of each outcome in this set. So, to respond to this objection: I agree that if our goal were simply to turn an infinite lottery into a simple one by lopping off low probability outcomes, then there would be nothing to favour my approach over the alternative(s) just considered. But recall the dialectic. My approach was not *motivated* by the fact that it turns infinite gambles into simple ones. Rather, this was an advantage which *results from* the approach. The motivation came from quite general considerations (discussed in Sect. 6) concerning normative theories of practical activities, and more specifically from the idea that such theories must make room for tolerances. In particular, my proposal arose from placing a tolerance (ϵ) on the norm of classical decision theory which directs decision makers to ignore outcomes whose probability is zero. The alternative proposal now under consideration, by contrast, is not motivated by any such general considerations: it is suggested purely as a means of turning infinite gambles into simple ones. Thus, what distinguishes my proposal from the myriad alternatives which lead to the same destination (i.e. to the transformation of infinite gambles into simple ones) is that my proposal is motivated 'from behind' by quite general considerations, while the alternatives are purely goal-driven and ad hoc. So step 2 of my proposal is not an arbitrary choice from a multitude of possibilities: it is a well-motivated choice.

8. Evaluative compositionality

At this point it will be useful to distinguish two theses. The *weak thesis* is that rational agents *may* apply the truncation method: faced with a decision problem *d* that involves a lottery *L*, an agent is rationally *permitted* to pick a probability ϵ that is rationally negligible with respect to *d*, *L*, and herself and then set her value for *L* to her value for L/ϵ. The *strong thesis* is that rational agents *must* apply the truncation method: faced with a decision problem *d* that involves a lottery *L*, an agent is rationally *mandated* to pick a probability ϵ that is rationally negligible with respect to *d*, *L*, and herself and then set her value for *L* to her value for L/ϵ. Note that the weak thesis *permits* valuing the St Petersburg gamble at some finite value (i.e. the value of one of its truncations), while the strong thesis *demands* this: it *forbids* valuing the St Petersburg gamble at its expected value.

Note that the strong thesis, in itself, does not specify a truncation point. It says that, faced with a decision problem *d* that involves a lottery *L*, an agent is rationally mandated to pick a probability ϵ that is rationally negligible with respect to *d*, *L*, and herself and then set her value for *L* to her value for L/ϵ. The specification of ϵ is left to the agent: beyond specifying that the chosen ϵ must be rationally negligible (with respect to the agent and the given *d* and *L*), the strong thesis, in itself, does not say what ϵ should be.

It is the weak thesis that is supported by the considerations put forward so far, and it is the weak thesis to which this paper is committed. Nothing that we have said so far supports the strong thesis that rational agents *must* apply the truncation method. When applied to infinite lotteries, (RNP) asserts the existence of rationally *negligible* probabilities: positive probabilities that *may* be ignored. The strong thesis, by contrast, asserts that there are probabilities *non grata*: probabilities that *must* be ignored. (As noted in the previous paragraph, however, the strong thesis, in itself, does not specify *which* probabilities must be ignored, in any case: only that *some* must be.)[21]

The weak thesis leads to a violation of evaluative compositionality. Consider, by way of example, the St Petersburg gamble (henceforth referred to as *S*). Let us suppose that we are faced with a particular decision problem involving *S*: say, whether or not to pay \$100 to play *S*. (RNP) asserts the existence of a threshold (ϵ) such that one

[21] Recall n. 19. We can now see why it makes little difference to the arguments of this paper if, given a choice of where to truncate a lottery *L* (i.e. a choice of ϵ), there is more than one acceptable substitute gamble that may be considered in place of *L* (e.g. L/ϵ or $L\backslash\epsilon$). The arguments to be presented below turn on the weak thesis: that one *may* apply Truncation. What *else* one may do, if one does *not* apply Truncation — e.g. whether one may apply a different version of Truncation, which features $L\backslash\epsilon$ in place of L/ϵ — makes no difference to these arguments.

may ignore outcomes of S whose probability is below the threshold when coming to a decision with regards to this problem. (RNP) does not say what the threshold is in this (or any other) case—but let us suppose, for the sake of argument and without loss of generality, that a rational agent may in this case ignore outcomes of S whose probability is less than 0.01.[22] So a rational agent may value S as she values $S/0.01$. The latter is a simple gamble, so she values it at its expected value:[23]

$$\frac{64}{126}\cdot\$2+\frac{64}{252}\cdot\$4+\frac{64}{504}\cdot\$8+\frac{64}{1008}\cdot\$16+\frac{64}{2016}\cdot\$32+\frac{64}{4032}\cdot\$64=\$6.10$$

But any probability that is *smaller* than a rationally negligible probability is itself rationally negligible: by the way (RNP) is stated, one may ignore *every* outcome whose probability is *less* than ϵ. So given that our agent may truncate at 0.01, she may also truncate at (e.g.) 0.001. But if she does *that*, her value for S will be higher: approximately \$9.[24] Hence evaluative compositionality fails. Even once the agent's values for the outcomes of the St Petersburg gamble are fixed—and the probabilities of these outcomes are given by the gamble—still no unique value for the gamble itself is mandated by purely rational considerations.

In section 5 we argued that (IFF) and (URP) together imply that evaluative compositionality is a requirement of rationality. It is (URP) that we are now denying—not (IFF).[25] The idea is not that it takes some further factor, in addition to the values and probabilities of the outcomes, in order to fix the unique rational value of some gambles. Rather, the view is that there simply is no unique rational value: there is a range of rationally admissible values. Evaluative compositionality fails for infinite gambles—given (RNP)—not because rational agents must take into account *more* than the values and probabilities of the outcomes, but because they may take into account *less* (they may

[22] I am *not* suggesting that it really would be rational to ignore outcomes of S whose probability is below 0.01. Picking such a high value for ϵ simply makes the example easy to present.

[23] Presented in decreasing order, the probabilities of the possible outcomes of S are $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{32}, \frac{1}{64}, \frac{1}{128}, \ldots$. The first of these to fall below 0.01 is the seventh, $\frac{1}{128}$. The sum of the first six is $\frac{63}{64}$. The probabilities shown in the expected utility calculation—i.e. the non-zero probabilities assigned by $S/0.01$—are the results of dividing each of the first six probabilities assigned by S by $\frac{63}{64}$ (i.e. $\frac{1}{2}\div\frac{63}{64}=\frac{64}{126}, \ldots, \frac{1}{64}\div\frac{63}{64}=\frac{64}{4032}$).

[24] S assigns nine probabilities above 0.001, namely $\frac{1}{2}$ through $\frac{1}{512}$.

[25] We are not forced to accept (IFF): the point here is simply that we need not deny it—whereas we do need to deny (URP).

ignore some outcomes altogether). However, (RNP) says simply that a rational agent need not consider outcomes of *arbitrarily low* probability: she need not consider smaller and smaller probabilities *ad infinitum*. Beyond that, it places no constraints on where a rational agent may 'draw the line'. Different agents may draw the line in different places, leading to different values for the gamble — but provided the probabilities ignored are indeed rationally negligible, all of these agents are proceeding in a rationally admissible way.[26]

To obtain a *unique* value for each gamble, in the present framework, we would need to adopt the strong thesis *and* specify, for each gamble and each possible assignment of values to its outcomes, the particular ϵ at which agents who assign those values to the outcomes must truncate the gamble. But apart from the fact that the strong thesis is not supported by the considerations presented so far, it seems clear that any such choice of ϵ could only be *conventional*. That is, it would not be dictated by pure rationality. Therefore, while we could determine a *unique* price for each gamble in this way, it would not be a unique *rational* price (and evaluative compositionality would not be restored).

Indeed, if we assume (IFF), then it follows that any such choice of ϵ could not be dictated by pure rationality. By fixing an ϵ, we can rule out some values of a gamble (i.e. values got by truncating at a point other than ϵ). If these values were ruled out as *irrational*, then the choice of ϵ would be a further factor — beyond the values and probabilities of the outcomes — affecting whether it is rationally permissible for an agent to place a certain value on a certain gamble. (IFF) asserts that there are no such further factors. Thus, if we fix on an ϵ and thereby rule out certain values for a gamble, it cannot be that we are ruling them out as *irrational* (i.e. as being values that are not rationally admissible): we can only be ruling them out as something like *unconventional*.

Another thing that follows, if we assume (IFF), is that the number of degrees of freedom in (RNP) is reduced. For example, it is compatible with (RNP) alone that there be two agents *a* and *b* who assign the same values to the outcomes of lottery *L*, such that it is admissible for *a* to truncate *L* at ϵ, but not admissible for *b* (i.e. *b* may truncate, but only beyond a threshold lower than ϵ). Given (IFF), this possibility is ruled

[26] (RNP) says that rationally negligible probabilities must *exist*; it says nothing about how to *identify* them. The question arises how we determine, in a particular case, whether a given probability really is rationally negligible. I think that this is an important question — but it is not one that I try to answer here: nothing that I say in this paper turns on knowing, of any particular probability, whether it is rationally negligible (in a given context).

out: given that *a* and *b* assign the same values to the outcomes of *L*, if it is rationally permissible for *a* to place a certain value on *L*, then it is rationally permissible for *b* to place that value on *L* (and vice versa).

9. Constancy and Consistency

Suppose that at some time — say, 11am — you are offered a St Petersburg gamble *S*. In accordance with the truncation method, you value it at the value you place on S/ϵ, for some rationally negligible ϵ. Now suppose that at a second time — say, 11.30am — you are offered a second St Petersburg gamble S'. The truncation method directs you to pick a rationally negligible ϵ' and value S' at the value you place on S'/ϵ'. The *only* constraint — for all that we have said so far — is that ϵ' be rationally negligible (and (RNP) ensures that at least one such ϵ' exists). In particular, there is no requirement that $\epsilon' = \epsilon$: no requirement that you truncate S' at the point at which you truncated *S*. This holds no matter how similar *S* and S' are: indeed, they might both be based on the *very same* set of outcomes (say, some series of tosses of a particular coin to be carried out from 12 p.m. until the coin lands Heads for the first time). It also holds if you are offered *S* and S' at the *same* time: for example, suppose that a single decision problem involves a choice between paying \$*x* for *S* or \$$x'$ for S'. In accordance with the truncation method, you value *S* at the value you place on S/ϵ, for some rationally negligible ϵ, and you value S' at the value you place on S'/ϵ', for some rationally negligible ϵ'. The *only* constraint is that ϵ and ϵ' be rationally negligible: there is no requirement that $\epsilon' = \epsilon$.

Should we add an additional requirement to the truncation method which mandates truncating different lotteries at the *same* point, other things being equal? More specifically, let us consider four possible principles.[27]

Weak Consistency: If a single decision problem *d* involves two lotteries L_1 and L_2 over the same set of outcomes, then if one

[27] A note on the naming of these principles, to aid in their comprehension: principles governing how an agent should evaluate two gambles within the context of a single decision problem are labelled Consistency principles. Principles governing how an agent should evaluate two gambles, one in the context of one decision and the other in the context of a different decision, are labelled Constancy principles. We shall say more below about what it takes for two lotteries to be relevantly similar, but I assume that two lotteries which have the very same possible outcomes are relevantly similar. Hence, in each case — i.e. Constancy and Consistency — the Weak principle is supposed to be a special case of the Strong principle.

truncates L_1 at ϵ for purposes of addressing d, one must also truncate L_2 at ϵ for purposes of addressing d.[28]

Strong Consistency: If a single decision problem d involves two relevantly similar lotteries L_1 and L_2, then if one truncates L_1 at ϵ for purposes of addressing d, one must also truncate L_2 at ϵ for purposes of addressing d.

Weak Constancy: If two decision problems d_1 and d_2 involve (respectively) two lotteries L_1 and L_2 over the same set of outcomes, then if one truncates L_1 at ϵ for purposes of addressing d_1, one must truncate L_2 at ϵ for purposes of addressing d_2.

Strong Constancy: If two decision problems d_1 and d_2 involve (respectively) two relevantly similar lotteries L_1 and L_2, then if one truncates L_1 at ϵ for purposes of addressing d_1, one must truncate L_2 at ϵ for purposes of addressing d_2.

Before discussing the merits (or otherwise) of these principles, let us clarify their content by considering some examples. Suppose we have some coins c and c' which are indistinguishable (they were produced by the same machine at the same mint, one after the other). We also have two coin-tossing machines, t and t', which are likewise indistinguishable (they were produced at the same factory, one after the other). Each machine features a dial which can be used to set the time between successive tosses: from 5 seconds to 86,400 seconds (24 hours). Now consider the following experimental setups:

(1) Coin c will be tossed by machine t with its dial set to 5 seconds, beginning at 12 p.m., until the coin lands Heads.
(2) Coin c' will be tossed by machine t' with its dial set to 5 seconds, beginning at 12 p.m., until the coin lands Heads.
(3) After experiment 1 concludes, coin c will be tossed by machine t with its dial set to 5 seconds, until the coin lands Heads.

Suppose you are offered the choice between a St Petersburg game (at price \$$a$) to be decided by experiment 1, and a Pasadena game (at price \$$b$) to be decided by experiment 1. This is a single decision

[28] If we are allowing that an agent may apply Truncation, *or* a different version of Truncation which features $L\backslash\epsilon$ in place of L/ϵ (recall n. 19 and n. 21), then it is to be understood here that the agent must furthermore apply the *same* decision method to L_1 and L_2. Similar remarks apply to the three principles to follow.

problem, and the two lotteries involved are defined over the very same set of outcomes — so if we truncate one of them at ϵ, Weak Consistency demands truncating the other at the same point.

Suppose you are offered the choice between a St Petersburg game (at price \$*a*) to be decided by experiment 1, and a St Petersburg game (at price \$*b*) to be decided by experiment 2. This is a single decision problem, but the two lotteries involved are not defined over the very same set of outcomes — so Weak Consistency does not apply. Presumably the two lotteries are, however, relevantly similar — so if we truncate one of them at ϵ, Strong Consistency demands truncating the other at the same point. (The same applies if you are offered the choice between a St Petersburg game (at price \$*a*) to be decided by experiment 1, and a St Petersburg game (at price \$*b*) to be decided by experiment 3.)

Suppose that at 11 a.m. you are offered a choice between \$*a*, and a St Petersburg game to be decided by experiment 1. At 11.30 a.m. you are offered a choice between \$*b*, and a St Petersburg game to be decided by experiment 1. These are two different decision problems, so neither Consistency principle applies. The lotteries involved in the two decision problems are, however, defined over the very same set of outcomes, so if we truncate the first of them at ϵ (for purposes of addressing the 11 a.m. decision problem), Weak Constancy demands truncating the second of them at the same point (for purposes of addressing the 11.30 a.m. decision problem).

Suppose that at 11 a.m. you are offered a choice between \$*a*, and a St Petersburg game to be decided by experiment 1. At 11.30 a.m. you are offered a choice between \$*b*, and a St Petersburg game to be decided by experiment 2. These are two different decision problems, so neither Consistency principle applies. The two lotteries involved in the two decision problems are not defined over the very same set of outcomes, so Weak Constancy does not apply. Presumably the two lotteries are, however, relevantly similar — so if we truncate the first of them at ϵ (for purposes of addressing the 11 a.m. decision problem), Strong Constancy demands truncating the second of them at the same point (for purposes of addressing the 11.30 a.m. decision problem).

When are two lotteries relevantly similar? In other words, what exactly do the two Strong principles state? For example, consider a series of experimental setups 2.1, 2.2, 2.3, … , 2.86395 which differ from 2 above in the setting of the dial on machine t': in each setup, the setting is increased by one second, until it reaches the maximum. Presumably a lottery L_1 to be decided by Experiment 1 is relevantly similar to a lottery L_2 to be decided by experiment 2. But suppose L_2 is to be decided instead by some experiment in the above series. How far down the series do we have to go before L_1 and L_2 are no longer relevantly similar (or are they always relevantly similar)? This is a difficult question — and we could come up with other, equally difficult questions concerning the notion of relevant similarity. Fortunately we do not have to answer such questions here. We shall leave the two Strong principles vague — by leaving the notion of two lotteries being 'relevantly similar' vague. However, we *apply* the Strong principles only in cases where it is intuitively obvious that the lotteries involved are indeed relevantly similar by any reasonable measure.

It is time now to consider the pros and cons of the four principles. Note first that agents who violate the Weak principles are subject to Dutch book: in the case of Weak Consistency, a synchronic Dutch book; in the case of Weak Constancy, a diachronic Dutch book.[29] For example, suppose a bookie knows that at 11 a.m. you will truncate bets based on experiment 1 at $2^{-1{,}000{,}000}$ (i.e. you treat the probability of not getting Heads by the millionth toss as being zero) and at 11.30 a.m. you will truncate such bets at $2^{-1{,}000{,}000{,}000}$ (i.e. you treat the probability of not getting Heads by the billionth toss as being zero) — thereby violating Weak Constancy.[30] Then she can offer you at 11 a.m. a bet which pays out \$1 if Heads comes up first at toss n for $1 \leq n \leq 1{,}000{,}000$ and \$0 otherwise, and at 11.30 a.m. a bet which pays out \$1 if Heads comes up first at toss n for $1{,}000{,}000 < n \leq 1{,}000{,}000{,}000$ and \$0 otherwise. At 11 a.m. you will value the first bet at \$1 and at 11.30 a.m. you will value the second bet at some positive amount (for — unlike at 11 a.m. — you think the probability of first getting Heads at toss n for $1{,}000{,}000 < n \leq 1{,}000{,}000{,}000$ is positive). So you will be prepared to buy the first bet for \$1 and the second for some positive amount. But your combined return from both bets will be either \$1 (if the coin first lands Heads on toss n for $1 \leq n \leq 1{,}000{,}000{,}000$) or \$0 (if the coin first lands Heads on toss n for $n > 1{,}000{,}000{,}000$) — so you face a guaranteed loss, no matter what outcome eventuates. It is similarly

[29] A (synchronic) Dutch book is a package of bets, offered (at one time — as a package) at prices each of which the agent considers fair, such that buying all the bets in the package (at these prices) guarantees the agent a loss. A diachronic Dutch book — or Dutch strategy — is a sequence of bets, each offered at a price the agent considers fair (at the time it is offered), such that buying them all (at these prices) guarantees the agent a loss.

[30] The particular numbers do not matter: all that matters is that they are different.

easy to show that an agent who violates Weak Consistency is susceptible to a (synchronic) Dutch book.

However, it does not follow that the Weak principles are requirements of rationality. Susceptibility to Dutch book reveals the existence of what I shall call *bivaluation*: it reveals that the same outcome has been assigned multiple distinct evaluations.[31] Now, susceptibility to Dutch book is always a matter for practical concern (assuming one does not want to lose money) — but whether it is anything *more* than that is not something which the mere susceptibility itself can decide. The susceptibility *reveals* the existence of bivaluation: whether the bivaluation *matters* must be determined by other means. As Christensen (1991, p. 242) puts the point: ‘the inconsistency [i.e. bivaluation, in my terms] should not concern us at all unless the set of beliefs in question *should* be consistent. Moreover (and this is a crucial point), the question of whether the beliefs in a certain set should fit with one another has nothing to do with anyone's financial prospects. Vulnerability to the Dutch bookie, while it reveals an inconsistency in a certain set of beliefs, simply does not speak to this prior question at all.’

To appreciate this point, contrast two cases. The first concerns an agent whose degrees of belief do not conform to the probability axioms, and who is therefore susceptible to Dutch book. Here is what Ramsey famously wrote about this case:

> These are the laws of probability, which we have proved to be necessarily true of any consistent set of degrees of belief. Any definite set of degrees of belief which broke them would be inconsistent in the sense that it violated the laws of preference between options, such as that preferability is a transitive asymmetrical relation, and that if α is preferable to β, β for certain cannot be preferable to α if *p*, β if not-*p*. If anyone's mental condition violated these laws, his choice would depend on the precise form in which the options were offered him, which would be absurd. He could have a book made against him by a cunning better and would then stand to lose in any event. (Ramsey 1990a, p. 78)

Ramsey's point is that it is *clearly absurd* for one's choice to depend on the precise form in which the options are offered. Susceptibility to Dutch book *reveals* that one is bivaluating in this way (i.e. evaluating an option one way when presented thus and another way when

[31] As Armendt (1992, p. 218) puts it: ‘The idea underlying a Dutch Book argument is that an agent whose beliefs violate the recommended constraint is making the mistake of evaluating the same option in two or more different ways. Since these evaluations involve (according to Bayesians) dispositions to choose and act, the distinct evaluations could be exploited by a bettor (who realizes what the agent is doing)’.

presented so) — but it does not *make* this kind of bivaluation bad. That it *is* bad is established *independently* of the susceptibility to Dutch book. The second case is due to Christensen:

> Suppose that I am shopping with my wife. My credence in rain today is 25%. My wife … sets the probability of rain at 50%. I am approached by a bookie, who offers to bet me $1 to my $3 that it will rain … Given my credence, I regard this bet as fair [and] accept it … The bookie then approaches my wife, offering her a bet at $2 to $2, which he will win if it doesn't rain. Given her credence, she regards this bet as fair, and accepts it. The bookie has now assured himself of a $1 profit … my wife and I hold all our assets in common, so that not only has the bookie made a sure profit, but we have sustained a sure loss. (Christensen 1991, pp. 239–40)

The susceptibility to Dutch book of the husband and wife as a couple reveals the existence of bivaluation: one member of the couple places one probability on rain, the other member places a different probability. Whether or not such bivaluation is *bad*, however, must be settled by other means — and as Christensen notes, it obviously is *not* bad: 'Consistency in degrees of belief … is a rational ideal for individuals, not couples — even couples with joint checking accounts' (Christensen 1991, p. 240).

So, Dutch book reasoning does not, in itself, show that we should adopt the Weak principles as requirements on all rational agents. The agent who truncates different bets over the same set of outcomes at different points is susceptible to Dutch book. The susceptibility may alert us to the existence of bivaluation (treating an event as having zero probability, and as having positive probability) — although in this case it was obvious from the outset that truncating two gambles over the same outcomes at different points means treating one and the same outcome as having two different probabilities — but the question whether such bivaluation is irrational remains open. Let us then consider the four principles directly, and see whether the case for adopting each of them as requirements on all rational agents can be made on other grounds.

Let us start with Weak Consistency. Suppose that I am offered a choice between two gambles over the same set of outcomes: say, a Pasadena game P and a St Petersburg game S, each to be decided by experiment 1. Suppose it is rationally permissible to truncate P at ϵ. Then it is also rationally permissible to truncate P at $\delta < \epsilon$. Let us suppose that exactly the same is true of S. Still, it seems that it is *not* permissible to truncate P at ϵ and S at δ, for purposes of making the decision at hand. That would involve treating some outcome as having probability zero (in the context of *P*) and also, in the same breath as it were — that is, in the course of addressing a single decision problem — as having positive probability (in the context of *S*). That would involve just the sort of double-think — or divided-mind inconsistency as Armendt (1992, p. 219) calls it — that Ramsey was talking about in the passage quoted above: the sort where one's choice depends on the precise form in which the options are offered. This really does seem absurd (irrational, inconsistent). So, it seems, we should accept Weak Consistency as a requirement on all rational agents.

Given Weak Consistency, there is a rather strong intuitive pull to accept Strong Consistency as well. Given the perfect similarity between experiments 1 and 2, it would seem extremely odd for an agent to treat the outcome of Heads on toss *n* in experiment 1 as having positive probability, while — in the very same breath — treating the outcome of Heads on toss *n* in experiment 2 as having zero probability. However, it is not clear that such an agent could be accused of flat-out inconsistency — as could an agent who violates Weak Consistency — for there is no single outcome to which she is assigning both zero and non-zero probability. It seems that, rather than being *inconsistent*, the agent is being objectionably *arbitrary*. There is therefore a strong pull towards accepting Strong Consistency as a requirement on all rational agents: but the case is not as clear as it is for Weak Consistency.

Let us jump now to Strong Constancy. It seems that there is no good case for imposing this as a requirement of rationality. Consider a St Petersburg game *S*, to be decided by experiment 1, offered at 11 a.m., and a St Petersburg game S', to be decided by experiment 2, offered at 11.30 a.m. Would an agent who truncated *S* at ϵ and truncated S' at $\epsilon' \neq \epsilon$ thereby be irrational? I think not. The feeling that it is irrational to value *S* and S' differently seems to me to be a hangover from commitment to evaluative compositionality. If we suppose that every gamble has a unique rational price — and as we mentioned earlier, this is a prima facie attractive view — then it looks as though the rational prices of *S* and S' must be the same.[32] The picture we are working with now, however — in light of (RNP) — is that not every lottery *has* a unique rational price. The St Petersburg gamble, for example, has no unique rational price: it has many *rationally admissible* prices, each one determined by truncating at a rationally negligible ϵ. Once we grasp this point, it no longer seems irrational to value

[32] Assuming the values placed on the outcomes remain constant between 11 a.m. and 11.30 a.m.

S and S' differently. As long as we value each at one of its *admissible* prices — that is, we truncate at an ϵ that is indeed rationally negligible — then we have done nothing irrational. Compare the situation of a student choosing a seat in a lecture hall. Suppose students are allowed to sit anywhere — except in the first two rows, which are reserved for tutors. Our student picks a seat in the fifth row in the first week of classes. Does she then do something wrong if she picks a seat in the tenth row — or a different seat in the fifth row — in week two? Of course not. As long as she sits in an admissible seat each time, it does not matter if she sits in different seats each week. Given (RNP), an agent is in the same boat with respect to Strong Constancy: as long as she truncates at a rationally negligible ϵ each time, there seems to be no reason why she must truncate at the same point both times.

That leaves Weak Constancy. Recall that the problem with violating Weak Consistency is that one treats some outcome as having both zero probability and positive probability — in one and the same decision problem. If one violates Weak Constancy, one likewise treats some outcome as having both zero probability and positive probability — only not in the same breath: rather, one treats the outcome as having zero probability for purposes of making one decision and treats it as having positive probability for purposes of making another decision. Is this any better than violating Weak Consistency? I think it is. The case seems rather similar to (although admittedly not exactly the same as) that of violating Strong Constancy. (RNP) says that there are outcomes that a rational agent may ignore. So suppose she ignores some of them. Now rationality did not *demand* this: rationality said that they *could* be ignored, not that they *had to* be ignored. So, it seems, it would not be *irrational* for the agent later on to decide not to ignore them after all — or to ignore even more outcomes, provided they too are rationally negligible.

In sum, there are strong cases for accepting Weak Consistency as a requirement of rationality and rejecting Strong Constancy, and slightly weaker cases for accepting Strong Consistency and rejecting Weak Constancy. These tentative conclusions are sufficient in the present context: it is not necessary, for purposes of this paper, to make a definitive statement on whether to add each principle to the basic decision method of Truncation; they are all possible bolt-on options.

Before moving on, there is one further point that needs to be discussed. The Weak principles talk of two lotteries 'over the same set of outcomes'. This is to be interpreted in a strict sense: two lotteries are to be counted as being defined over the same set of outcomes only if they are *explicitly* so defined. An example will make this clear. Suppose that one is offered a choice between a St Petersburg gamble, based on a sequence of tosses of coin c, and a gamble S^* which is based on the very same sequence of coin tosses *and* on a sequence of die-rolls which are to be carried out alongside the coin tosses. That is, coin c is to be tossed until it comes up Heads, and each time it is tossed, a six-sided die d is to be rolled. If the coin lands Tails, we proceed to the next toss and roll. Once the coin lands Heads, the process of tossing and rolling ceases. At this point, S^* pays out the same as the St Petersburg gamble, unless (a) the die landed on consecutive numbers on consecutive rolls throughout the experiment (with 1 counted as coming after 6), in which case S^* pays out ten times what the St Petersburg gamble pays, or (b) the die landed on 6 on every roll, in which case S^* pays out nothing. The possible outcomes of S^* may be represented thus:

$$\begin{array}{ll} \text{Process ends at 1 toss/roll:} & \langle H,1\rangle, \langle H,2\rangle, \ldots, \langle H,6\rangle \\ \text{Process ends at 2 tosses/rolls:} & \langle T,1,H,1\rangle, \langle T,2,H,1\rangle, \ldots, \langle T,6,H,1\rangle \\ & \langle T,1,H,2\rangle, \langle T,2,H,2\rangle, \ldots, \langle T,6,H,2\rangle \\ & \vdots \\ & \langle T,1,H,6\rangle, \langle T,2,H,6\rangle, \ldots, \langle T,6,H,6\rangle \\ \text{Process ends at 3 tosses/rolls:} & \langle T,1,T,1,H,1\rangle, \ldots \\ \vdots & \vdots \end{array}$$

Suppose we explicitly define the St Petersburg gamble in the same way: that is, its possible outcomes are as above; if any of the first six outcomes listed above (i.e. outcomes of the form $\langle H,i\rangle$ for $i \in \{1,\ldots,6\}$) occurs, the payout is \$2; if any of the next thirty-six outcomes listed above (i.e. outcomes of the form $\langle T,i,H,j\rangle$ for $i,j \in \{1,\ldots,6\}$) occurs, the payout is \$4; and so on. In this case, the two lotteries (i.e. the St Petersburg gamble as just defined, and S^*) are *explicitly* defined over the same outcomes, and Weak Constancy applies. But suppose instead we define the St Petersburg gamble in the usual way—with outcomes as follows:

$$\begin{array}{ll} \text{Process ends at 1 toss:} & \langle H\rangle \\ \text{Process ends at 2 tosses:} & \langle T,H\rangle \\ \text{Process ends at 3 tosses:} & \langle T,T,H\rangle \\ \text{Process ends at 4 tosses:} & \langle T,T,T,H\rangle \\ \vdots & \vdots \end{array}$$

Now you might say that, in some sense, the outcomes of the St Petersburg gamble as just defined are the same as the outcomes of S^*: after all, it is not as though, when we are considering the St Petersburg gamble alone (and not thinking about S^*), the die-rolls somehow *do not happen*: they still happen — and there are still just as many possible ways they might happen — it is just that in so far as the St Petersburg gamble is concerned, it does not *matter* how they happen. That is why, when considering the St Petersburg gamble alone, it is convenient to ignore the die rolls: to clump together possible outcomes which differ in die-face but not in coin-side, as in the second presentation above. This is all well and good, but the fact remains that S^* and the St Petersburg gamble as most recently defined are not *explicitly* defined over the same outcomes, and Weak Constancy should *not* be applied. Look what happens if we (mistakenly) attempt to apply Weak Consistency to S^* and the St Petersburg gamble as most recently defined (henceforth S). Suppose we truncate S at $\epsilon = \frac{1}{12}$.[33] Then we are treating the probability of the outcome $\langle T, T, H \rangle$ as positive ($\frac{1}{8}$) and the probability of the outcome $\langle T, T, T, H \rangle$ (which, prior to truncation, is $\frac{1}{16}$, which is less than $\frac{1}{12}$) as zero. Now suppose we truncate S^* at the same point. Then we are treating each of the possible outcomes in which the process ends at 3 tosses/rolls as having probability zero (because prior to truncation, each has probability $\frac{1}{2^3} \times \frac{1}{6^3}$, which is less than $\frac{1}{12}$). But then we are treating the possibility of the coin landing Tails–Tails–Heads both as having positive probability — when presented as outcome $\langle T, T, H \rangle$ of S — and as having zero probability — when presented as the union of the outcomes $\langle T, i, T, j, H, k \rangle$ $(i, j, k \in \{1, \ldots, 6\})$ of S^*. And that is precisely the kind of divided-mind inconsistency that Weak Consistency is supposed to *avoid*.

10. Benefits of truncation

I have now presented the view that rational decision makers need not factor in outcomes of arbitrarily low probability and shown how it leads to a violation of evaluative compositionality. I said in section 6 that the question of whether to reject (RNP) on the basis of adherence to evaluative compositionality or vice versa was to be decided on the basis of consideration of the overall theoretical situation — by tracing out the consequences of each position and then forming a judgement

[33] Recall n. 22: here too, this ϵ seems too high actually to be a rationally negligible probability — but it makes the example easier to present if we pick a large ϵ.

as to which leads to the better overall view. I turn to such consideration now.

The first major advantage of the present proposal is that it provides a complete solution to the problems posed by the Pasadena game and its relatives. By contrast, no complete solution exists which retains evaluative compositionality. If we are to retain the latter principle for infinite gambles, we will need new decision rules: for the rule of maximizing expected utility does *not* determine a unique rational value for gambles such as the Pasadena game which have no unique expected value. But what will these rules be? Easwaran (2008) suggests valuing the Pasadena game at what he calls its *weak expectation*. Sprenger and Heesen (MS) offer some objections to this view — but more importantly in the present context, Easwaran himself acknowledges that the proposal is not a general solution to the problem of how to value any infinite gamble: 'it seems very likely that just as … expectations fail to exist for many games, so do weak expectations' (Easwaran 2008, p. 639). So at this stage, any plea to retain evaluative compositionality must be essentially empty: for there is no proposal on the table as to *what* function it is that takes us from the probabilities and values of its outcomes to the value of the gamble itself, which purports to work for all infinite gambles.

But wait: there is an option we have not yet discussed. We can retain the rule of maximizing expected utility for all infinite lotteries if we suppose that utilities are bounded (Hammond 1998, Sect. 8; Sprenger and Heesen MS). From a technical point of view, this solution is very attractive. It is well known that a bounded random variable has an expectation — so we can certainly make every gamble have an expected value by making the utility function bounded.[34] From a conceptual point of view, however, this seems unmotivated: it cuts the utility function to fit the decision theory (where the decision theory is thought of as telling us to maximize expected utility), whereas what we want is a decision theory which tells us what a rational agent would do — and there seems to be nothing *irrational* about having an unbounded utility function. Imagine a person who gets more value out of more money — or more chocolate ice cream, or whatever — at a linear rate. Lucky chap! Certainly this fact about his value structure does not make him irrational. (Contrast someone who has symmetric preferences: that really does seem irrational.) Now what should he do

[34] The utility function is (given certain natural assumptions) a random variable; the expected value of a gamble (if it exists) is the expectation of this random variable.

when offered the Pasadena game (framed in terms of monetary outcomes, or bowls of ice cream, or whatever)? Decision theory (thought of as telling us to maximize expected utility) gives him no advice. But then this decision theory is *not* a theory of ideal rational decision making. It is, at best, a theory of what rational decision makers *who happen to have bounded utility functions* should do. Thus the bounded utility approach does not ultimately avoid the problem that any plea to retain evaluative compositionality must be, at this stage, essentially empty: for there is no proposal on the table as to *what* function it is that takes us from the probabilities and values of its outcomes to the value of the gamble itself, which purports to work for all infinite gambles *and all rational decision makers*, no matter what their preference structure (as long as that preference structure is not in itself irrational).

The proposal that a rational agent's value for an infinite gamble is the same as her value for some truncation of it, on the other hand, requires no new decision-theoretic machinery (beyond Truncation itself): orthodox classical decision theory is sufficient. When faced with an infinite gamble, we truncate it. We then have a simple gamble: and the existing theory tells us how to handle it. Thus, the proposal does not work only for the Pasadena game: it is guaranteed to work for every infinite gamble. There is no possibility of someone's inventing an infinite gamble which the proposal cannot handle: for any infinite gamble can be truncated; every truncated gamble is a simple gamble; and the orthodox theory handles all simple gambles nicely. This stands in stark contrast to other proposals in the literature. As we have mentioned, Easwaran's proposal works (arguably) for the Pasadena gamble, but is not generally applicable — and Sprenger and Heesen's proposal, while it works for all gambles, does not work for all rational *agents*: it works only for agents who have bounded utility functions. Similarly, Baker (2007) puts forward a proposal for handling the Pasadena game, but then presents a new gamble — the Alternating St Petersburg Game — which his proposal does *not* handle.

The second major advantage of the proposal to solve decision problems by Truncation is that not only does it yield a solution to all decision problems involving infinite gambles: it furthermore gives solutions which — at least in regards to existing problem cases — are intuitively right. For example, in the case of the St Petersburg game, the proposal to value a gamble at its expected value leads to the result that we should assign infinite value to the game. Intuitively, this is absurd: that is why the term 'paradox' is widely used in connection with the St Petersburg game.[35] The present proposal, however, allows a rational agent to assign finite value to the game. This is an important point. The present proposal is *not* a conservative extension of the proposal to value a gamble at its expected value: that is, an extension which fills the gaps by providing a value for gambles which have no expected value, but values gambles which *do* have expected values at their expected value. The present proposal allows one to value the St Petersburg game at *less* than its expected value. But this seems *right*: the paradox was precisely that the game does not seem worth its expected value. Of course, no particular finite value is rationally mandated: the farther out one truncates, the higher the value will be. But to my ears at least, this sounds right. It is rationally permitted to ignore outcomes below a certain probability — but for any proposed stopping place, it would also not be irrational to factor in further, smaller probabilities. Rationality does not require consideration of arbitrarily small probabilities — but for any probability one decides to ignore, rationality permits one to consider a smaller probability. Thus, rationality allows multiple possible valuations of the St Petersburg gamble.

Some think that the problem posed by the St Petersburg game is not simply that intuitively, a rational agent would not pay an *arbitrarily* high amount to play it: the problem is that a rational agent might not even want to pay a *very* high amount to play it. My view accommodates this sort of intuition: someone who wants to pay less than $\$x$ to play the game can be seen as thinking that probabilities below y are rationally negligible (e.g. if x is 10 then y is 0.001).

A third significant benefit of the proposal is that it allows us to retain dominance reasoning. Suppose we are faced with a decision problem where we must decide which of two infinite lotteries to accept. On the proposal of this paper, we may truncate the lotteries at some point. Now if one infinite lottery dominates another (as Altadena dominates Pasadena, and the Petrograd game — which is just like the St Petersburg game except that each payoff is $1 higher Colyvan (2008, p. 37) — dominates the St Petersburg game) then the truncation of the first will have a higher expected value than the truncation of the second — and so, because the expected utility

[35] Not everyone agrees that it is wrong to value the St Petersburg game at its expected value; for example, Hájek and Nover (2006, p. 706) present 'a principled reason for accepting that it is worth paying any finite amount to play the St Petersburg game'.

property holds for simple lotteries, we will prefer the former. Of course this assumes that we truncate both infinite lotteries at the same point: but that we must do so (for purposes of addressing a single decision problem, such as a choice between Pasadena and Altadena) is mandated by Weak Consistency.[36]

Note that if one lottery only *weakly* dominates another (i.e. it never gives worse outcomes, but does not always give better outcomes) then Truncation will not always demand that one prefer the weakly dominant lottery. For example, consider a hybrid of the Pasadena and Altadena games whose payoffs match those of Pasadena for outcomes of Heads on tosses 1 through n, and then match those of Altadena for outcomes of Heads on subsequent tosses. If we truncate early, this game will have the *same* value as Pasadena (even though it weakly dominates the latter); if we truncate late, it will have a higher value. This seems exactly right. (RNP) tells us that there are outcomes which may rationally be ignored. We do not have to ignore them — but we may. If we do ignore them, then we ignore the differences between the hybrid game and the Pasadena game — and so the games are equally valuable (in our eyes). If we do not ignore them, then the hybrid game is preferable.

A fourth benefit of the present proposal is that it solves various other decision-theoretic puzzles that have been posed in the Pasadena literature. Consider a problem raised by Easwaran (2007, p. 12). Let A_H and P_H be the Altadena and Pasadena gambles, and A_T and P_T be analogous gambles in which the coin is tossed until it first lands Tails (rather than Heads). It seems that one should be indifferent between A_H and A_T and between P_H and P_T, and that one should prefer A_H to P_T and A_T to P_H. Classical decision theory cannot underwrite these judgements. The judgements *are* underwritten, however, by the truncation method (assuming Strong Consistency). For whatever ϵ we

[36] Recall that in Sect. 9 we accepted Weak Consistency as a requirement of rationality, but tentatively rejected its diachronic analogue, Weak Constancy. Without Weak Constancy, we cannot say that the value one places on Altadena in some context c must always be higher than the value one places on Pasadena in any other context c' (even assuming one's values for the outcomes of these gambles remain constant). This does not undermine the point being made about dominance — for the requirement that the value one places on Altadena in some context c always be higher than the value one places on Pasadena in any *other* context c' is overkill when it comes to underwriting dominance reasoning. What is required to underwrite dominance reasoning — 'choose the dominant lottery' — is that one prefer the dominant of two lotteries L_1 and L_2 when presented with a choice between them, i.e. when faced with a single decision problem involving L_1 and L_2. This is precisely what Truncation — together with Weak Consistency — ensures.

pick, A_H/ϵ and A_T/ϵ will have the *same* expected value (and likewise for P_H/ϵ and P_T/ϵ) while A_H/ϵ will have a *higher* expected value than P_T/ϵ (and likewise for A_T/ϵ and P_H/ϵ).[37]

Similarly, the truncation method solves a problem raised by Colyvan:

> consider a variation on the Petrograd game. Like the Petrograd, this game has payoffs $1 higher than the corresponding payoffs of the St. Petersburg game — except for one. The exception is a payoff for some very low probability state and this is $1 *less* than the corresponding St. Petersburg payoff. Call this game *the Leningrad game*. Here expected utility theory suggests that we ought to be indifferent between the Leningrad game and the St. Petersburg game; dominance reasoning is not applicable and so is silent. But there is a very strong intuition that the Leningrad game is better than the St. Petersburg game. After all, the Leningrad game almost dominates the St. Petersburg game and the probability of finding oneself in the non-dominant state is, by construction, very low. So here is the … challenge: either find a decision rule that supports this intuition or explain away the intuition. (Colyvan 2008, p. 38)

If we truncate the Leningrad and St Petersburg games before the low probability state, the truncated Leningrad gamble dominates the truncated St Petersburg gamble and so has higher expected value. If we truncate after the low probability state, then the truncated Leningrad gamble does not dominate the truncated St Petersburg gamble — but it *still* has a higher expected value (its low-probability less-good outcome is swamped by all the other better outcomes). Thus, Truncation (together with Weak Consistency) underwrites the intuition that the Leningrad game is better than the St Petersburg game.

11. Conclusion

Evaluative compositionality — the idea that the value that a rational agent places on a gamble is a function of the values that she places on the possible outcomes of the gamble, together with the probabilities assigned to those outcomes by the gamble — is a highly appealing principle. Unfortunately, no-one has shown that there is a way of retaining it that gives plausible results and works for all infinite gambles —

[37] Recall that in Sect. 9 we accepted Strong Consistency only somewhat tentatively. This is not a problem here, because the strength of the case for Strong Consistency seems to match precisely the strength of the intuitions about Easwaran's problem. It does not seem flat-out *inconsistent* to treat the probability of getting Heads first on toss n (in the sequence of tosses used to decide A_H) as zero, while regarding the probability of getting Tails first on toss n (in the sequence of tosses used to decide A_T) as positive (or vice versa) — but this does seem objectionably *arbitrary*.

including the Pasadena game and its relatives. That is, there is no rule on the table which takes as inputs values of outcomes and the probabilities assigned to those outcomes by a gamble, and gives as output a rational value for the gamble, which is known to assign a unique and plausible value to every well-defined gamble, finite or infinite. One response is the chocolate ice cream view: we can value the Pasadena game and its relatives however we like; *any* value is rationally permissible. This view involves a wholesale abandonment of evaluative compositionality — and it faces a number of objections. I have argued that there is a much more attractive option. On the basis of general considerations about normative theories of practical activities, it is plausible to adopt (RNP): the view that there is a *tolerance* on the norm of decision theory that tells decision makers to ignore outcomes whose probability is zero; that is, that infinite precision with respect to this norm is not a requirement of rationality. Given (RNP), we should accept that Truncation is a rational method of making decisions. Evaluative compositionality is then lost — and intuitively, that is a cost of the proposal. However, as already mentioned, it is a cost that no-one knows how to avoid. Furthermore, unlike in the case of the chocolate ice cream response, evaluative compositionality is not abandoned wholesale — and many benefits accrue. First, we get a complete account — and it involves no new decision-theoretic machinery (beyond Truncation itself): standard expected utility theory now suffices for the evaluation of all gambles, finite and infinite. Second, the values assigned to gambles are intuitively plausible. Third, by adopting a natural additional constraint — Weak Consistency — we can vindicate dominance reasoning. In short, the picture of decision theory that we get if we adopt (RNP) is more unified, tractable, and plausible than any alternative picture currently available.[38]

References

Armendt, Brad 1992: 'Dutch Strategies for Diachronic Rules: When Believers See the Sure Loss Coming'. *PSA: Proceedings of the*

[38] For comments and discussions, I am grateful to *Mind*'s Editor and anonymous referees; to Nick Agar, Arif Ahmed, Mark Colyvan, John Cusbert, Alan Hájek, Ed Mares, Michael McDermott, Kristie Miller, Huw Price, Dave Ripley, and Lionel Shapiro; to audiences at the Centre for Time at the University of Sydney on 12 April 2010, the AAP conference at the University of New South Wales on 9 July 2010, and the Philosophy department at Victoria University of Wellington on 1 December 2011; and to the participants in a reading group at ANU in July 2010. Thanks to the Australian Research Council for research support.

Biennial Meeting of the Philosophy of Science Association, 1 (Contributed Papers), pp. 217–29.

Baker, Alan 2007: 'Putting Expectations in Order'. *Philosophy of Science*, 74, pp. 692–700.

Barberà, Salvadore, Peter J. Hammond, and Christian Seidl (eds) 1998: *Handbook of Utility Theory, volume I: Principles.* Dordrecht: Kluwer.

Christensen, David 1991: 'Clever Bookies and Coherent Beliefs'. *Philosophical Review*, 100, pp. 229–47.

Colyvan, Mark 2006: 'No Expectations'. *Mind*, 115, pp. 695–702.

—— 2008: 'Relative Expectation Theory'. *The Journal of Philosophy*, 105, pp. 37–44.

Easwaran, Kenny MS: 'Dominance-based Decision Theory' (Typescript date 24 September 2007.).

—— 2008: 'Strong and Weak Expectations'. *Mind*, 117, pp. 633–41.

Fine, Terrence L. 2008: 'Evaluating the Pasadena, Altadena, and St Petersburg gambles'. *Mind*, 117, pp. 613–32.

Hájek, Alan and Harris Nover 2006: 'Perplexing Expectations'. *Mind*, 115, pp. 703–20.

—— 2008: 'Complex Expectations'. *Mind*, 117, pp. 643–64.

Hammond, Peter J. 1998: 'Objective Expected Utility'. In Barberà, Hammond, and Seidl 1998, pp. 143–211.

Nover, Harris and Alan Hájek 2004: 'Vexing Expectations'. *Mind*, 113, pp. 237–49.

Ramsey, F. P. 1990a: 'Truth and Probability', in his 1990b, pp. 52–94.

—— 1990b: *Philosophical Papers*, ed. D. H. Mellor. Cambridge University Press, Cambridge.

Resnik, Michael D. 1987: *Choices: An Introduction to Decision Theory.* Minneapolis: University of Minnesota Press.

Shrader-Frechette, Kristin 1985: 'Technological Risks and Small Probabilities'. *Journal of Business Ethics*, 4, pp. 431–45.

Sprenger, Jan and Remco Heesen MS: 'The Bounded Strength of Weak Expectations' (Revised version published in *Mind*, 120, 2011, pp. 819–32.).

von Neumann, John and Oskar Morgenstern 1944: *Theory of Games and Economic Behavior*. Princeton: Princeton University Press.

