

Weight for Stephen Finlay

Daan Evers

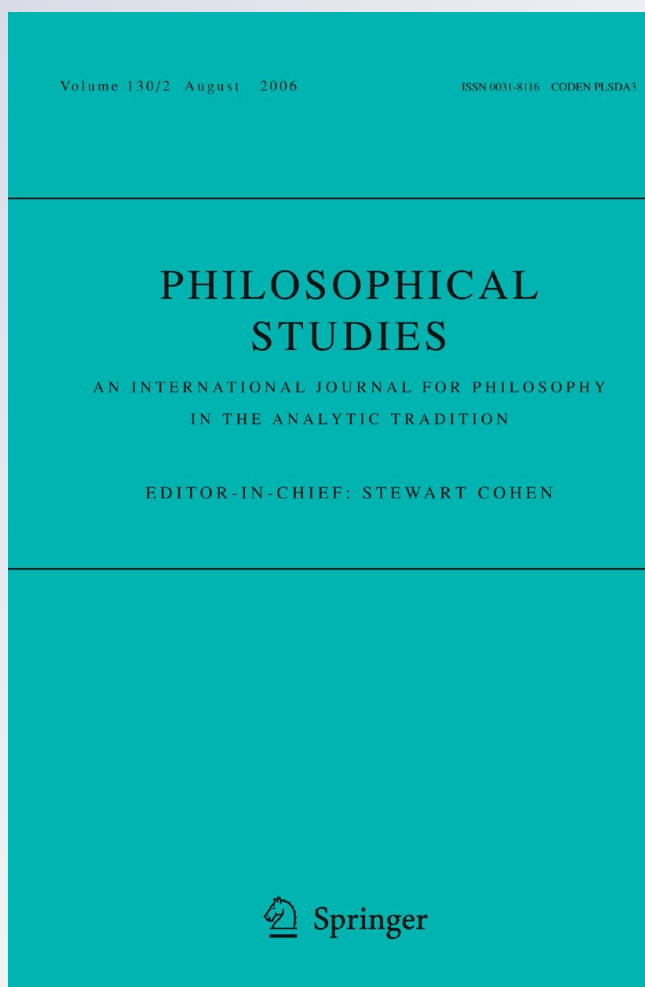
Philosophical Studies

An International Journal for Philosophy
in the Analytic Tradition

ISSN 0031-8116

Philos Stud

DOI 10.1007/s11098-011-9842-y



Your article is published under the Creative Commons Attribution Non-Commercial license which allows users to read, copy, distribute and make derivative works for noncommercial purposes from the material, as long as the author of the original work is cited. All commercial rights are exclusively held by Springer Science + Business Media. You may self-archive this article on your own website, an institutional repository or funder's repository and make it publicly available immediately.

Weight for Stephen Finlay

Daan Evers

© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract According to Stephen Finlay, ‘A ought to X’ means that X-ing is more conducive to contextually salient ends than relevant alternatives. This in turn is analysed in terms of probability. I show why this theory of ‘ought’ is hard to square with a theory of a reason’s weight which could explain why ‘A ought to X’ logically entails that the balance of reasons favours that A X-es. I develop two theories of weight to illustrate my point. I first look at the prospects of a theory of weight based on expected utility theory. I then suggest a simpler theory. Although neither allows that ‘A ought to X’ logically entails that the balance of reasons favours that A X-es, this price may be accepted. For there remains a strong pragmatic relation between these claims.

Keywords Stephen Finlay · Ought · Reasons · Weight of reasons · Expected utility theory · Moral semantics

1 Introduction

An important theory of normative discourse is Stephen Finlay’s (2004, 2005, 2006, 2008, 2009, 2010a, b). Finlay has developed theories of all the main normative words, like ‘good’, ‘ought’, ‘must’, ‘may’ and ‘reason’. One thing he has not yet done is develop a theory of a reason’s weight. This is not unimportant, for it may affect our assessment of his theory of ‘ought’. After all, there seem to be relations of entailment between ‘A ought to X’ and ‘The balance of reasons favours that A X-es’.

In this paper, I argue that an adequate theory of weight could not allow that ‘A ought to X’ (as interpreted by Finlay) logically entails that the balance of reasons

D. Evers (✉)

Department of Philosophy, Utrecht University, Janskerkhof 13a, 3512BL Utrecht, The Netherlands
e-mail: daan.evers@phil.uu.nl

favours that A X -es. Nor the other way around. This is the main object of the paper, but I briefly suggest that this result may not be deadly.

I will proceed as follows. In Sect. 2, I summarise Finlay's probabilistic theory of 'ought' in (2009). In Sect. 3, I show why the weight of a reason cannot be understood merely in terms of the degree to which an act raises the probability of an end. In Sect. 4, I develop a suggestion made by Finlay himself: to understand the weight of reasons in terms of expected utility. In Sect. 5, I show why this implies that ' A ought to X ' does not logically entail that the balance of reasons favours that A X -es. In Sect. 6, I argue that a simpler theory of weight will do, but that it too is incompatible with logical entailments.

2 Finlay's theory of 'ought'

Following Sloman (1970) and Kratzer (1977, 1981), Finlay observes that 'ought' is in the same family as nonnormative modal auxiliary verbs, like 'must', 'need to', 'have to', 'may', 'might', 'can', etc. They can be analysed in terms of possible worlds, as follows:

Modal Must: It must be the case that p = (df.) Every possible world in which circumstances/laws C obtain is such that p , and

Modal May: It may be the case that p = (df.) At least one possible world in which circumstances/laws C obtain is such that p (Finlay 2009, p. 319).

If you insert an 'in order that...' modifier into constructions with nonnormative modal auxiliaries, they start to look like instrumental normative statements. For example:

In order that I fulfil my duty as a judge, I must sentence you to death.

In order to sell your home, you could try advertising it online (2009, p. 320).

So Finlay hypothesizes that '[n]ormative modal terms (concepts) in their instrumental uses are simply ordinary modal terms (concepts) occurring under an "in order that..." modifier' (p. 320). In other words, instrumental normative terms are ordinary modals occurring in an end-relational context. Thus:

Must_E: In order that E it must be the case that p = (df.) Every possible world in which circumstances/laws C obtain, *including its being the case that E* , is such that p ;

May_E: in order that E it may be the case that p = (df.) At least one possible world in which circumstances/laws C obtain, *including its being the case that E* , is such that p (2009, p. 320).^{1,2}

'Ought' also has a nonnormative counterpart in its predictive use, as in: 'It ought to rain tomorrow'. A plausible analysis is:

¹ For the sake of uniformity, I have replaced Finlay's small case variable ' e ' in this and subsequent quotations with an upper case ' E '.

² Shouldn't there be some indication that E obtains *because* of p ? Finlay denies that this is part of the truth conditions of normative statements, although pragmatics will make it salient. These issues are discussed more fully in his (2010a).

Modal Ought: It ought to be the case that p = (df.) It is more likely, given circumstances C , that p than that any member of $\mathfrak{R}$ [set of alternatives] obtains (2009, p. 322).³

What about the normative ‘ought’? Since ‘ought’ is itself a modal auxiliary verb, Finlay thinks it too can be understood as the ordinary modal under an ‘in order that...’ modifier. If you ought (instrumentally) to do p , then p is *better* than the relevant alternatives. This is the case if p makes the end more likely than the relevant alternatives do. So for normative uses of the instrumental ‘ought’, Finlay proposes:

Ought_E: In order that E it ought to be the case that p = (df.) It is more likely, given circumstances C *including its being the case that E* , that p than that any other member of $\mathfrak{R}$ obtain (2009, p. 323, my italics).

Note that Finlay’s order of conditionalization is reversed: instead of making ‘ought’s into statements about the likelihood of ends given that some means is taken, Finlay makes them into statements about the likelihood of means given that some end has been achieved. This is (among other things) necessary to preserve uniformity with the predictive ‘ought’. It is also necessary to explain why ‘ A must X ’ entails that A ought to X , but not the other way around.

The suggestion that ‘ought’ statements are about the likelihood of means seems counterintuitive: normative ‘ought’s are not estimations of what an agent is in fact most likely to do. In order to avoid this, Finlay says that the circumstances relative to which the means are likely do not include facts about the agent’s psychological dispositions (Finlay 2009, p. 323, 2010b, p. 79). Context determines that it is part of C that any course of action is as likely to be taken as any other (Finlay calls this assumption *symmetry of choice* in 2010b, p. 80). Given symmetry of choice, if X is the most effective way of achieving E , then, given that E obtains, it is most likely that X was taken as a means (2009, p. 323).

Finlay conjectures that ‘[e]very normative ‘ought’ is [...] the ordinary modal ‘ought’ under an ‘in order that...’ modifier (p. 328, my italics). This includes *categorical* uses of ‘ought’: uses that seem unconditional (not relative to some end). Finlay believes that normative ‘ought’s without explicit ‘in order that...’ modifiers presuppose them contextually. The point of leaving out the modifier is rhetorical: it helps to communicate a demand that the agent acts in accordance with the values of the speaker (for details, see his 2009).

3 Weight and probabilities

So Finlay claims that ‘ A ought to X ’ means ‘given C including E , it is more likely that X than that any other member of $\mathfrak{R}$ obtains’. Given that Finlay analyses ‘ought’ in terms of probability, it seems natural to analyse the strength of reasons in terms of probability too. But before we look into this, we need to know what reasons are.

³ If, as Finlay does, we think of probability as a measure of the space of possibilities, then these analyses of normative modals lead to exceptional unification (unification of normative and nonnormative modals).

In (2006), Finlay says that a reason for an act X is a fact which explains why X -ing is conducive to E . This in turn can be analysed in terms of probability as follows:

Reason: A fact is a reason to X (relative to E) iff that fact explains why $P(E/C \& X) > P(E/C \& \text{not-}X)$.

A brief illustration: We're playing pool and you want to pocket the ball (your end E). Given this end, you have a reason to chalk the cue tip: chalking prevents the tip from slipping when it makes contact with the ball. This is a reason because it explains why you have a better chance of pocketing the ball if you chalk the tip than if you don't. Slightly more formally: the fact that chalking prevents the cue tip from slipping explains why P (pocketing the ball / C & chalking the cue tip) $>$ P (pocketing the ball / C & not chalking the cue tip).

Now for the weight of reasons: If what matters is the eventuation of E , we have most reason to do the act on the supposition of which E is most likely. Since 'most reason' means 'weightiest reason' (not 'greatest number of reasons'), the collective strength of reasons seems correlated with how likely the act is to realise the end. If so, we could formulate the following principle of weight:

Most Reason (prob): A has most reason to X (relative to E), iff X -ing raises the probability of E to a value higher than any relevant alternative in $\mathcal{R}$.⁴

Most Reason (prob) straightforwardly explains why ' A ought to X in order that E ' entails that the balance of reasons favours that A X -es (and vice versa). For symmetry of choice guarantees that if X is most conducive to E , then if E obtains, X is the means most likely chosen.

However, reasons derive from many different ends. But *Most Reason (prob)* identifies what we have most reason to do only *relative to E* , and we can't extend *Most Reason (prob)* to cover reasons derived from different ends. Suppose that X -ing makes E more likely than any other act makes likely any other salient end. Does it follow that we have most reason to X ? No, because E may not be as *important* as other salient ends. So we cannot determine what the balance of reasons requires by comparing the values to which acts, respectively, raise the probabilities of salient ends.

It seems, then, that an end-relational theory of weight has to take into account the importance of ends. One way to do this was suggested to me in personal communication by Finlay himself. The idea is to make the weight of reasons depend on expected utility. In other words: what one has most reason to do is what maximises expected utility. In the next section, I will flesh out this suggestion.

4 Weight and expected utility

To make headway, I will briefly state what expected utility theory is (ignoring various details). Expected utility theory (EUT) can be used as a normative and as a descriptive theory. In its normative guise, it is a theory of rational preference

⁴ For a more detailed discussion of this principle, see my (2010).

(a theory of which preference ordering it would be rational to have).⁵ Its essence consists in an expected utility function, which assigns an expected utility value to each act in a specified domain. Normative EUT assumes that it is rational to prefer one alternative to another just in case the first's expected utility is greater than the second's. In other words: it is rational to maximise expected utility.

What is expected utility? This notion is defined in different ways in different approaches, but we needn't go into the debate between the rivals.⁶ The problems for Finlay arise on all versions of EUT. So I will confine myself to an influential one developed by Richard Jeffrey (1965).⁷

Most versions of EUT assume that *rational* preferences satisfy a number of axioms, first formulated by Von Neumann and Morgenstern (1947).⁸ Examples are the following (where ' $X \geq Y$ ' means that one either prefers X to Y or is indifferent between them):

Transitivity: for every X, Y and Z with $X \geq Y$ and $Y \geq Z$ we must have $X \geq Z$.

Reflexivity: for every X , $X \geq X$.

Completeness: for every X and Y either $X \geq Y$, or $Y \geq X$.

This list is not exhaustive, but that is not important for our purposes. Von Neumann and Morgenstern proved that, provided the relevant axioms are satisfied, there is always some expected utility function which represents a subject's rational preferences.⁹ This also holds for Jeffrey's function.¹⁰

In Jeffrey's version of EUT, the expected utility of an act is determined by two factors: (1) the value of each of its possible outcomes and (2) the probability of each outcome conditional on performance of the act. More specifically, the expected utility of an act is the *sum* of the values of each possible outcome multiplied by that outcome's conditional probability.

In formal notation:

Let X be an act in a set of alternatives.

Let O_1, O_2 (etc.) be possible outcomes of X .

Let $P(O_i|X)$ be the conditional probability of O_i given X .

Let $V(O_i)$ be the value of O_i .

The expected utility of X can then be represented by the following sum:

$$EU(X) = (V(O_1) \times P(O_1/X)) + (V(O_2) \times P(O_2/X)) \dots + V(O_n) \times P(O_n/X).$$

⁵ I will return to the distinction between normative and descriptive uses of EUT in Sect. 6.

⁶ For a detailed overview, see Joyce (1999).

⁷ This version is commonly known as Evidential Decision Theory, as opposed to Causal Decision Theory (see, e.g. Joyce 1999, chapters 4 and 5). My discussion of Jeffrey is based on Joyce's book.

⁸ Or at least, most versions assume that there is some way of fleshing out the subject's rational preferences so as to make them satisfy the axioms (Joyce 1999, p. 45).

⁹ The sense in which such a function *represents* a subject's rational preferences is that it can be proven that X is preferred over Y if and only if X 's expected utility is greater than Y 's (see Joyce 1999, p. 43).

¹⁰ For a discussion of Von Neumann and Morgenstern's work, see Joyce (1999), chapter 1. For Jeffrey, see chapter 4. One difference between Jeffrey's version of EUT and earlier ones is that Jeffrey defined expected utility in terms of conditional probabilities.

In order to understand how Finlay might put EUT to use in his semantic framework, we have to consider what he might take the value of an outcome to consist in. But first we need to know what an outcome is.

In the context of Finlay's theory, I will take an outcome to be an end. This makes sense because we're interested in the weight of end-relational reasons. End-relational reasons are facts which explain why an act is conducive to an end. So it makes sense to make the weight of such reasons depend (at least in part) on the value or importance of the end which makes them into reasons. So the required kind of EUT is one where the expected utility of an act is a function of (1) the importance of the end to which it is conducive and (2) the probability that this end is realised conditional on performance of the act.

It should be clear that the notion of importance here is not end-relational (not a matter of conduciveness to some substantial end, like avoiding pain or getting richer). That would invite the problem raised in Sect. 3. So what is it determined by? For a philosopher like Finlay, it will be determined by the subject's relation to the end (his or her mental states with respect to it). Let us say that the degree to which the subject cares about the end determines its importance (for that subject). So we will be working with a theory according to which the expected utility of an act (for a subject S) is determined by (1) the degree to which S cares about the ends to which it is conducive and (2) the probability of their realisation conditional on performance of the act.

This brings us to the application of this theory to the problem of the weight of reasons. The idea is to make the weight of a reason to X in some sense depend on X 's expected utility. But, of course, we shouldn't make the weight of an individual reason to X depend on the sum of the products of the probabilities and importance of all the (relevant) ends to which X is conducive. Rather, we should make the weight of a reason to X depend only on the product of the probability and importance of the end to which the reason explains that X is conducive.

What we need to account for the weight of reasons is something like the following principle (where F_1 and F_2 are both reasons and C_1 and C_2 the background assumptions conditional on which the likelihoods of ends are determined):

More Reason (EUT)*: An E_1 -based reason F_1 is a stronger reason to X than an E_2 -based reason F_2 iff $V(E_1) \times P(E_1/C_1 \& X) > V(E_2) \times P(E_2/C_2 \& X)$.

This definition covers cases where E_1 and E_2 are different ends and when they are identical. In case they are identical (i.e. in case F_1 and F_2 derive their status as reasons from the same end), there can be a difference in weight between F_1 and F_2 only if there is a difference in the value to which they, respectively, explain that the probability of the end is raised. In such cases, we have to compare the probability of E given X and background information which includes F_1 but not F_2 with the probability of E given X and background information which includes F_2 but not F_1 .¹¹ So in these cases C_1 and C_2 are not identical.

¹¹ Of course, this makes sense only if C_1 and C_2 do not entail the relevant facts. So we'll have to assume that the context will always supply an appropriate set of background assumptions, via a principle of accommodation (thanks to Ralph Wedgwood). See also my (2010).

Our definition also needs to cover cases where F_1 and F_2 are each reasons for a different act. In case F_1 and F_2 are, respectively, reasons for X and Y in virtue of the same end E , any difference in strength will boil down to the difference between $V(E) \times P(E/C_1 \& X)$ and $V(E) \times P(E/C_2 \& Y)$. In case F_1 and F_2 are each reasons for a different act in virtue of a different end, then the difference in strength will be determined by the difference between $V(E_1) \times P(E_1/C_1 \& X)$ and $V(E_2) \times P(E_2/C_2 \& Y)$. A definition that covers all these cases is this:

More Reason (EUT): a reason F_1 derived from an end E_1 is stronger reason to X than a reason F_2 derived from an end E_2 is a reason to Y iff $V(E_1) \times P(E_1/C_1 \& X) > V(E_2) \times P(E_2/C_2 \& Y)$ (where ' X ' and ' Y ' are either the same or different acts in $\mathfrak{R}$, ' E_1 ' and ' E_2 ' are either the same or different ends and ' C_1 ' and ' C_2 ' are either the same or relevantly different background information).

The corresponding principle determining the balance of reasons reads as follows:

Most Reason (EUT): there is most reason to X iff the sum of all products $V(E_i) \times P(E_i/C_i \& X) >$ the sum of all products $V(E_j) \times P(E_j/C_j \& Y)$ (for all Y in $\mathfrak{R}$ nonidentical to X , all relevant ends to which X and Y are conducive and all relevant C_i and C_j , where ' i ' and ' j ' are variables ranging over $1, 2, \dots, n$).

This amounts to saying that, for any act X , there is most reason to X just in case X -ing maximises expected utility. For the expected utility of X is equal to $(V(O_1) \times P(O_1/C_1 \& X)) + (V(O_2) \times P(O_2/C_2 \& X)) \dots + V(O_n) \times P(O_n/C_n \& X)$. So the act Y which maximises expected utility is the act for which the sum of all products $V(E_i) \times P(E_i/C_i \& Y)$ is greatest.¹²

5 Weight as expected utility and the entailments

Suppose that *More Reason (EUT)* and *Most Reason (EUT)* give us the truth conditions of 'reason F_1 is weightier than reason F_2 ' and 'There is most reason to X '. Can Finlay now explain why ' A ought to X ' entails that there is most reason for A to X ? This primarily depends on what ends ' A ought to X ' is relative to. Suppose that it is a single end E_1 . From the fact that out of a set of alternatives, X is most conducive to E_1 , it does not follow that X maximises expected utility. After all, we

¹² John Broome (personal communication) suggested that *More Reason (EUT)* might be vulnerable to counterexample. He also suggested that this can be avoided by making the weight of an individual reason depend on the degree to which it increases the expected utility achieved from an end (see my 2011 for details). We can determine this degree by subtracting the value of $V(E_1) \times P(E_1/X)$ in the light of background information which excludes F_1 from the value of $V(E_1) \times P(E_1/X)$ in the light of background information which includes F_1 . This would also in principle allow for differences in strength between reasons for the same act that are based on the same end.

Broome's idea requires some changes to the principles in the main text (which I have done in 2011), but I have stuck to the simpler principles here for two reasons: first, the revisions make no difference to the main point of the paper, which is that any account of the weight of reasons in terms of expected utility does not allow logical entailments between ' A ought to X ' as interpreted by Finlay and 'The balance of reasons favours that A X -es' (see also footnote 13). Second, the revised principle for the collective weight of reasons still entails that we have most reason to X just in case X -ing maximises expected utility.

may not care about E_1 all that much. So it needn't be true that the sum of all products $V(E_1) \times P(E_1/X) >$ the sum of all products $V(E_2) \times P(E_2/Y)$.

Of course, the same holds if 'A ought to X' is relative to two or more ends. From the fact that out of a set of alternatives, X is most conducive to E_1 and E_2 (etc.), it does not follow that X maximises expected utility. After all, we may not care about E_1 and E_2 (etc.) all that much.¹³

What happens if we make 'A ought to X' itself relative to the end of maximising expected utility? From the fact that out of a set of alternatives, X is most conducive to maximising expected utility, it does follow that X maximises expected utility (relative to that set). And so it follows from 'A ought to X' that the balance of reasons favours that A X-es.

It may not be obvious that this follows, since being conducive to something means raising the probability that it transpires (which is compatible with its not transpiring). But remember that Finlay's direction of conditionalization is reversed. 'A ought to X in order that E' means 'Given E, it is most likely that A X-ed'. Now substitute for 'E' the maximisation of expected utility. It is obvious that, given that A has maximised expected utility, the act which has greatest expected utility is the one most likely performed by A.

However, I don't think Finlay can or wants to make 'ought' relative to the end of maximising expected utility for at least two reasons (the first of which he anticipated).

First, it collapses the distinction between 'ought' and 'must' in at least some cases. Remember that 'A must X in order that E' means that no worlds where E obtains are such that A does not X. But if there is only one act that maximises expected utility, then there are no worlds in which that end obtains in which the act is not performed. (The distinction still exists for cases where more than one act is optimal in terms of expected utility.)

Second, making 'ought' relative to the maximisation of expected utility would introduce information about the speaker's mental states into the meaning of 'ought' statements. That Finlay wants to avoid this is clear from his (2011).

I maintain that our moral claims are relativized to our standards or ends, but this is to be read *de re*, not *de dicto*; i.e. if my relevant moral end is E then my moral claim is to be interpreted as ought-relative-to-E, not as ought-relative-to-my-ends (Finlay 2011, p. 542).

If we think of E as the maximisation of expected utility, we have introduced information about the speaker into the meaning of moral statements, because being such as to maximise expected utility is (in Finlay's case) to stand in a certain relation to the subject's states of caring. This relation cannot be eliminated from our conception of expected utility, since we cannot make the value or importance of ends depend on conduciveness to substantial ends.

¹³ Notice that this would be so even if *Most Reason* (EUT) is not quite right. So long as the weight of reasons depends at least in part on states of caring or desire, 'A ought to X' will not logically entail that the balance of reasons favours that A X-es, nor the other way around (unless 'ought' statements are themselves relativized to the end of maximising expected utility, as we will see in the main text).

The point can be put more simply: the utility function represents the speaker's preference ordering (which is a function of the degrees to which the speaker cares about ends and their probabilities). So 'A ought to X in order that expected utility is maximised' can be unpacked as meaning 'A ought to X in order that we act in accordance with my preferences'.

Thus, making 'ought's' relative to maximising expected utility would make unavailable to Finlay his response to an objection by Richard Joyce (2011). Joyce complains about relativistic analyses of moral language as follows:

[Suppose] the judge at the Nuremberg trials kept relativizing his condemnation of the war criminals with the suffix "...by our moral standards." This would not just be weird and irritating; it would be scandalous; there would be protests. [...] The relativist Nuremberg judge [...] will be interpreted not as adding something unnecessary, but as revealing himself, in adding the suffix, to be saying too little (Joyce 2011, p. 527).

Finlay retorts that his end-relational analyses do not include positive determiners (like 'my', 'your', 'our', etc.).

Suppose for example that the [...] end is promoting general human wellbeing. In demanding respect for general human wellbeing and asserting that the Nazis acted in ways detrimental to that end, the judge would not be directing our attention to his/our own attitudes at all, but simply to the ideal of general human wellbeing, and its relation to the actions of the Nazis. [...] On this kind of relativist view, it is no essential part of what we as moral speakers communicate that we demand concern or respect for these ends because they are our ends (Finlay 2011, p. 543).

If Finlay made 'ought's' relative to maximising expected utility, then this response would be unavailable to him. For this end is intrinsically relational: it is the end of acting in accordance with one's preferences.

One might suggest that expected utility itself is not a matter of preferences. It is whatever values are determined by the utility function. And so the end of maximising expected utility can also be thought of as the end of securing the greatest value. No mental states involved! However, I don't think this is an option. Expected utility values have to be interpreted (they are the values of something which they represent). What they represent is a relationship between degrees of caring and the likelihoods of ends.

What this suggests, in my view, is that Finlay has no option but to deny that the entailments between 'A ought to X' and 'The balance of reasons favours that A X-es' are semantic or logical entailments. I will indicate why I don't think that is a disaster in the next section. But before I do that, we need to tackle an important issue regarding EUT. This issue partly motivates an alternative theory of weight.

6 A simpler theory of weight

In Sect. 4, I said that EUT comes in a normative and a descriptive guise. In its normative guise, it is a theory of rational preference. In its descriptive guise, it is a theory of people's actual preferences. Which one is relevant to Finlay?

The following line of thought seems to support the idea that it is really descriptive EUT that could be of use to Finlay. In order to retain a pragmatic (as opposed to logical) link between ‘A ought to X’ and ‘The balance of reasons favours that A X-es’, Finlay should say that a subject who says ‘A ought to X’ ordinarily does so because s/he prefers X over relevant alternatives. It is plausible that a speaker’s willingness to make this assertion is explained by the fact that s/he *actually* prefers X over relevant alternatives. But it is not plausible that a speaker’s willingness to make the assertion is explained by the fact that s/he would prefer the act if only s/he were rational. For if s/he does not actually prefer X, s/he would not choose to recommend X in the first place. So the pragmatic relation between ‘A ought to X’ and ‘The balance of reasons favours that A X-es’ exists only if what the balance of reasons requires is determined by the subject’s actual preferences.

If this is correct, then there is a problem. Empirical studies seem to show that people’s actual preferences don’t always conform to EUT (see for example Yaqub, Saz and Hussain 2009). People seem to violate various of the axioms formulated by Von Neumann and Morgenstern at least some of the time. Furthermore, it does not seem plausible that people’s states of caring are always determinate enough for precise assignments of values to $V(E)$ (which means that the completeness axiom is often violated; see also Joyce 1999, pp. 43–46).

This creates the following problem: if people’s judgments about the weight of reasons is plausibly correlated with their actual preferences, but people’s actual preferences cannot be described by a utility function, then why would *Most Reason* (EUT) give the truth conditions of judgments about the balance of reasons?

The solution to the problem of indeterminate mental states may consist in weakening the completeness axiom. Instead of demanding that the subject is either indifferent between X and Y, or prefers one over the other (for all relevant X and Y), we should demand that ‘it [...] be possible in principle to “flesh out” [an] agent’s preferences (usually in more than one way) to obtain a complete ranking that does not violate [the completeness axiom]’ (Joyce 1999, p. 45). As long as the subject clearly prefers one option over the others, it can be indeterminate exactly how much a subject cares about some of the other options.

What about other violations of the axioms? For example, it seems that people’s preferences differ when the same options are presented in a different way (violating an axiom of consistency, not listed above). People also seem to violate transitivity, although this is perhaps more controversial: some apparent violations may result from insufficiently fine-grained descriptions of the outcomes (see Broome 1993, pp. 100–104); although he concentrates on normative EUT, his reflections are useful in this context).

Even if this happens, it may still be plausible to assign *Most Reason* (EUT)’s truth condition to statements about the balance of reasons (and *mutatis mutandis* for *More Reason* (EUT)). What would be required is that in judging that the balance of reasons requires X people implicitly commit themselves to the claim that their preference for X is rational as specified by EUT.

What sort of evidence might there be for this claim? If people would retract their judgment that X is required by the balance of reasons upon discovering that their preferences violated axioms of EUT, we would have some reason to think that

judgments about the balance of reasons are intended as expressions of rational (instead of actual) preferences. But it would be difficult to establish this. An empirical study would have to make sure that everyone understood EUT, and there would have to be agreement on the interpretation of apparent violations. This is extremely difficult.

However, I am not convinced we need to take a stand on these issues. It is not clear to me that we need anything as sophisticated as EUT in the context of Finlay's semantics. What we do need is the two elements that I've suggested EUT should work with: degrees of caring and probabilities.

For Finlay, normativity is a matter of conduciveness to ends that matter to us. So how good a reason something is should depend on these two factors: how much we care about the end and its conditional probability. Now we basically have two options to consider: reasons derived from the same and different ends. I will start with the latter.

More Reason (informal): (A) In cases where two reasons F_1 and F_2 derive from different ends: F_1 is weightier than F_2 (for a subject S) iff S cares more that the end from which F_1 derives is promoted to the extent it is by the act for which it is a reason than S cares that the end from which F_2 derives is promoted to the extent it is by the act for which it is a reason.

(B) For cases where two reasons F_1 and F_2 derive from the same end: F_1 is a stronger reason than F_2 (for a subject S) iff F_1 raises the probability of the end to which the act for which it is a reason is conducive to a value higher than F_2 .

(A) and (B) specify a disjunctive truth condition for ' F_1 is stronger reason than F_2 '. We can formulate a simple truth condition for 'The balance of reasons requires that A X-es' as follows:

Most Reason (informal): There is most reason for A to X (for a subject S) iff S cares more that the ends to which X is conducive are promoted to the extent they are by X -ing than S cares that the ends to which alternatives are conducive are promoted to the extent they are by these alternatives.

In short: there is most reason for A to X iff one cares more that A X-es than one cares that any of the alternatives in $\mathfrak{R}$ are performed by A .

I would argue that it is an advantage that these truth conditions do not depend on the truth of EUT as a descriptive theory. After all, it is controversial whether EUT is adequate as a theory of actual preferences. It is also advantageous that they don't require the correctness of EUT as a theory of rational preferences. For this too is controversial. Another advantage is that the definitions do not involve more metaphysical material than Finlay is committed to already: states of caring and probabilities.

The major downside is that the theory requires us to deny that ' A ought to X ' logically entails that the balance of reasons favours that A X-es. For ' A ought to X ' (being a statement of conditional probability) does not entail that the speaker cares more that A X-es than s/he cares about any of the alternatives in $\mathfrak{R}$. And so ' A ought

to X does not logically entail that the balance of reasons favours that A X -es (nor will the latter entail the former). But I think this is acceptable, for two reasons.

First, although ' A ought to X ' does not logically entail that the balance of reasons favours that A X -es (or the other way around), there is still a strong pragmatic relationship between the two. Speakers will be prepared to say ' A ought to X ' only if they care more that X is done than they care about alternatives. And when they care more that X is done, they will be prepared to say that A ought to X . So we can deduce that the balance of reasons favours X (at least from the point of view of the speaker) if we understand what motivates 'ought' statements in the first place.

Second, phrases like 'the balance of reasons favours X ' are hardly colloquial. So we don't have many data from everyday language to suggest that the entailments are logical, rather than pragmatic. It is perhaps worth noting that it has been denied before that these entailments are logical. David Wong (1984) analysed ' A ought morally to X ' in terms of the requirements of a system of moral rules. Relevant systems were ultimately determined by the speaker's mental states. But Wong was a Humean about reasons. And Humeans think that the truth of ' A has a reason to X ' depends on A 's desires. Since what is required by moral rules to which the speaker is committed does not entail anything about the agent's desires, Wong denied that ' A ought to X ' logically entails that there are moral reasons for A to X . Instead, he offered an explanation of why we normally expect other people to have reasons to act in accordance with the moral systems to which we are referring. Finlay can do something similar. For example, he could appeal to the fact that normative discussion usually presupposes shared ends and assignments of importance (as indeed he does in 2008).

7 Conclusion

I set out to show that a plausible theory of weight for Stephen Finlay could not explain that ' A ought to X ' logically entails that the balance of reasons favours that A X -es (nor the other way around). Finlay's probabilistic theory of 'ought' could not be paired with a probabilistic theory of weight according to which we have most reason to X just in case X -ing makes the end to which it is conducive more likely than other relevant acts make likely other ends. For this theory ignored the importance of ends as a factor in the weight of reasons. I set out to develop a more promising suggestion by Finlay. The suggestion was to understand weight as determined by an act's expected utility. I developed one version of this idea, and argued that it was incompatible with logical entailments between ' A ought to X ' and 'The balance of reasons favours that A X -es'. It also seemed preferable to avoid issues concerning EUT as a theory of actual and rational preferences. So I developed another, simpler theory of weight. This theory made the balance of reasons depend on what the speaker cares about most. Neither theory could explain why ' A ought to X ' would logically entail that the balance of reasons favours that A X -es. In my view, this may be accepted. But others may believe it is a weighty problem.

Acknowledgments I would like to thank Stephen Finlay, Ralph Wedgwood, John Broome, Natalja Deng and an anonymous referee for *Philosophical Studies* for discussion and helpful comments. Work on this paper was made possible by the Royal Institute of Philosophy, the Prins Bernhard Cultuurfonds, the DAAD's Michael Foster Memorial Scholarship, Jesus College Oxford and the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013)/ERC Grant agreement No 263227.

Open Access This article is distributed under the terms of the Creative Commons Attribution Non-commercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Broome, J. (1993). *Weighing goods*. Oxford: Basil Blackwell.
- Evers, D. (2010). The end-relational theory of 'ought' and the weight of reasons. *Dialectica*, 64(3), 405–417.
- Evers, D. (2011). *Subjectivist theories of normative language*, D.Phil. thesis, University of Oxford.
- Finlay, S. (2004). The conversational practicality of value judgement. *The Journal of Ethics*, 8(3), 205–223.
- Finlay, S. (2005). Value and implicature. *Philosophers' Imprint*, 5(4), 1–20.
- Finlay, S. (2006). The reasons that matter. *Australasian Journal of Philosophy*, 84(1), 1–20.
- Finlay, S. (2008). The error in the error theory. *Australasian Journal of Philosophy*, 86(3), 347–369.
- Finlay, S. (2009). Oughts and ends. *Philosophical Studies*, 143(3), 315–340.
- Finlay, S. (2010a). Normativity, necessity, and tense. A recipe for homebaked normativity. In R. Shafer-Landau (Ed.), *Oxford studies in metaethics* (Vol. 5, pp. 57–85). Oxford: Clarendon.
- Finlay, S. (2010b). What *ought* probably means, and why you can't detach it. *Synthese*, 177(1), 67–89.
- Finlay, S. (2011). Errors upon errors: a reply to joyce. *Australasian Journal of Philosophy*, 89(3), 535–547.
- Jeffrey, R. (1965). *The logic of decision*. Chicago: The University of Chicago Press.
- Joyce, J. (1999). *The foundations of causal decision theory*. Cambridge: Cambridge University Press.
- Joyce, R. (2011). The error in "the error in the error theory". *Australasian Journal of Philosophy*, 89(3), 519–534.
- Kratzer, A. (1977). What 'must' and 'can' must and can mean. *Linguistics and Philosophy*, 1(3), 337–355.
- Kratzer, A. (1981). The notional category of modality. In H. J. Eikmeyer & H. Rieser (Eds.), *Words, worlds, and contexts* (pp. 38–74). Berlin: Walter de Gruyter.
- Sloman, A. (1970). Ought and better. *Mind*, 79(315), 385–394.
- Von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior* (2nd ed.). Princeton: Princeton University Press.
- Wong, D. (1984). *Moral relativity*. Berkely: University of California Press.
- Yaqub, M. Z., Saz, G., & Hussain, D. (2009). A meta-analysis of the empirical evidence on expected utility theory. *European Journal of Economics, Finance and Administrative Sciences*, 15, 117–133.