

*Dutch Books, Additivity, and Utility Theory*¹

Brad Armendt
Arizona State University

One guide to an argument's significance is the number and variety of refutations it attracts. By this measure, the Dutch book argument has considerable importance.² Of course this measure alone is not a sure guide to locating arguments deserving of our attention—if a decisive refutation has really been given, we are better off pursuing other topics. But the presence of *many* and *varied* counterarguments at least suggests that either the refutations are controversial, or that their target admits of more than one interpretation, or both. The main point of this paper is to focus on a way of understanding the Dutch Book argument (*DBA*) that avoids many of the well-known criticisms, and to consider how it fares against an important criticism that still remains: the objection that the *DBA* presupposes value-independence of bets.³

It is safe to say that the influence of Bayesian approaches to current work in decision theory, epistemology, and philosophy of science, is not on the wane—in fact the opposite. Yet the acceptance of the *DBA*, which is heuristically the most potent argument for that approach, may be in decline. If so, perhaps the explanation is that we have come to rely more on the decision-theoretic foundations for Bayesian theory, and one good argument is enough. Another point of this paper therefore, is to compare, in the light of objections to it, the status of the *DBA* with that of the decision-theoretic argument(s)

¹ Thanks to Stewart Cohen, Jim Joyce, Mark Kaplan, Bernie Kobes, Patrick Maher, Brian Skyrms, and Bas van Fraassen for helpful discussions related to this paper. Special thanks to Jim Joyce for his comments on an earlier version; this is much improved as a result. But they do not necessarily agree with all the views expressed here.

² Discussions of the argument, with references to some of the very extensive literature, can be found in Ellery Eells, *Rational Decision and Causality* (Cambridge: Cambridge University Press, 1982), pp. 46-52, and in Patrick Maher, *Betting on Theories* (Cambridge: Cambridge University Press, 1992), chapter 4.

³ This objection appears in Henry Kyburg, "Subjective Probability: Criticisms, Reflections, and Problems," *Journal of Philosophical Logic* 7 (1978): 82-83, reprinted in Kyburg's *Epistemology and Inference* (Minneapolis: University of Minnesota Press, 1983). It is prominent in Frederic Schick's "Dutch Bookies and Money Pumps," *Journal of Philosophy* 83 (February 1986): 112-119.

based on rational preference constraints. My overall purpose is to set out what a *DBA* really says and presupposes, rather than to argue that we are all compelled to accept its conclusion. But we will find that the argument has more value, and a more legitimate claim on our attention, than most of its critics think.

The Dutch book argument as drama.

We will not proceed by cataloging and discussing all the interesting objections to the *DBA*; nor will we consider a multitude of interpretations. Instead we will move directly to a sketch of the interpretation of the argument that eludes many of the standard criticisms, and to the value-independence objection. This will presuppose some acquaintance with the *DBA*.

In an often-quoted passage, Frank Ramsey writes:

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 ... 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.⁴

This passage comes after Ramsey's presentation of his axioms for rational preference; the inconsistency he explicitly mentions involves the violation of those axioms. Two things are worth pointing out about the way Ramsey raises the idea of the *DBA* in the last two sentences. First, he ties it to the "absurdity" of letting one's choice be governed by the manner of presentation, or description, of options. When is it absurd to do this? Perhaps when we evaluate the same option differently under different descriptions, and let our choice depend on which description is used. Second, it is at least a very plausible interpretation, if not an inevitable one, that Ramsey means the Dutch book to be an illustration of an already bad situation, rather than the sole basis for criticism of the beliefs that generate it. Both of these points have been made and emphasized by Brian Skyrms.⁵

⁴ "Truth and Probability," (1926), in Ramsey's *The Foundations of Mathematics* (London: Routledge and Kegan Paul, 1931), p. 182. Ramsey's axioms are for preferences over gambles. At this point he has already discussed at some length the idea that it is a useful idealization to regard action-guiding degrees of belief as betting odds. In his 1929 note, "Probability and Partial Belief," Ramsey says, "A theory is a set of propositions [closed under entailment]. The interest of such sets comes from the possibility of our adopting one of them as all we believe. A probability-theory is a set of numbers associated with pairs of propositions obeying the calculus of probabilities. The interest of such a set comes from the possibility of acting on it consistently," *Foundations of Mathematics*, p.256.

⁵ "Higher Order Degrees of Belief," in *Prospects for Pragmatism*, ed. D.H. Mellor (Cambridge: Cambridge University Press, 1980), pp. 109-137; also *Pragmatics and Empiricism* (New Haven:

Here is my version of the way to view the *DBA*; a reader familiar with Skyrms' writings will recognize how heavily I rely on them: Begin with the bare bones of the argument: It assumes an ideal, artificial *betting scenario*. In this scenario, the agent announces fair betting quotients for wagers on propositions, and commits himself to taking bets at those rates for stakes of any size.⁶ The scenario includes a cunning bettor, who knows the agent's advertised odds, but who has no information about the world the agent lacks. The agent is vulnerable to a Dutch book if the cunning bettor can offer him a set of bets at the agent's own odds, such that, however the bets end up paying off, the agent is bound to suffer a net loss. A sure loss is not a good situation for the agent; it is what we may call a *pragmatically defective outcome (DO)*.⁷ When the *DO* is produced by the agent's own fair betting quotients (and when other fair betting quotients would avoid it), we have a reason for judging the agent's betting behavior flawed.

What has this got to do with defending the principle that rational degrees of belief are probabilities?⁸ The fair betting quotients are supposed to *be* the agent's degrees of belief; when their violation of the laws of probability yields Dutch book vulnerability in the betting scenario, the stigma carried by *DO* is transferred to the betting quotients, *i.e.* the agent's beliefs. The stigma is only avoidable by complying with the probability laws. (A converse Dutch book argument shows that compliance is sufficient for avoiding a Dutch book.)

Now as Ramsey was well aware, the betting scenario is a contrived one, and much has been made of it since. Why should the announced odds be the agent's beliefs? Bets may never pay off. There are not actually any cunningbettors. If there were, the sure loss under compulsion to wager is really no indication of irrationality. And so on. But these objections miss the point of the argument, we now say. We should resist the temptation to think that a *DBA* demonstrates that violations of probability are bound to lead to dire outcomes for the unfortunate agent. The problem is not that violators are bound to suffer, it is that their action-guiding beliefs are flawed. The flaw is that they are

Yale University Press, 1984), 21-26; also "Coherence," in *Scientific Inquiry in Philosophical Perspective*, ed. N. Rescher (Pittsburgh: University of Pittsburgh Press, 1987), pp. 225-242.

⁶ Take a fair betting quotient on *p* to be the least favorable acceptable betting quotient on *p*, rather than the betting quotient set without knowing the direction of the bet. The latter characterization assumes too much when additivity is at issue, as it will be below.

⁷ Sure losses may not be the only situations that count as pragmatically defective outcomes, but they are what standard Dutch books yield. In some other arguments (e.g. the argument for strict coherence, and some diachronic arguments for Jeffrey Conditionalization) the *DO* is possible loss, with no possibility of gain.

⁸ Or, with defending other norms; Dutch book arguments have been given in defense of Conditionalization, Jeffrey Conditionalization, and Reflection. See Brad Armendt, "Dutch Strategies for Diachronic Rules: When Believers See the Sure Loss Coming," in *PSA 1992, Vol.1*, ed. D. Hull et al. (East Lansing: Philosophy of Science Association), pp. 217-229.

tied to inconsistency, of the kind Ramsey suggests: an inconsistent evaluation of a single option under different descriptions. That inconsistency can be vividly depicted by imagining the betting scenario and what would befall the violators were they in it. The idea is that the irrationality lies in the inconsistency, when it is present; the inconsistency is portrayed in a dramatic fashion when it is linked to the willing acceptance of certain loss. The objections just mentioned attack the drama, but not the situation it illustrates.

In judging the worth of the *DBA*, then, we should replace the question, "On what grounds do you suppose that what you demonstrate for the highly artificial betting scenario has also been demonstrated in general?" with the questions, "How well does the betting scenario serve as a setting in which to illustrate features of belief systems? How well do wagers stand in for choices we make under imperfect knowledge? How well do payoffs represent our goals and interests? And how well do betting quotients represent dispositional degrees of belief *qua* basis for action?" Imperfectly, no doubt. But is the imperfection largely a mismatch of generality and precision, or is it a complete absence of recognizable similarity between our beliefs and choices, on the one hand, and the betting scenario and its devices on the other? To display the betting scenario as a valuable illustration, it is not necessary to argue that no other treatment could be as good. For those unsympathetic to the betting scenario, it is worth pausing for a moment and reflecting on what superior treatment might be given, one having a device more general, as familiar, and/or more manageable than the wagers of the betting scenario, for dealing with action-guiding beliefs.

The Dutch book argument connects violations of the probability laws with choices that yield a *DO*. It proceeds by drawing on the action-guiding character of the violators (degrees of belief); by supposing a scenario in which those guides are operative; and by invoking (through the cunning bettor's offers) a particular pattern of guidance that yields the *DO*. When the *DO* is sure loss, derived from exchanges constructed only by reference to the agent's action-guides, what goes on is this: The agent is susceptible to exploitation because the agent displays pragmatic inconsistency, in that he gives conflicting evaluations to the same option(s). When this is so, the violation of laws of probability is tied to the inconsistency. Hence when the inconsistency is objectionable, so is the violation. To support this interpretation, we need to say more about the inconsistency. In what sense is it objectionable? How is it the source of the believer's dramatic loss? Is *DB* vulnerability a reliable indicator of it?

Ramsey calls it "absurd" to let one's choice depend on the way one's options are presented. I say it *is* a flaw of rationality to give, at the same time, two different choice-guiding evaluations to the same thing. Call this *divided-mind* inconsistency.⁹ It is difficult to find principles more fundamental, judgments more secure, from which to

⁹ Diachronic Dutch strategy arguments raise another possibility, change-of-mind inconsistency. See Armendt, "Dutch Strategies," p. 219. See also David Christensen, "Clever Bookies and Coherent Beliefs", *Philosophical Review* 100 (1991): 229-247.

argue that divided mind inconsistency is a flaw. Notice that the norm of consistency to which we appeal, consistency in evaluating our options, is not the pragmatic norm recommending that one always act so as to promote one's goals. The norm that divided-mind inconsistency is to be avoided stands on its own, not on a justification by the pragmatic norm. It is instead a norm that regulates how we should conceive of, or specify, our interests (*i.e.* consistently). There can be cases in which pursuit of our other interests is overriding. About them we say, "Avoidance of divided-mind inconsistency is important, but its importance may sometimes be overridden by rich rewards for being inconsistent. This shows not that the probability rules have not been justified, but simply that rational choice might sometimes yield irrational belief."¹⁰ So we can add this: In claiming that divided-mind inconsistency is objectionable, we are not necessarily claiming that it is more objectionable than all other things. And in attributing it to an agent, we need not be saying more than that he exhibits an imperfection of pragmatic rationality. It does not automatically follow that the agent is crazy, hopelessly confused, or blameworthy; it could be a minor inconsistency, or inconsistency in a complex situation, or inconsistency whose importance is overridden by other considerations. The point remains, however, that divided-mind inconsistency is a flaw in the agent's set of judgments and evaluations, if not necessarily in the agent.

What basis is there for thinking that Dutch book vulnerability, when it occurs, arises from such inconsistency? *I.e.* what basis is there for Ramsey's claim in the next-to-last sentence quoted? The answer must come from the structure of the Dutch book argument, and will soon take us to the objection about value-independence of the bets.¹¹ The idea is this: a believer whose partial beliefs violate the probability rules 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) in a way that is essentially quite simple: the bettor sells the option to the agent at the higher of the

¹⁰ I discussed this, particularly in connection with diachronic rules for belief change, in "Wanted: Irrational Belief Changes (Reward)," colloquium paper for the 1992 APA Eastern Division meeting, Washington, DC.

¹¹ Vagueness in our evaluations, or in our beliefs, can be a source of inconsistency, in that their refinements may disagree; here we follow Ramsey in assuming that vagueness is not an issue, while agreeing that in reality our beliefs are vague. How to think about the connection between norms for precise rational belief, and beliefs that are actually vague is something we will not explore in this paper. It is inevitable, even (especially) with respect to full belief, that many of our beliefs will be vague; we still recommend against having contradictory precise full beliefs. And we typically do not investigate matters of importance with an eye to blurring our opinions, rather than making them more precise. I subscribe to the view that good norms and defenses of norms for precise belief regulate vaguer belief indirectly. For an opposing view, particularly on the *DBA*, see Henry Kyburg, "Subjective Probability," and "Bets and Beliefs," *American Philosophical Quarterly* 5 (January 1968): 54-63, reprinted in *Epistemology and Inference*, 63-78.

prices, and buys it back at the lower price.¹² The bettor's strategy is slightly complicated by the nature of the goods exchanged (bets), and by what is involved in obtaining value from them (payoffs when outcomes are determined), but his *basic* strategy is simple: sell the agent something at a high price, and buy what is the same thing, or what is in all relevant respects an equivalent thing, back at a low price, and pocket the difference.

Of course an important element of the simple strategy just described, one that is crucial in judging the rationality or irrationality of leaving oneself open to it, is the identity, or equivalence, of the options that are bought and sold. After all, however irrational is the behavior just described, it need not be irrational to evaluate *different* things differently, and buy the expensive one while selling the cheap one. A net outlay will occur, but if net value is received, so what? To pursue this, we must look in more detail at the recipes for the Dutch books.

Bets, books, and value.

The *DBA* is supposed to show that rational degrees of (partial) belief are probabilities. There is little doubt that it has also been taken as support for the idea that we *have*, or can be fruitfully theorized to have, degrees of belief (if not always rational ones). But any such support is at best indirect, by displaying the *fruits* of an underlying view about dispositional action-guiding belief. It is a mistake to suppose that the *DBA*, or anything like it, could prove out of thin air, on the basis of no substantial assumptions, that there *are* partial action-guiding beliefs. This is worth mentioning because it suggests a distinction that can be made among critical responses to the *DBA*. One might agree, if only for purposes of discussion, that action-guiding partial beliefs can be attributed to us, while rejecting what the *DBA* claims about them (that when they violate probability, they are flawed). Or, on the other hand, one might reject the whole idea of action-guiding partial belief (as something we have). A criticism of the *DBA* which denies that there is anything we do that could be illustrated by the betting scenario (sometimes the denial is not explicit), amounts to a rejection of the subject of the *DBA*. If the criticism is made by saying that the *DBA* has failed to prove that the betting scenario illustrates something about us, then it shows misunderstanding of the point of the argument; worse than that, it may amount to an impossible demand to prove something, while assuming nothing. Either way, such a criticism differs from saying (even just for the purposes of argument) that we do have partial action-guiding beliefs, but the *DBA* is a fallacious attempt to say something interesting about them. We can better see the significance of this general point by looking at several of the *DBA*'s presuppositions.

Consider, for example, the passage of time.

¹² A complication arises when diachronic Dutch Books are constructed: they involve extending a conditional sure loss for the agent, one that can be imposed should a particular eventuality occur, to an unconditional sure loss for the agent, by offering him an additional side bet against the eventuality.

The *DBA* involves the construction of books, combinations of bets. Making bets takes time; so the *DBA* essentially involves a procedure that requires the passage of time (which makes relevant the possibility that the believer is changing his/her mind). The temporal interval during which the process occurs undermines its connection and relevance to the synchronic norm. The *DO* turns out to be no worse than what might happen when we do what the *DBA* cannot rule out, namely, change our minds.

This is a weak criticism. It might be a more telling one if we regarded beliefs as states incapable of enduring over time. If any presupposition that they do endure could be properly rejected as contrary to their nature, then the book-making procedure might be an objectionable flaw in the betting scenario. But beliefs *do* endure from one moment to the next. Or put another way, if there are good objections to be made to the Bayesian's presupposition that we have dispositional action-guiding beliefs, the claim that beliefs cannot endure over time is not one of them. It is obvious from the start that the *DBA* is meant to be a defense of a synchronic norm; hence the dramatic betting scenario must be applied to simultaneous action-guiding beliefs. Can we imagine sufficient stability of the relevant beliefs to permit the book-making procedure into the illustrative drama? We can, once we realize that no one is claiming to prove the stability *must* be there.¹³

Here is another supposition: The beliefs that the Bayesian proposes to theorize about are represented by betting quotients, which are ratios of payoffs. It is supposed to make sense to think of a degree of belief in p as the (fair) betting quotient on p ; there is supposed to be a single betting quotient for p . Among the many points that can be raised, consider this one: The quotient or ratio is fixed, whatever the size and scale of the payoffs.¹⁴ Now this could immediately lead us to a distinction between quantities of goods and utilities (to which we will attend soon), but let us set that aside (focus on

¹³ Compare this exchange: *A* says, "What's wrong with (fully) believing the denials of what your (full) beliefs entail? Well, then your beliefs are inconsistent, and at least one of the beliefs you have, and one of the things you express when you express your beliefs, is false." *B* replies, "Expressing one's beliefs takes time, and it might be that really you expressed those beliefs because you changed your mind, not because you had inconsistent beliefs." I take it *B* has not decisively refuted *A*.

¹⁴ Freedom to choose stakes of various sizes is needed for the cunning bettor in the synchronic *DBA* for probability, only to blow up tiny violations so that sure losses are not so tiny (it is strictly necessary if the goal is to demonstrate sure loss of at least some minimal size); it is necessary in diachronic *DBAs* so that the proper side bets can be arranged. It isn't actually otherwise necessary for the synchronic *DBA* for probability (but as Jim Joyce reminds me, it is needed for the converse *DBA*). This freedom is usually granted to the bettor, and the view that betting quotients/degrees of belief do not vary with the stakes is generally held. The relevant philosophical question here is whether, or to what extent, beliefs are like this (and not, e.g., whether it is wise to wager with a bettor having such freedom). Whether belief is affected by the size of the stakes is not the same question as whether bets are value independent, but a connection will be noted below.

payoffs in utility, if that helps) and ask, *do* we have beliefs like that? That is, do I typically have a single belief in, say, *storm tomorrow* whether my concern is losing my sailboat and life, or the inconvenience of carrying an umbrella? It is surely true that the scale of the payoffs affects how we make choices, perhaps even in ways that are not entirely, or at least not easily, explained by the move to utility. But the present question is whether *what I believe* (to what extent I believe it) varies according to the importance of the interests at hand. The view embodied in the betting scenario on this issue is that, while we may adjust our decision-making when the importance of the decision changes (*e.g.*, life-or-death decisions are more conservative than decisions on small matters of convenience), those adjustments are in our utility function and decision-making, rather than variation in our belief.¹⁵ This view accords as well as any with how we think of belief, but its truth is not, I think, rigorously demonstrable from neutral assumptions. The point to be made here, though, is that rejection of this view amounts to rejection of the Bayesian's conception of degree of belief; criticism of the *DBA* on these grounds is a rejection of it before it starts, irrespective of its content, and not a demonstration that it is fallacious.

We should note another common way of putting this presupposition: The betting scenario presupposes not that fair betting quotients are invariant over stakes of all sizes; instead it presupposes that all wagers are made at stakes falling within some suitable range of sizes, and the fair betting quotients are fixed within that range. On this way of putting the assumption, the agent's degrees of belief are represented by some of the odds at which he would bet, namely when the wagers are at suitable stakes; wagers at extreme stakes, outside the suitable range, are not reliable indicators of his degrees of belief. This way of putting the assumption has the effect of ruling out some of the more unrealistic wagers in the betting scenario, but for our purposes there is little difference in the two ways of putting the assumption. The difference would be important if our project were to *measure* the agent's beliefs by use of the betting scenario. As it would also be, if we regarded fair betting quotients as *defining* the degrees of belief. Subjective probability theory has been much influenced through its history by behaviorism. In the first place, it

¹⁵ How we talk about what we believe, especially when we use full-belief language, is influenced by the stakes, as well as by other contextual factors. So we may find ourselves describing an attitude as belief on one occasion, but as conjecture on another. When the difference results from a shift in the standards for applying these terms, the point about the invariance of the epistemic attitude still holds. See Stewart Cohen, "How to Be a Fallibilist," *Philosophical Perspectives*, 2, *Epistemology*, James Tomberlin, ed., 1988. Context-sensitive acceptance is a central idea in much of the work of Henry Kyburg and Isaac Levi. A related point is that we tend to have vaguer beliefs when only small matters of concern are involved, than when we are confronted by important ones. But this reflects differences in the efforts we make to resolve vagueness on different occasions, rather than simultaneous possession of different beliefs for different stakes. (Thanks to Bernie Kobes for discussion of this, though he may disagree.) Kyburg, "Bets and Beliefs," p. 71, points out that we choose/wager differently on vague beliefs than on precise beliefs, but gives us no reason to suppose that we have the different beliefs simultaneously.

seems a kind of quasi-behavioristic move to focus on belief *qua* guide to action. But the further move sometimes made, taking betting odds to define partial belief, is more so, and it is inessential to Bayesian theory. We can regard partial belief as a dispositional state that influences deliberation and choice, and we can discuss partial belief by citing its real or hypothetical effects, while allowing that the effects we cite are good guides to the belief only usually, or under certain conditions. The *DBA* illustrates features of belief by assuming the conditions are met.

One more supposition: The value of each wager, *e.g.* a wager on *p*, lies in its payoffs, and not on any inherent value attaching to *p*'s obtaining (or not). We should *not* claim that this is so in general, when betting in real life. (The betting scenario is not about real-life betting, except insofar as it is a kind of acting.) Instead, this supposition is one of the main points of the betting scenario, which is supposed to illustrate, with a manageable stand-in for the complexities of real deliberation and choice, the interplay of action-guiding belief and choice. Actions are represented by wagers, ends by their payoffs. The advocate of the *DBA* does not ask us to believe that all choices are bets, nor that actions cannot carry inherent value. He does suppose that we are familiar with choices and actions that are like this, enough so that the illustration makes sense. One further point: the objection that the *DBA* is undermined by the possibility that a believer is well rewarded for violating the probability rules (or well punished for conforming to them), so that the believer will rationally incur the cunning bettor's sure penalty for the sake of the greater reward, is another way of attacking only the illustrative device. The payoffs defining the betting quotients are meant to represent all relevant values. This assumption is not made because real betting (let alone real action—say, child-rearing) is always like that, but because we seek a clear and general illustrative model.¹⁶

To summarize, each of these suppositions underlying the *DBA* may be questioned. One way of challenging them, as unproven assertions that the betting scenario constraints must apply to all belief, choice, and action, mistakes the point of the *DBA* and the suppositions. Another way of questioning them remains—one could simply hold that the betting scenario in which they apply is so remote from anything that we really have or do in our believing, choosing, acting, that nothing we might say about the betting scenario matters; it sheds no light on belief, choice, and action. There is little hope of producing an argument to the effect that anyone with such doubts is demonstrably mistaken—and the *DBA* is not even an attempt to do so. To repeat what we said above, if one's criticism of the *DBA* is entirely of this sort, it amounts to a rejection of partial action-guiding belief, and depends not at all on the content of the *DBA*.

¹⁶ And recall our earlier point, that the norm recommending avoidance of divided-mind inconsistency is not justified by appeal to the norm that recommends acting so as to further one's ends

Additivity.

To establish that rational degrees of belief are probabilities, the *DBA* must consider each of the probability rules and show that an arbitrary violation is exploitable by the cunning bettor. The interesting rule, the one we concentrate on, is additivity.¹⁷ When additivity is violated, we have, for incompatible p, q ,

$$dob(p \vee q) \neq dob(p) + dob(q).$$

The cunning bettor's response to this is an offer to sell (buy) a bet on the disjunction $(p \vee q)$, and buy (sell) a pair of bets on p and on q for equal stakes, when the left-hand side is $>$ ($<$) than the right. Let Bp represent a bet on p , and suppose that a wager on x pays \$1 if x , and \$0 if not. The bets are bought at the price judged fair by the agent, $\$dob(x)$, and we have:

		Payoffs				Prices	
p	q	Bp	Bq	$Bp + Bq$	$B(p \vee q)$	$Bp + Bq$	$B(p \vee q)$
T	F	1	0	1	1	$dob(p) +$ $dob(q)$	$dob(p \vee q)$
F	T	0	1	1	1		
F	F	0	0	0	0		

¹⁷ In particular, finite additivity. We won't go into the (easy) recipes for violations of the two rules, 1) $dob(T) = 1$; and 2) for all p , $dob(p) \leq 1$. But the question may arise, how are those violations tied to divided-mind inconsistency? In the case of (1), we can point out that a fair betting quotient of other than 1 for a tautology T is an assessment that a bet paying \$1 (or 1 utile) if T , has value different from a gift of \$1 (1 utile), which comes to the same thing. When (2) is violated and we have a $FBQ > 1$ for p , we are judging that it is worth a sacrifice of b (\$ or utiles) for the privilege of paying a (\$ or utiles) if p . (Rather than paying for insurance, we pay a fallible arsonist to hurt us.) This conflicts with the idea that avoidable exposure to a possible *overall* loss (and no prospect of gain) is not something of positive value, worth paying for. Making this idea precise involves more attention than we will give here to the assumption that payoffs are bearers of value, and to the location and value of the *status quo*. On the former, see above. There may be more to this idea than a consistency principle, but its normative status is not exactly shaky. In any case notice that the rule in question, that $dob \leq 1$ and—in the presence of the other rules— $dob \geq 0$, has the least content among the three rules: it fixes the probability scale in a very convenient but still arbitrary way.

When the buying and selling is complete, the payoffs of the bets will cancel, and the net exchange is determined by the difference in the prices, which the cunning bettor has arranged in his favor (he sold high and bought low).

Notice that the wager on the disjunction $B(p \vee q)$ is, insofar as the amounts and conditions of the payoffs go, identical to the combination $Bp + Bq$ of wagers at equal stakes on p and on q . When that combination represents the sum $dob(p) + dob(q)$, then the agent who violates additivity can justifiably be accused of the absurdity Ramsey mentions: he gives to the same thing (the wager on the disjunction, which is the same as the combination of individual wagers) two evaluations which differ, in that additivity is violated and an inequality in one direction or the other holds. I take it that what Ramsey had in mind was, at least in part, this fact.

The identity of the disjunctive wager and this combination of wagers is not in doubt, nor is the inconsistency of evaluating them differently. Notice that this remains so even if the bets are changed so that their costs and payoffs are no longer quantities of some commodity (dollars, pounds of gold, etc.), but are instead quantities of value, should we know how to measure those. They are the same bet, and however their value depends on the values of the payoffs, consistency demands that they have the same value.

But doubts can be raised about whether violations of additivity must involve inconsistency. So far we have just seen that the book $Bp + Bq$ is identical to the wager $B(p \vee q)$. Supposing again that wagers pay 1 and that $dob(p)$ is the fair price for such a wager on p , which we will also write as $V(Bp)$, additivity requires that the sum of the prices of the individual bets equals the price of the book/disjunctive bet:

$$V(Bp) + V(Bq) = V(Bp + Bq) = V(Bp \vee q) \quad (A)$$

The second equality is what obviously follows from the irrationality of divided-mind inconsistency. But we must have the first equality as well, to get the desired equality between the leftmost and rightmost expressions in (A). Why should the first equality hold? When goods are combined, the value of their combination need not be the sum of their values individually. And if we are not entitled to presuppose the first equality, there is a gap in the *DBA*: The argument claims that an agent with nonadditive evaluations of the bets (nonadditive degrees of belief) will judge acceptable a book producing sure loss. Its underlying significance is supposed to be that such acceptance is a mark of inconsistency. But neither must be so: The book involves only the sum bet $Bp + Bq$ (in further combination with $B\sim(p \vee q)$). In the absence of additivity this shows nothing about the values of Bp and Bq individually. The agent, when offered the Dutch book, will evaluate it according to the values of the sum bet and the bet on the disjunction—and these can be in perfect agreement even though the first equality in (A) fails to hold. Similarly, no inconsistency can be attributed, since no principle of consistency forces values to sum up as in (A).

The crux of the objection then, is this: I have $dob(p)$, $dob(q)$, and $dob(p \vee q)$. When you ask me to evaluate a wager on p , taken by itself, I assess it at price $dob(p)$. Similarly for q . When you ask me to evaluate a wager on $(p \vee q)$, or what comes to the

same thing (recall p and q are incompatible), a pair of wagers at equal stakes on p and on q , I evaluate it either way at the same price. I will consistently use $V(Bpvq)$ and escape the Dutch book, even while violating (A). So the *DBA* fails as a defense of additivity and probability; it fails as a criticism of additivity violations.¹⁸

We might have made the point in terms of sequential choice: In following his recipe for constructing the Dutch book, the cunning bettor will offer to sell me first a bet on say, p , and then a bet on q (to be followed by an offer to buy a bet on their disjunction). I will buy the bet on p , for $dob(p)$, but *then*, when a bet on q is offered my evaluation of it will be made in the light of my previous acceptance of the bet on p . This might well affect my evaluation of the bet on q (e.g. the exposure to risk on p may change the risk I am willing to run on q). In fact, my evaluation of that bet (call it $V(Bq\text{-after-making-}Bp)$) is not my simple $dob(q)$; but it is the increment by which $V(Bp + Bq)$ exceeds $V(Bp)$.

This way of putting the objection may help make it clearer, and it may lend plausibility to the point, but it is really out of place. Recall that the *DBA* defends a *synchronic* norm: simultaneous beliefs are to be probabilities; simultaneous evaluations are to be consistent. To be relevant, a criticism must involve non-additive synchronic beliefs, and simultaneous evaluations that violate (A). So the plausible view that making some bets leads to adjustments in making later bets is not relevant. The critic may still point to the possible violation of (A) as a gap in the *DBA*, but to be a problem for the *DBA*, the violation must be synchronic: the *dobs* are arranged, at a single time t , so as to violate (A).

That this can clearly happen when we evaluate bets is shown by cases of a familiar kind: Bus fare is \$1, and I have \$2. But I'd rather take a cab, which costs \$6. If I am offered a wager in which I risk \$1 on a fair die coming up 6, to win \$6, I am willing to take the bet. Similarly for a wager at the same stakes on the die showing a 5. But I am not willing to take both wagers at those prices, since losing the bus fare means I have to walk. The sum of the values of the individual bets exceeds the value of the sum bet.¹⁹

What is brought out here is a clear failure of additivity in the payoff goods (\$). But the dollars do not measure all the relevant value in this example, so the full payoffs of the wagers really involve more. The ratios of payoff dollars do not reflect the real odds that guide my choice (The sum bet's possible \$6 gain, i.e. a cab ride and change, does not weigh in at three times the value of a \$2 loss plus a walk home.) A simple version of the betting scenario does ask us to regard ratios of payoffs in dollars as real guides to choice. Or, better, it asks us to accept this as a useful general model of ends,

¹⁸ This line of criticism is well presented in Schick, "Dutch Bookies and Money Pumps."

¹⁹ A case in which I am willing to overpay for insurance with broad coverage, compared to what one would pay for each of two narrower, nonoverlapping policies alone, illustrates the inequality in the other direction. See remarks on this case at the end of the present section.

actions, and belief. But we are familiar enough with objections to the idea that value is generally represented by quantities of some commodity, that this model may well strike us as inadequate.²⁰

There are advantages, then, to a more sophisticated betting scenario, where wagers can be specified with payoffs in quantities of value, rather than piles of dollars. To the obvious question, how are quantities of value measured, we can distinguish two lines of response. One line interprets the question as a challenge to make sufficient sense of quantities of value so that the betting scenario and *DBA* carry some force as an illustration of belief and action. A more demanding challenge, and one on which many critics of the *DBA* may insist, is to provide a general demonstration that value is measurable, and measurable in a non-arbitrary way. The first challenge may or may not be easier to meet; a very prevalent view is that the only hope of meeting the first challenge is by providing a theory strong enough to meet the second, in which case that further theory tells us all that is really of interest (as when that theory provides an account of belief along with one of value). Adherents of this view are divided in their opinions about whether the subjective theory of rational preference does meet the second challenge. So there are both Bayesian and non-Bayesian criticisms of the *DBA* along these lines.

We will postpone consideration of the second challenge, and the relative merits of the *DBA* and the theory of rational preference, until the next section. How much can be said about the first challenge without appeal to full-blown utility theory? The following strategy can be tried: appeal to our familiarity with rough measures of value, and to the fact that we do manage, when we make choices, to weigh the relative importance of ends that are not pure commodities. Then ask, what plausible bases can be found for additivity failures? And how badly do they undermine the betting scenario? At best, this approach would not yield a general argument that additivity failures cannot reasonably arise, since it will not attempt to catalog and demonstrably rule out all possible sources. But the approach has some potential, given the interpretation we put on the *DBA*.

Suppose then some rough measure of value, without assuming a full utility theory. The betting scenario specifies wagers that pay off in quantities of value, and ratios of these quantities are betting quotients; an agent's fair betting quotients represent his degrees of belief. Should the degrees of belief be additive? Why not? Consider further our example: the value of the wager on the die showing 6 is increased by including nonmonetary goods in its payoffs (I might win a cab ride, at the risk of the expendable \$1). Similarly for the wager on the die showing a 5. Plausibly neither bet reflects my fair betting quotient, since I would agree to risk more than the \$1 for a chance

²⁰ The desirability of some move to quantities of value is already suggested, prior to these complications, by the familiar point that the value of a heap of a commodity is not linear with its size, and the assumption that fair betting quotients are insensitive to the size of the stake. More on this below.

at its payoff (I would submit to listening to the bettor's bad poetry while waiting for my bus),²¹ or I would accept a lesser payoff (a \$4 gain is enough for the cab), or both. Adjust the payoffs so that they fit my fair betting quotient. Would a pair of bets on these two propositions, each with stakes that match the values of the adjusted bets, be barely acceptable to me?

Here's one reason why I might say no: risk of compounded losses weighs more heavily with me than additivity dictates. I just don't operate, in my wagering or more generally in my choosing, by valuing compounded losses at the sum of the values of their components. So the value of the sum bet, which carries a risk of compounded loss, is less than the sum of the values of the component bets.²² If this is my reason, then it is worth noticing that I should already be worried by, I should already object to, an assumption that came earlier: namely, the assumption that my betting quotients are stable (though my choices may change) under shifts in the size of the stakes. If the risk of compounded loss on a combined bet carries extra weight, then ought not a doubled risk, when the stakes are doubled, do the same? My objection to additivity amounts to little or nothing more than an objection I should have made earlier to the assumed insensitivity of betting quotients to stakes. That assumption is rooted in the view that what I believe is not a function of the importance of the matter at hand, though my care in attending to it may be.²³

It is true that the response given so far to the objection that additivity is an unwarranted, significant assumption does not amount to a proof that it is innocuous, or that it surely holds in all our deliberation. But the response has some force: It proceeds from an assumption that what we believe is not a function of the stakes at hand, and builds that assumption into the betting scenario (either as an assumption of invariance of odds over all stakes, or as an assumption that the odds are invariant over a limited range of stakes, together with the view that beliefs are reliably indicated by wagers only in that range). It then points out that the most natural reasons for doubting additivity are also

²¹ But not if he's a Vagon.

²² Recall that the propositions in the example, and in the general case, are incompatible, so compounded gains are not at issue, when we buy the bets on the propositions. When we sell the bets, the question becomes my treatment of compounded gains. Remarks analogous to those that follow then apply. Also recall the betting scenario's assumption that values of wagers lie in their payoffs, rather than in any inherent value the truth of proposition may carry.

²³ The effect of the assumed invariance over stakes is even easier seen if we go back to the simple view that dollars measure value: if I would risk \$1 to win \$6 should a 5 show on the die, then insensitivity to the stake would have me willing to risk \$2 to win \$12. The question is whether there are good grounds for objecting to additivity which are not also objections to insensitivity to the size of stakes. It is interesting that they are not easy to find (or so I say).

reasons for questioning this stability of belief.²⁴ We can reject the assumption of stability, but in doing so we are more rejecting the whole approach to belief than finding a fallacy in the *DBA*.

Notice, however, that we have really only addressed the suggestion that value is subadditive, that $V(Bp) + V(Bq) > V(Bpvq)$. Superadditivity must also be ruled out for additivity to hold, and plausible bases for expecting superadditivity will also count against the betting scenario's suitability as an illustration. It is interesting that failures of additivity in this direction are much less often suggested by critics of additivity than are subadditive failures, and plausible examples are not so easily given, though no doubt examples can be constructed. (For one, see the insurance-policy example given above.) The natural motivation for $V(Bp) + V(Bq) > V(Bpvq)$ is the interaction of compounded losses, when buying the bets, or of compounded gains, when selling. But this doesn't apply to the opposite inequality. Why would $V(Bp) + V(Bq) < V(Bpvq)$? To fix the direction, imagine we are buying these bets. The bet on the disjunction does not get extra value, compared to the sum of the others, because its payoffs involve compounded gain—its payoff is the same as either individual bet. Compared to the individual bets, it risks compounded loss (*i.e.* it has a higher price), but that is no reason to overvalue it. The reason, if there is one, is the increased chance of winning, compared to that of just one of the individual bets. The assumed invariance of belief to the stakes does not speak to this; there appears to be no presupposition of the *DBA* already in force that undermines the suggestion that the values of the relevant wagers are superadditive. On the other hand, the question remains as to whether this possibility is so strong and pervasive that the betting scenario and the *DBA* lose all force as illustrations of belief and action. Unless a good case for pervasive superadditivity is made, it is open to the defender of the *DBA* to say, "You may insist that additivity still fails, for other reasons or for no particular reason, but this is now a weak sort of objection to an argument intended as an illustration, and having the merit of depending on assumptions that recognizably often apply."

Utility theory.

An advocate of the *DBA* wants to be able to use bets that pay off in quantities of value. The second, more demanding challenge raised in the last section calls for a thorough answer to the question, how is value measured, and measured non-arbitrarily? A thorough answer is available: we can give a general theory of subjective utility—in fact, there are many theories—founded on principles governing rational *preference*. This

²⁴ The connection between additivity and stake invariance is that the latter is equivalent to additivity of bets at different stakes on the same proposition. The connection to the stability of belief is *via* other assumptions we have mentioned: that payoffs are in quantities of value, and that all value lies in payoffs. This suggests that reasons, if there are any, for failure of additivity have to do with the way I treat various quantities of value; presumably they would hold when the quantities are attached to bets at different stakes on the same proposition.

is not the place to go into great detail about such theories; we will keep the discussion at a rather general level. Since we do not strive to notice and remark on all the varieties of these theories, our description may at times apply more straightforwardly to some theories than to others. As was true of our discussion of the *DBA*, some familiarity with utility theory is here assumed.

Best for the purpose of meeting the challenge is a utility theory that does not presuppose a theory of rational partial belief is already settled, before developing the theory of value/utility. (This is not to rule out theories that develop both, but in two stages; we are just setting aside theories that, without addressing rational belief, simply presuppose it.) A standard way of proceeding is to set out principles of rational preference, and then to demonstrate that whenever a system of preferences satisfies the principles, theories of the rational beliefs and values that accompany the preferences "fall out". Since the theory of value gives a utility measure that satisfies an expected utility principle, we not only have an answer to the challenge of measuring value, we have an answer that rules out failures of additivity, because expected utilities are additive.²⁵ The quality of the answer hinges on a number of factors, including of course the quality of the principles of rational preference. Not all sets of axioms for rational preference systems are equally good or equally general, but here we will not go into a detailed comparison of the axioms of the leading examples.²⁶

The general outline of one of these theories is this: A set of axioms governing preferences is put forward; these constrain the preference relation in suitable ways, and also guarantee that preference systems are sufficiently rich. A demonstration is given that for any system of preferences that satisfies the axioms (call such system *coherent*), two functions Pr and U with desirable features can be defined. The desirable features include Pr 's obedience of the rules of probability, U 's order-preservation of the entities ranked by the preference relation, and the conformity of both Pr and U to an expected utility principle EU (hence U 's additivity). The demonstration shows first the existence of a pair of such functions, which is certainly significant; it gives us reason to suppose that we have found a measure of value (since U is order-preserving and satisfies EU). But at least as interesting is the further demonstration that the pair $\langle Pr, U \rangle$ is unique (actually, unique to some interesting degree, up to some interesting set of other pairs); this is what justifies the view that the measure of value is not arbitrary. We could call

²⁵ Skyrms develops this point in "Coherence," p.228-232.

²⁶ There are many of these theories. Those best known to philosophers include Ramsey's in "Truth and Probability," Leonard Savage's in *The Foundations of Statistics* (New York: Wiley, 1954), and the Jeffrey-Bolker theory in Richard Jeffrey's *Logic of Decision* (Second edition, Chicago: University of Chicago Press, 1983). I sketch one for causal decision theory in "Conditional Preference and Causal Expected Utility," in *Causation in Decision, Belief Change, and Statistics*, W. Harper and B. Skyrms, eds.(Dordrecht: Kluwer, 1988), pp. 3-24.

this approach a preference-theoretic one, but instead let us call it a *decision-theoretic foundation* for utility and partial belief, or just a *decision-theoretic argument*.

Such arguments provide powerful results; their generality and rigor give them the best credentials for being the basis of Bayesian theory. And it is a popular view, especially among Bayesians, but also held by others, that though the seductive *DBA* is really fallacious, the decision-theoretic argument is unaffected by objections usually raised against the *DBA*. So, say the Bayesians, we can be satisfied with the theoretical foundation of our approach, while no longer really believing the *DBA*. This view is popular, but it is not universal. Of course critics object to the decision-theoretic argument for a variety of reasons, which cannot all be considered here. But we need not go very far into the structure of the decision-theoretic foundation to see the possibility that the failure-of-additivity objection to the *DBA*, pushed by a persistent critic, will go unanswered by the decision-theoretic argument.

An informal observation about the typical development of a decision-theoretic foundation will introduce the problem: A theory that yields the strong results we have described is a theory designed to do just that. And since the number of fully general, really obvious principles of rational preference is limited, the constraints a decision theory will impose on preferences go further. The further axioms are chosen with the goal of deriving the strong results in mind, including the result that the unique *Pr* and *U* measures satisfy a rule of expected utility. To the extent this is so, *EU* is more a constraint on what the theory is to be, than it is a derived property of the one and only pair $\langle Pr, U \rangle$ that rational preferences generate. To the extent it is so, then, the critic of the additivity of value will be skeptical of the defense of additivity that the decision-theoretic argument is supposed to provide. The non-arbitrary measure of value has not been derived so much as assumed, the critic may say, pointing to statements like these:

...But sometimes the number [the degree of belief] is used itself in making a practical decision. How? I want to say in accordance with the law of mathematical expectation; but I cannot do this, for we could only use that rule if we had measured goods and bads. But perhaps in some sort of way we approximate to it, as we are supposed in economics to maximize an unmeasured utility. The question also arises why just this law of mathematical expectation. The answer to this is that if we use probability to measure utility, as explained in my paper, then consistency requires just this law. Of course if utility were measured in any other way, *e.g.* in money, we should not use mathematical expectation.²⁷

...The impartiality axiom is obviously true, if the agent's preferences are represented by a pair *prob, des* ... and it plays an essential role But it is not the sort of assumption that is particularly plausible simply because

²⁷ Ramsey, "Probability and Partial Belief," p. 256.

we are taking prospects to be propositions. The axiom is there because we need it, and it is justified by our antecedent belief in the plausibility of the result we mean to deduce from it.²⁸

We should not be too quick to agree with the critic. For one thing, the "informal observation" above is more an objection *ad hominem* to the decision-theoretic argument, than it is a direct criticism. Real criticism should instead be directed toward the preference axioms that are actually used. This is a topic that deserves more attention than we can give it here, but it is not hard to point to axioms, and to defenses of axioms, which the critic can attack: Jeffrey's Impartiality axiom, Savages' sure-thing principle (his *P2*), and Ramsey's assumptions supporting his definition of degree of belief are each essential to the respective theories, and each likely targets of a critic who questions the universal additivity of value/utility. The money-pump argument, which is offered in defense of the transitivity of preference, is criticized by Schick precisely for its presupposition of value additivity.²⁹

The challenge we are exploring was to show that value is measurable in a nonarbitrary way, with an eye to responding to the *DBA* critic's suggestion that value need not be additive. The decision-theoretic foundation seemed to promise a decisive answer on both counts: value is measurable, and not just arbitrarily (uniqueness), and it must be additive after all. But the suspicion is now raised that the additivity is more assumed than demonstrated. If the suspicion is borne out, it is hard to see why the persistent critic of the *DBA* who issued the challenge should now think it has been met. In fact, there is room for the critic of Bayesian theory to pursue the objection, enlarging its scope to include the decision-theoretic argument.³⁰ A defender of Bayesian theory should think again about whether our response of last section to the less demanding first challenge has some force after all; and it is possible that the *DBA*, so defended, will lend helpful support to the decision-theoretic foundation.

²⁸ Jeffrey, *Logic of Decision*, p. 147. It should not be assumed that Ramsey's and Jeffrey's theories are unusual in this respect.

²⁹ "Dutch Bookies and Money Pumps," pp. 116-118. The money-pump argument has its attractions, but it is *prima facie* more clearly a diachronic argument (for the synchronic norm) than is the *DBA*, and so a better target for the value-independence objection. But one might interpret the money-pump argument in a fashion paralleling our interpretation of the *DBA* (as an illustration presupposing, in particular, diachronic stability of preference), and then Schick's similar criticisms of the two arguments are in agreement with the present claim that problems for the *DBA* also tend to be problems for the decision-theoretic argument.

³⁰ In "Coherence," p. 232, Skyrms, who is sympathetic to both arguments says, "...This is not to say that the qualitative assumptions of the representation theorems for probability and utility are above question; but rather to emphasize that criticism of the dutch book arguments, if it is to be more than superficial, must question utility theory and the representation theorems as well."

How would the expanded criticism go? The line of attack was against the *DBA* as an underpinning for subjective probability; *i.e.* for the norm that partial beliefs obey the probability rules. Now, let us suppose, the critic asserts that the attempt to turn back the attack with a powerful decision theory showing that values must indeed be additive has failed. To which a Bayesian willing to give up the *DBA* replies, "Oh well—but the *DBA* is not really my justification for the theory of partial belief. The real justification is the decision-theoretic foundation, showing that rational preference is uniquely representable by Pr and U . The Pr function is probability, and it represents rational partial belief. That my account is not arbitrary wishful thinking is supported by the uniqueness result."

But now the critic asks, "Why should we think so? What supports this interpretation of the measure Pr ? It may be a measure attached to propositions, and it may be unique, but what makes it a measure of belief?"³¹ What reply can be given to this? The decision theory's formal representation theorem itself does not give one. The Bayesian can point to his particular preference axioms and the entities to which they apply, and try to argue the plausibility of taking Pr as degree of belief. On this score, Ramsey's theory probably fares as well or better than any (if we overlook the familiar problem about gambles that conflict with how we think the world works): Why should Pr be degree of belief? Because the measure Pr attaches to those propositions p appearing in gambles of the form (α if p , β if not), and it makes good sense that our rankings of the gambles depend on our beliefs in those propositions, in just the ways the axioms say. That is, they do if value obeys expected utility—when we say it makes good sense, we have in mind uncomplicated cases where *EU* is most appropriate ...

I suggest that we would find, should we continue to push this exchange, that the Bayesian's best attempts to defend his interpretation of the measures falling out of the preference axioms will mimic the moves that we made in the previous sections to defend the *DBA*. If so, why not take the *DBA* seriously as a motivation for subjective probability? Neither it nor the decision-theoretic foundation can demonstrate that we have the partial beliefs about which Bayesians theorize. But they can, in fact, lend support to each other: the decision-theoretic foundation provides an assurance of the generalizability of the theory. And if last section's response to the first challenge carries some force, the *DBA* as we have understood it provides motivation for the interpretation given to the measure Pr , particularly for theories (*e.g.* Savage's and Jeffrey's, it seems to me) in which the interpretation is less obvious than it is for Ramsey's.

Conclusion.

The recognition that the *DBA* is an illustration, rather than the sole basis for attributing irrationality, allows us to see it in a better light as a motivation for the Bayesian's constraints on rational degrees of belief. The decision-theoretic argument,

³¹ Kyburg questions the interpretation of Pr as degree of belief; see "Two World Views," *Nous* (1970): 337-348, reprinted in *Epistemology and Inference*, pp. 18-27.

powerful and interesting as it is, is nevertheless helped by the kind of illustration provided by the *DBA*, particularly with regard to the question, why think *Pr* measures degrees of belief? In this paper we have been more concerned with how the *DBA* should be understood, and with how it illustrates features of belief, than with arguing that we must accept its conclusion.

We will add to those rather mild conclusions a bit of moralizing, however: Demands that we assume nothing and prove strong conclusions, however the demands are disguised, are unreasonable. We will make better progress by finding salient models for belief (and action), and exploring them. An appropriate response to such demands is to request from the critics something better. A Bayesian's admission that his theory can be improved, seen in these terms, is not thereby an admission that the current theory is nonsense. And the fact that nobody can (correctly) prove something from nothing does not make every theory equally good or bad.