Nozick, Ramsey, and Symbolic Utility

WESLEY COOPER
University of Alberta

I explore a connection between Robert Nozick's account of decision value/symbolic utility in *The Nature of Rationality*[1] and F. P. Ramsey's discussion of ethically neutral propositions in his 1926 essay 'Truth and Probability',[2] a discussion that Brian Skyrms in *Choice and Chance*[3] credits with disclosing deeper foundations for expected utility than the celebrated *Theory of Games and Economic Behavior*[4] of von Neumann and Morgenstern. Ramsey's recognition of ethically non-neutral propositions is essential to his foundational work, and the similarity of these propositions to symbolic utility helps make the case that the latter belongs to the apparatus that constructs expected utility, rather than being reducible to it or being part of a proposal that can be cheerfully ignored. I conclude that decision value replaces expected utility as the central idea in (normative) decision theory. Expected utility becomes an approximation that is good enough when symbolic utility is not at stake.

EXPECTED AND SYMBOLIC UTILITY

In *The Nature of Rationality* Nozick introduces the idea of symbolic utility, the subjective utility that an action may have intrinsically or for its own sake. This utility is subjective because it is determined by the agent's considered preferences rather than an objective ideal. It is still normative, however, because it stipulates that preference should be consistent with being fully informed and thinking clearly. What explains an agent's having preferences (or aversions) about an act itself may be its expressing or representing or meaning something. Whatever the explanation, the intrinsic preference signifies utility that attaches to the action or belief itself, not to its further outcomes. By contrast, the standard view is that the expected utility of an act is a product of the utility of its possible outcomes multiplied by the (subjective) probabilities of those outcomes:

$$\sum_{(i=1)}^{n} prob(Oi) \times u(Oi)$$

[1] Robert Nozick, *The Nature of Rationality* (Princeton, 1993).
[2] F. P. Ramsey, 'Truth and Probability', *The Foundations of Mathematics and Other Logical Essays* (New York, 1960).
[3] Brian Skyrms, *Choice and Chance* (Belmont, Calif., 2000).
[4] J. Von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior* (Princeton, 2004).

© 2008 Cambridge University Press
doi:10.1017/S0953820808003166
Utilitas Vol. 20, No. 3, September 2008
Printed in the United Kingdom

https://doi.org/10.1017/S0953820808003166 Published online by Cambridge University Press

subjective utility Subjective utility is utility disclosed by preference. In normative rational-choice theory this is considered preference.

subjective probability Subjective probability is probability relative to an agent's beliefs, especially as measured by techniques such as von Neumann and Morgenstern's and F. P. Ramsey's that use betting behavior to elicit the degree of probability that the agent assigns to an outcome.

expected utility (EU) EU is the product of the subjective utility of an act's outcome times the subjective probability of attaining it.

symbolic utility (SU) SU is (subjective) utility that an agent assigns to an act itself. In possible-worlds terms, an agent has a preference or aversion for a world simply in virtue of his or her performing that act in that world.

Figure 1: Utility and probability

Nozick is suggesting that the standard view of expected utility is misleading because it does not take into account the possibility of a decision-maker's considered preference or aversion for the acts that are available, in addition to the possible outcomes of those actions. These acts are not merely instrumental to outcomes, but rather they are intrinsically valenced, positively or negatively: their valence – that is, the preference–driven tendency to perform those acts – would not be extinguished if their instrumental value were believed nil. One would still be averse to burning the flag, still prone to tell the truth. Such acts seem to require *weight* that the standard view does not bestow. Some moral theories, notably consequentialist ones, also withhold this weight. The difference between an act and an omission, killing and letting die, does not matter because only consequences matter. There is a formal fit between the standard view and consequentialist moral theories, but also a gap between the demands of these theories and what EU-rational agents are willing to offer. This is the duality of reason that utilitarians such as Sidgwick wrestle with, the polarity of what's rational from the point of view of the universe, on one hand, and what's rational from the individual's point of view. One motivation for Nozick's suggestion (others will be noted below) is to render the gap narrower than the standard view can manage.

Confidence in the scientific credentials of subjective utility has been bolstered by the demonstration that it can be rendered objective by measuring an agent's disposition to make bets with the aid of dice, lotteries and the like. This measurement yields in principle an ordinal ranking of a decision-maker's strengths of desires and degrees of belief. It encourages the idea that a science of rational choice can be built up from subjective utility. It eases concern that subjective utility deals in

https://doi.org/10.1017/S0953820808003166 Published online by Cambridge University Press

Evidential expected utility (EEU) EEU replaces EU [Probability(Outcome) × Utility(Outcome)] with the probability of the outcome given the act. This probabilistic conditionalization may affect an agent's choice in Newcomb's problem.

Causal expected utility (CEU) CEU replaces EU with a causal-cum-probabilistic conditionalization. The probabilities relevant to CEU are restricted to what the agent more or less probably can bring about in the choice situation, which may affect an agent's choice in Newcomb's problem.

Decision Value (DV) DV is the (subjectively) weighted sum of EEU, CEU, and SU.

Figure 2: Two kinds of expected utility, and decision value

the intractably inner or arbitrary. Von Neumann and Morgenstern set out the theory for measuring these bets, but earlier work by Ramsey will be credited here with providing a measurement that does without the mentioned technologies. Ramsey's process filters out symbolic utility in deriving the conception of expected utility, suggesting that subjective utility is implicated at the deepest foundations of decision theory.

EVIDENTIAL AND CAUSAL EXPECTED UTILITY

Nozick explores symbolic utility for implications about decision theory's ideal of rational choice, which is centred on maximizing expected utility. This is part of a larger project of recommending a *factored* ideal of rational choice, in which maximizing expected utility is replaced by maximizing decision value. DV maximizes the weighted sum of two kinds of expected utility and SU. It is act-sensitive as well as consequence-sensitive. His approach to the Prisoner's Dilemma turns on this factoring, as it implies the rationality of taking into account utilities other than the expected utilities in the payoff matrix for the PD, notably the symbolic utility of expressing oneself as a cooperative person, by choosing the 'optimal' action of doing what's best for both prisoners collectively, instead of the 'dominant' action of doing what's best for me whatever the other prisoner decides. Also, expected utility is factored into the weighted sum of (1) purely probabilistic expected utility, or what he calls 'evidential expected utility', and (2) causal expected utility, which calculates the probabilities of outcomes conditional exclusively upon what the agent can make happen in the choice situation. This further factoring informs his approach to Newcomb's Problem, calling upon the rational agent to switch between purely probabilistic and causal/probabilistic reasoning depending on whether there is much to gain by reasoning in causal/probabilistic terms ('taking both boxes') or reasoning in purely probabilistic terms

https://doi.org/10.1017/S0953820808003166 Published online by Cambridge University Press

(taking only the opaque box that may have a million dollars inside). If there is almost a million dollars in the transparent box, take both boxes; if there is only a penny in the transparent box, take only the opaque box. Both of these situations have the formal structure of Newcomb's Problem, but they differ in their 'cash value' from a decision-value perspective.

To summarize, the three accounts of probability in expected utility can be formulated as follows.

1. EU (unconditional expected utility)

$$\sum_{(i=1)}^{n} prob(Oi) \times u(Oi)$$

2. EEU (evidential expected utility, EU as probabilistically conditional upon the choice of action)

$$\sum_{(i=1)}^{n} prob(Oi/A) \times u(Oi)$$

3. CEU (causal expected utility, EU as causally conditional upon the choice of action)

$$\sum_{(i=1)}^{n} prob(Oi//A) \times u(Oi)$$

(The double slash indicates causal influence of action upon outcome)

The general picture that emerges is that preference explains utility, and utility together with probability explains expected utility. Utility explains symbolic utility as a special case. Expected utility explains evidential expected utility and causal expected utility as special cases, and these latter together with symbolic utility explain decision value. This general picture is illustrated diagrammatically in Figure 3.

APPLICATIONS OF EEU/CEU FACTORING AND SU

Newcomb's Problem

DV's EEU/CEU factoring is particularly relevant to Nozick's solution to a long-standing bone of contention between causal theorists and those who take a EEU approach to Newcomb's Problem: A Very Good Predictor puts a million dollars in an opaque box prior to your choice just in case he predicts that you will take only that box, otherwise he will put nothing in it. You know about this. Also you are able to see a

https://doi.org/10.1017/S0953820808003166 Published online by Cambridge University Press

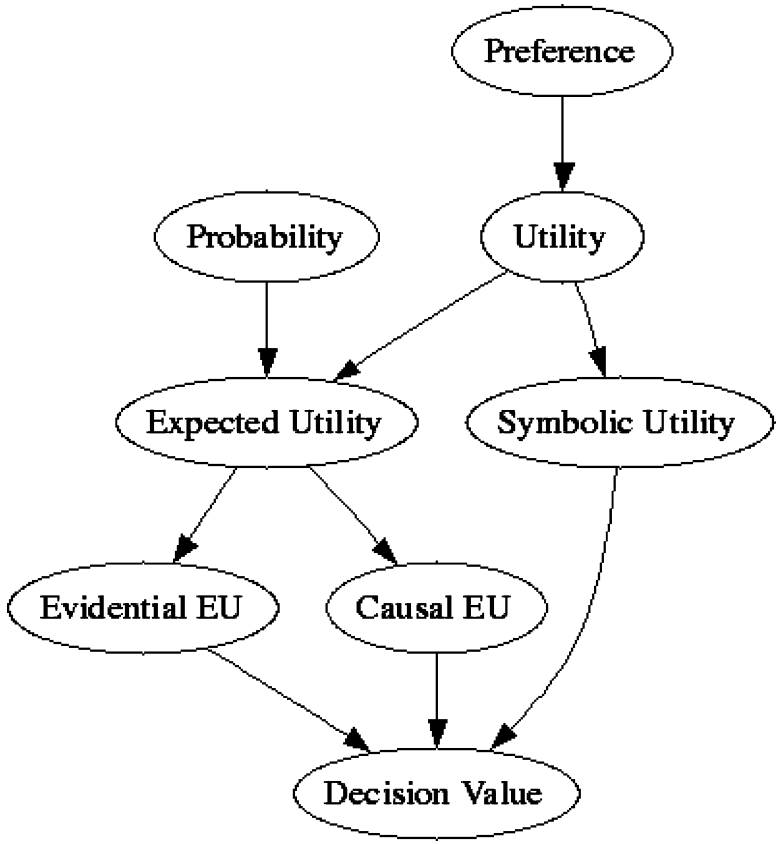


Figure 3: The arrows stand for 'is explanatorily more fundamental than'. This image will be adjusted when the argument is complete

Newcomb's Problem (NP) NP shows how EU reasoning tends to polarize into EEU and CEU reasoning. Factoring reveals amiddle-way solution.

Prisoner's Dilemma (PD) The PD shows how EU reasoning leads to a conflict between individual and collective rationality. Subjective utility reveals a solution that doesn't depart from individual maximizing of utility.

Magic Numbers and Algebraic Form Nozick's diagnosis of both NP and PD is that they initially pump one's intuitions with 'magic numbers' about utilities, dollars, years in jail, and the like. However, they have an algebraic form that abstracts from such magic numbers. Different numbers can maintain the abstract form of the NP or PD while revealing the appeal of DV's factoring.

Figure 4: Newcomb's Problem, Prisoner's Dilemma, Magic Numbers and Algebraic Form

https://doi.org/10.1017/S0953820808003166 Published online by Cambridge University Press

thousand dollars in a transparent box. EEU tells you, on the evidence of the Predictor's impressive record, to take only the opaque box. This is the rational choice, the choice that reflects conditional probabilities. CEU on the other hand tells you to take both boxes, on the grounds that the Predictor has put the million dollars in the opaque box or he hasn't; the only causal variable at play is your choice. So you might as well take both boxes. *That* is the rational choice, the 'dominant' choice that ensures the decision-maker does best whatever the Predictor has done.

Parsing EU into EEU and CEU slips through the dilemma of choosing between conditional probabilities and dominance. It allows the decision-maker to give more or less weight to either, depending especially on how much is in the transparent box. If there is a million dollars minus a loonie in the transparent box, one might well give great weight to CEU. There is little to lose, only a loonie. If there is only a penny in the transparent box, one might assign great weight to EEU. There is little to gain by taking the transparent box, only a penny. The weighted *decision value* account offers an alternative to the single-weight expected-utility account of rational choice, a middle way between the EEU and CEU strategies.

$$\mathrm{DV(A)} = \mathrm{W}_c \times \mathrm{CEU(A)} + \mathrm{W}_e \times \mathrm{EEU(A)}$$

Stages of Nozickian enlightenment about Newcomb's Problem

The learner's first conception of Newcomb's Problem presents two boxes with different amounts of money. Not any arbitrary amounts would create the Problem, but the amounts chosen do so. These 'magic numbers' may fixate the learner's attention on the given amounts, creating a frame that prevents him from exploring Nozick's solution.

For the advanced student the magic numbers give way to algebraic variables, standing for any sums such that the amount in the opaque box is more than the amount in the transparent box.

\$1M?	\$1K

Figure 5: Newcomb's Problem with Magic Numbers: There may or may not be a million dollars in the opaque box, and there are a thousand dollars in the transparent box

$x >?$	y

Figure 6: Newcomb's Problem with algebraic variables

https://doi.org/10.1017/S0953820808003166 Published online by Cambridge University Press

\$1M?	\$1M-\$1

Figure 7: Revenge of the Magic Numbers: some values of the algebraic variables allow a CEU solution: You have little to lose, so take both boxes

\$1M?	1 cent

Figure 8: Revenge of the Magic Numbers: some values of the algebraic variables allow an EEU solution: You have little to win, so take just the opaque box

The respective CEU and EEU strategies are fixed as long as x is greater than y. The causal theorist chooses the 'dominant' action, taking both boxes no matter how much or how little is in the transparent box; the evidential theorist takes only the opaque box, no matter how much or how little is in the transparent box.

The algebraic level of insight gives way to DV enlightenment when one assigns weight to different strategies, causal or evidential, according as there is a little or a lot in the transparent box.

Prisoner's Dilemma

Nozick recommends symbolic utility as a solution to another long-standing problem, the Prisoner's Dilemma. The payoff boxes in a (two-person, one-shot) PD matrix assign expected utilities to cooperating with the other prisoner or not cooperating (ratting), such that ratting dominates cooperation for both prisoners despite the fact that mutual ratting is non-optimal. They would do better if they were both to keep quiet.

The matrix gives expected utilities for act–outcome pairs, but it says nothing about the action itself, non-instrumentally or intrinsically. Symbolic utility does just this, and Nozick argues that it's rational to assign great weight to subjective utility when the downside of cooperation is not too high: when, for instance, it means only a few more minutes in jail if the other prisoner chooses to rat. Conversely, if the downside is an additional ten years in the slammer, it's rational to give little weight to subjective utility and much to EU. The full decision-value formula creates an explicit space for weighting of symbolic utility as well as whatever expected-utility weightings you favor.

$$\mathrm{DV(A)} = \mathrm{W}_c \times \mathrm{CEU(A)} + \mathrm{W}_e \times \mathrm{EEU(A)} + \mathrm{W}_s \times \mathrm{SU(A)}$$

https://doi.org/10.1017/S0953820808003166 Published online by Cambridge University Press

PD #1	Player 2: cooperate	Player 2: don't cooperate
Player 1: cooperate	10 yrs, 10 yrs	15 yrs, 1 yr
Player 1: don't cooperate	1 yr, 15 yrs	5 yrs, 5 yrs

Figure 9: A Prisoner's Dilemma with Magic Numbers

PD #2	Player 2: cooperate	Player 2: don't cooperate
Player 1: cooperate	x1, x2	w1, z2
Player 1: don't cooperate	z1, w2	y1, y2

Figure 10: A Prisoner's Dilemma with algebraic variables

Stages of Nozickian enlightenment about the Prisoner's Dilemma

The learner's first conception of the Prisoner's Dilemma presents a payoff matrix with numbers, denoting years in jail or some other representation of (dis)utility. Not any numbers create the dilemma. The chosen numbers 'magically' do so.

For the advanced student the magic numbers give way to greater or lesser expected utilities, where

$$w1 > x1 > y1 > z1$$

and

$$w2 > x2 > y2 > z2,$$

as shown in PD #2.

The algebraic level gives way to DV enlightenment, a partial solution to the PD when cooperation is only mildly punished by the expected-utility payoff boxes and the symbolic utility of cooperation (not shown in the boxes, because it is not an expected utility) is sufficient to outweigh the punishment. Let 'd' stand for days and 'm' for minutes, and assume that the positive symbolic utility of being cooperative outweighs an extra minute in jail. This gives PD #3a.

And at the more abstract level of utility, assume that the symbolic utility of being cooperative is +3. (Symbolic utility is not shown in the payoff matrix for outcomes because it is not a function of an action's *outcomes*.) Although non-cooperation is still the 'dominant choice' that yields the best payoff whatever the other player does, adding subjective utility of +3 to the payoff for cooperation makes it the rational choice.

https://doi.org/10.1017/S0953820808003166 Published online by Cambridge University Press

PD #3a	Player 2: cooperate	Player 2: don't cooperate
Player 1: cooperate	4d, 4d	4d + 1m, 4d − 1m
Player 1: don't cooperate	4d − 1m, 4d + 1m	5d, 5d

Figure 11: Revenge of the Magic Numbers: some areas of the algebraic space make cooperation rational

PD #3b	Player 2: cooperate	Player 2: don't cooperate
Player 1: cooperate	−1, −1	+1, −3
Player 1: don't cooperate	−3, +1	−2, −2

Figure 12: Revenge of the Magic Numbers: in terms of utility

Ethically Neutral A proposition is ethically neutral just in case one is indifferent between a world in which it is true and a world in which it is false.

Not Ethically Neutral A proposition is not ethically neutral just in case one is not indifferent between a world in which it is true and a world in which it is false.

The 'external' model SU happens to figure in an attractive solution to the PD, so it should be accepted as part of a DV alternative to EU.

The 'reductive' model SU reduces to EU. It should be possible to analyze the symbolic utility of an action in terms of expected utility.

The 'internal' account SU figures in the foundations of expected utility, belonging to its explanation.

Figure 13: Ramsey's definitions, and three models of the EU–SU relationship

THREE MODELS OF THE RELATIONSHIP BETWEEN SUBJECTIVE UTILITY AND EU

When an action has symbolic utility in Nozick's sense, the proposition that describes it is what F. P. Ramsey called *not ethically neutral*. In the 1926 essay 'Truth and Probability' he aimed to isolate such propositions in order to apply ethically *neutral* propositions to the foundations of probability, about which more later. Ethically *non-neutral* propositions are 'such that their truth or falsity is an object of desire to the subject', or 'not a matter of indifference'. Symbolic utility in Nozick's sense is tantamount to ethical non-neutrality in Ramsey's. Nozick goes beyond 'not a matter of indifference' to explain why this might be so. Specifically, the

https://doi.org/10.1017/S0953820808003166 Published online by Cambridge University Press

agent may not be indifferent to the act's expressive, representative or meaningful character. But at bottom Nozick's and Ramsey's conceptions are the same. Ramsey's interest in such propositions was different from Nozick's, however. He sets out the ideas of ethically neutral and ethically non-neutral propositions in the following passage. He states them using Wittgenstein's theory of propositions, noting that it would probably be possible to give an equivalent definition in terms of any other theory. I quote Ramsey at length:

> Suppose next that the subject is capable of doubt; then we could test his degree of belief in different propositions by making him offers of the following kind. Would you rather have world α in any event; or world β if p is true, and world γ if p is false? If, then, he were certain that p was true, simply compare α and β and choose between them as if no conditions were attached; but if he were doubtful his choice would not be decided so simply. I propose to lay down axioms and definitions concerning the principles governing choices of this kind. This is, of course, a very schematic version of the situation in real life, but it is, I think, easier to consider it in this form.
>
> There is first a difficulty which must be dealt with; the propositions like p in the above case which are used as conditions in the options offered may be such that their truth or falsity is an object of desire to the subject. This will be found to complicate the problem, and we have to assume that there are propositions for which this is not the case, which we shall call ethically neutral. More precisely, an atomic proposition p is called ethically neutral if two possible worlds differing only in regard to the truth of p are always of equal value; and a non-atomic proposition p is called ethically neutral if all its atomic truth-arguments are ethically neutral.[5]

The next stage of the argument shows that Ramsey's account of ethically non-neutral propositions supports the third of the following three interpretations of the relationship of symbolic utility to expected utility.

1 **DV as an optional alternative to EU** DV might replace EU at the center of decision theory, as Nozick proposed, because it solves long-standing problems such as Newcomb's Problem and the Prisoner's Dilemma. This might be called an *external* account of the relationship. It does not draw on the foundations on which expected utility is built, but rather declares those foundations to be inadequate and proposes an alternative external to those foundations.
2 **DV as reducible to EU** DV might be a 'high level' apparatus that could in principle be reduced to the 'low level' language of expected utility, comparable to the relationship between a high-level programming language and machine language.

[5] Ramsey, 'Truth and Probability', pp. 18–19.

https://doi.org/10.1017/S0953820808003166 Published online by Cambridge University Press

The Von Neumann–Morgenstern Test Determine subjective utility by reference to willingness to bet on some gambling technology in the external world.

The Ramsey Test Determine subjective utility by reference to willingness to bet where the truth or falsity of an objectively neutral proposition is variable.

scalar equivalance Both tests are scaling a decision maker's utilities (between 0 and 1, say)

parsimony in explanation The Ramsey test is more explanatorily fundamental because of its relative simplicity, notably its not requiring betting technologies in the external world.

Figure 14: The Von Neumann–Morgenstern and Ramsey tests for utility

3 **DV as explanatory of EU** The relationship between decision value and expected utility might have been implicit in the foundations of decision theory from the beginning, so that DV is neither simply an external replacement for EU nor high-level, but rather something that has to be recovered from the foundations on which EU was built. On this third interpretation EU may be useful as an approximation when ethical non-neutrality is not at stake, but otherwise the theory of rational choice requires DV-style factoring. This third relationship will be explored here.

RAMSEY VERSUS VON NEUMANN/MORGENSTERN

The Von Neumann–Morgenstern method picks the best payoff for the decision problem and gives it by convention utility 1, and likewise it gives the worst payoff utility 0. Then the utility of a payoff in between is determined by some gambling technology in the external world, such as dice or a lottery or a wheel of fortune, a chance device for which the chances are known. The utility of a payoff P is determined by the gamble with worst and best payoffs as possible outcomes that has value equal to P. To consider a case of interest to Searle that will be discussed later, a decision-maker's situation might be structured such that the payoffs are ranked from best to worst as follows:

$$\textit{Graceland} > \textit{little deuce coupe} > \textit{blue suede shoes} > \textit{son's death}$$

Gambler is indifferent between (A) a lottery ticket with 4/5 chance of owning Elvis's lovely property, Graceland, and 1/5 chance of son's death; and the little deuce coupe of Beach Boys fame, for sure. Or (B) a lottery ticket that gives 1/2 chance of Graceland and 1/2 chance of

https://doi.org/10.1017/S0953820808003166 Published online by Cambridge University Press

son's death, and one that gives Elvis's blue suede shoes, for sure. So Gambler's utility scale looks like this: Graceland 1, coupe .8, shoes .5, and son's death 0.

This gambler isn't showing much regard for his son's life, so consider instead ticket A with 4/5 chance of Graceland and 1/million chance of son's death, and the coupe for sure; and ticket B with 1/2 chance of Graceland and 1/billion chance of son's death, and one that gives shoes for sure. These gambles, which show higher regard for the son's life, also complicate the presentation of the utility scale (because the alternatives of Graceland and son's death don't between them exhaust the probability space between 0 and 1). If the intermediate values are measured by the distance from ownership of Graceland, the utility profile remains 1, .8, .5, 0. If measured by the distance from the son's death, it becomes 1, .999999999, .999999, 0.

So consider instead ticket A with million-1/million chance of Graceland and 1/million chance of son's death, and the coupe for sure; and ticket B with billion−1/billion chance of Graceland and 1/billion chance of son's death, and one that gives shoes for sure. Here the alternatives of Graceland and son's death exhaust the probability space between 0 and 1. These gambles give the utility scale 1, .000001, 000000001, 0. The utility of Graceland, given the revised probabilities, is relatively much higher.

Consider now Ramsey's method, which identifies propositions which are like the von Neumann/Morgenstern coin flips and lotteries in having instrumental value but no intrinsic value for the decision-maker. An ethically neutral proposition p doesn't affect preferences for payoffs. One is indifferent between payoff B with p true and B with p false. As Skyrms observes, 'The nice thing about ethically neutral propositions is that the expected utility of gambles on them depends only on their probability and the utility of their outcomes. Their own utility is not a complicating factor.'[6]

An ethically neutral proposition H has probability 1/2 for the decision-maker if there are two payoffs, A;B, such that he prefers A to B but is indifferent between the two gambles

1. Get A if H is true, B if H is false.
2. Get B if H is true, A if H is false.

If the gambler thought that H was more likely than not-H, he would prefer gamble 1. If he thought not-H more likely, he would prefer gamble 2. 'For the purpose of scaling the decision-maker's utilities,' Skyrms

[6] Skyrms, *Choice and Chance*, p. 142.

https://doi.org/10.1017/S0953820808003166 Published online by Cambridge University Press

notes, 'such a proposition is just as good as the proposition that a fair coin comes up heads.'[7]

So proposition H can be used to scale the decision-maker's utilities instead of a proposition that a gambler is indifferent about the outcome of a lottery.

Von Neumann/Morgenstern permits inference of degrees of belief from utilities, enabling the definition of expected utility as the product of probability and utility, but so too does Ramsey. So a decision-maker's degree of belief in the ethically neutral proposition p is just the utility he attaches to the gamble *Get G if p, B otherwise*, where G has utility 1 and B has utility 0.[8]

Ramsey's method in the 1926 essay covers the same ground as the von Neumann-Morgenstern theory, but it depends only on the decision-maker's preferences. This makes it theoretically more fundamental than the von Neumann–Morgenstern approach. It makes fewer assumptions, in particular doing without the assumption of a technology for determining objective chances (lotteries, coin tosses, etc.), while having the same power to explain subjective probability, subjective utility, and expected utility.

As Skyrms observes, the von Neumann–Morgenstern theory is really a rediscovery of ideas contained in Ramsey's essay, which 'goes even deeper into the foundations of utility and probability'.[9] But then recognition of actions as having utility independent from outcomes isn't something introduced simply as a replacement for the EU conception of utility maximizing, more or less persuasive depending on one's views about Newcomb's Problem and the Prisoner's Dilemma. Nor is it reducible to EU, on the analogy of a high-level programming language to a lower-level language. Rather, it is at the foundations of utility and probability.

THE 'EGG-SHELL' CONCEPTION

This completes a defense of Option 3, the 'internal' account of symbolic utility: it figures in the foundations of decision theory. It is not fundamental in the sense that it is related to expected utility in the way that an egg yolk is related to an omelet. That would be the 'reductive' account, and indeed ethically neutral propositions contribute to Ramsey's explanation of expected utility in this reductive way. But ethically non-neutral propositions are fundamental in the way that an egg shell is related to an omelet, as in the maxim that

[7] Skyrms, *Choice and Chance*, p. 142.
[8] Skyrms, *Choice and Chance*, p. 142.
[9] Skyrms, *Choice and Chance*, p. 141.

https://doi.org/10.1017/S0953820808003166 Published online by Cambridge University Press

you can't make an omelet without breaking eggs. Their existence must be acknowledged and they must be filtered out in order for Ramsey's reduction to work. Subjective utility does not reduce to expected utility (option 2), neither do they amount merely to an external appurtenance (option 1). SU's deep involvement in decision theory is a reason – additional to its contributing a solution to the Prisoner's Dilemma and other applications to be reviewed below – to acknowledge DV and SU's role in it.

Below in figure 15 is a revision of the 'is explanatorily more fundamental than' image in figure 3. Figure 15 represents the 'egg-shell' conception of symbolic utility that emerges from the comparison to Ramsey's ethically non–neutral propositions. This conception supports the third or 'internal' account of SU: it belongs to the foundations of decision theory, from which expected utility is derived.

SEARLE'S CRITIQUE OF DECISION THEORY

John Searle holds that decision-theoretic models of rationality 'are not satisfactory at all', because 'it is a consequence of Bayesian decision theory that if you value any two things, there must be some odds at which you would bet one against the other'.[10] He continues:

> Thus if you value a dime and you value your life, there must be some odds at which you would bet your life against a dime. Now I have to tell you, there are no odds at which I would bet my life for a dime, or if there were, there are certainly no odds at which I would bet my son's life for a dime.[11]

Recall that decision theory must always filter ethically non-neutral bets from those that are neutral. With the filtering done, it can proceed with standard expected-utility calculations for neutral bets. When the betting is ethically non-neutral, however – when one has a preference or aversion for the world in which that bet takes place – the factoring and weighting apparatus of decision value is required. It is rational for Searle to refuse the bet, for his betting his son's life is not ethically neutral for him, and the weight he attaches to the negative symbolic utility of the bet is great; he is averse to it, sufficiently so that the aversion outweighs the payoff in a decision-value calculation. He maximizes DV, but not EU, by refusing the bet. And generally symbolic utility is a motivational *basket* that collects side-constraints, notably moral ones, on maximizing expected utility.

[10] John Searle, *The Construction of Social Reality* (New York, 1995), p. 138.
[11] Searle, *Construction*, p. 138.

https://doi.org/10.1017/S0953820808003166 Published online by Cambridge University Press

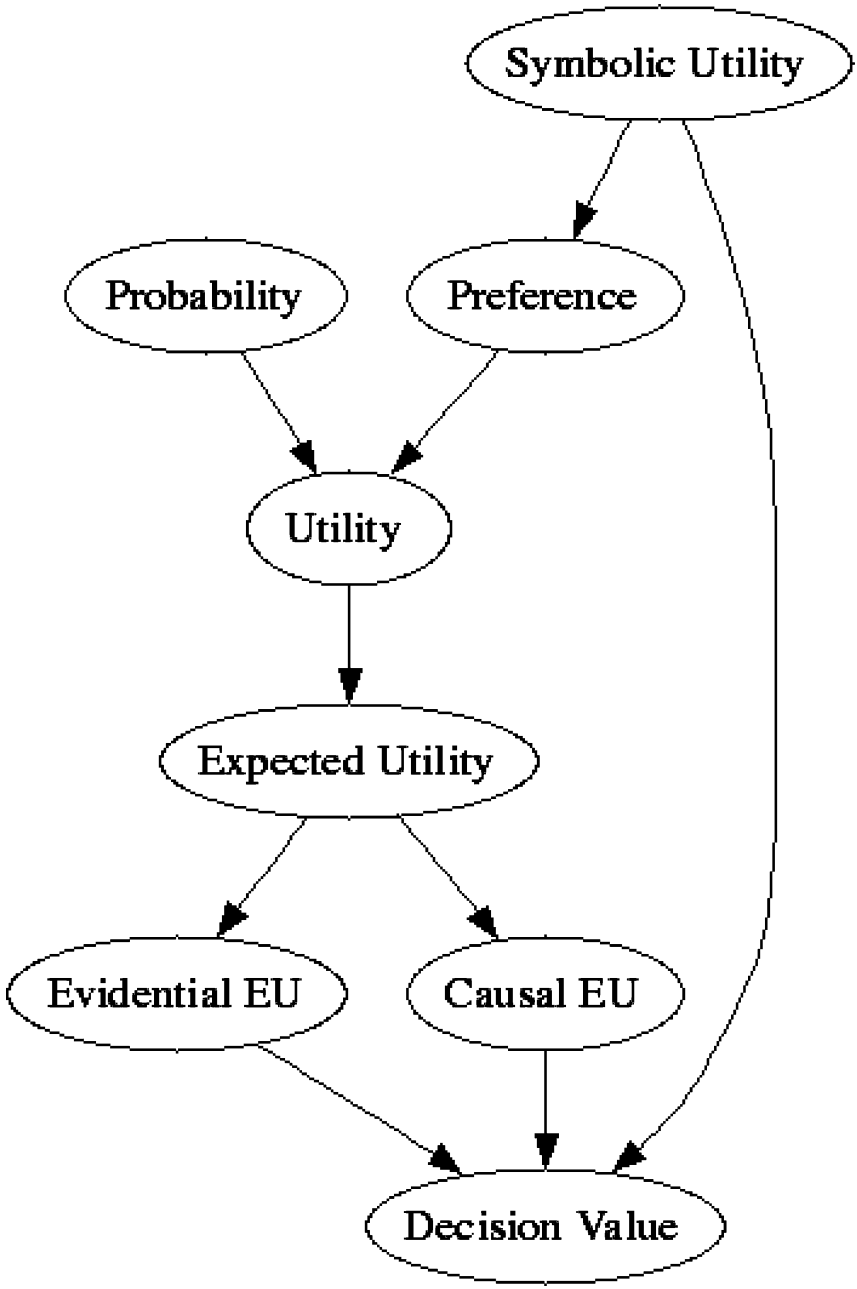


Figure 15: The 'egg-shell' conception of symbolic utility's relationship to expected utility

https://doi.org/10.1017/S0953820808003166 Published online by Cambridge University Press

A CONTRAST WITH BRINK'S OBJECTIVE UTILITARIANISM

David Brink defends what he calls objective utilitarianism both at the level of an individual's practical rationality and as a global moral principle. The goods that are 'intrinsically valuable' include reflective pursuit and realization of agents' reasonable projects, and certain personal and social relationships.[12] A rational agent maximizes these for his own life, and global rightness maximizes them for all lives. This doesn't guarantee convergence between self-interest and global rationality – the point of view of the individual and the point of view of the universe; but construal of self–interest in terms of objective utilities leads Brink to be optimistic about decreasing if not spanning the gap.

Symbolic utilities, though they are subjective rather than objective in Brink's sense, possess a similar gap-spanning quality. People often give weight to subjective utility about moral reasons that lead them to do the morally right thing (by their lights, and according to moral theories that support their moral beliefs). However, there is no overarching, transcultural conception of the good that informs SU, as there is for Brink's objective utilitarianism. Subjective utility gives weight to individuals' beliefs in such conceptions, as for instance when individuals' preferences are formed in a religious culture with a putatively objective moral credo. However, the DV/SU conception does not endorse (or reject) any such credo, but rather draws on the meanings available within a culture. Idiosyncratic symbolic utilities are particularly vulnerable to critique, by reference to a culture's standards, as immoral or uninformed or ill-considered. Under this pressure they are likely to change towards conformity with the culture's norms.

MAXIMIZING VERSUS SATISFICING

Like EU, DV is a maximizing theory. So something should be said about Satisficing (S), which aims at outcomes that are 'good enough' both at the level of the individual's practical agency and at the level of global moral rightness. Some decision theorists favor S because the maximizing alternatives require calculative rules that no human beings possess. But this seems to imply falsely that all calculation must take place in conscious mental life, whereas evolution and habit are capable of bearing much of the load without drawing on conscious resources. Still, real-world rational choice and belief may depart significantly from what's best to do and believe. Evolutionary,

[12] David Brink, *Moral Realism and the Foundations of Ethics* (Cambridge, 1989), p. 231.

https://doi.org/10.1017/S0953820808003166 Published online by Cambridge University Press

cultural and individual implementation of the ideal can be expected to impose S-like features, not because S is the ideal but because these features belong to the ideal's implementation. If ideally an agent is aware of all alternatives for action, the implementation would take into account that an agent's knowledge of the circumstances may fall short, not because of irrationality but because of limitations in the science and common-sense knowledge he has access to.

Information-processing with respect to what is known is also bounded, and the implementation would take this into account. The agent may need to use a 'stopping rule' that sets an 'aspiration level' that may pick an action different from the one that might be picked with knowledge of all relevant information and boundless information-processing capability. This is not because the ideal is satisficing rather than maximizing, but because maximizing in *this* context requires informational cherry-picking. Maximizing takes place relative to a Background (in Searle's sense) which shapes such dimensions of maximizing as 'relevant information' in the ideal formula's prescription that the rational agent takes into account all relevant information. Another example might be shaping the dimension of *calculations of probability*, in the ideal formula's conception of expected utility as a product of 'probability × utility'. Searle's objection to decision theory (see above), that there are no odds at which he would bet on his son's life for a dime, might best be understood as a stipulation about the background for maximizing: one has a moral reason for rejecting calculations of expected utility for bets on one's children's lives when these are gratuitous and offensive according to cultural norms. This qualification is necessary in order to permit probability calculations about sending one's children on airplane flights to visit relatives, etc.

AINSLIE'S PROBLEM, AND PERSONAL IDENTITY

How do our lives develop so that symbolic meaning attaches to our deeds? How does this meaning affect our identities? Consider the lives and actions of professionals. Professions evolve when a group of people becomes aware of a common interest in a skill and exercises collective intentionality in setting standards through constitutive rules which both define the profession more completely and assign symbolic utility to certain kinds of professional activity. In the academy, for instance, there are many rules defining honors and penalties, powers and permissions and privileges, and so forth. And typically there is a collective recognition of symbolic utility, positive or negative, that comes with an honor or penalty, apart from further consequences such as financial rewards or fines. This is not to say that all symbolic utility is attached to constitutive rules, however. To draw a parallel

https://doi.org/10.1017/S0953820808003166 Published online by Cambridge University Press

with chess, a brilliant gambit has high symbolic utility recognized by participants in that social institution. But gambits, unlike checkmate, are not themselves constitutive of chess. It is an important fact, though, that gambits would be impossible without the constitutive rules of chess, so even these non-constitutive symbolic utilities are closely tied to the status functions that define the game of chess. (Drawn upon in this section is Searle's familiar discussion of constitutive rules, status functions and collective intentionality.)

Participation in a profession may alter profoundly the character of rational decision-making, and not simply because one will have reason as a professional to do the things that are distinctive of one's profession, but also because the symbolic utility of doing those things well can tip the scales of rationally self-interested reflection towards conduct that would otherwise be imprudent or altruistic, or both. Notably there are professional duties and obligations which a professional rationally accedes to, even as a matter of self-interest, because of an expanded conception of the self and self-interest that attends the process of professionalization. Expansion of self-interest occurs when professionalization teaches one to take an interest in the interests of one's professional care.

Expansion of the self is closely related to this phenomenon, but it is a matter of decision and belief rather than interest and desire. Professionalism teaches one to define oneself partially in terms of exhibiting the professional virtues. One assigns considerable weight to this dimension of one's identity on Nozick's closest-continuer theory of personal identity, for instance.[13] In extreme and poignant cases, professionals' self-definitions lead them to judge that sacrifice of life for duty is in their best interest, because continued life otherwise would not be *their* life, which would have come to an end despite the continued existence of the living body and supervening psychological states.

This thought amounts to a twist on Nozick's solution to George Ainslie's problem about theorizing the irrationality of engaging in impulsive behavior that we know is against our long-term interests.[14] He is concerned with cases in which there is an earlier and lesser reward after an initial period A, followed by a later and greater reward in period C. Taking the B-reward will preclude taking the C-reward. (Let the B-reward be smoking and the C-reward be long life.) Although it is rational to stay aimed at the C-reward during the A-period, where the expected utility of the C-pursuing behavior is higher, this behavior has lower utility during the B-period. Nozick thoroughly

[13] Robert Nozick, *Philosophical Explanations* (Cambridge, MA, 1981), ch. 1, 'The Identity of the Self'.

[14] See the discussion of Ainslie in Nozick, *Rationality*, ch. 1, 'Overcoming Temptation'.

https://doi.org/10.1017/S0953820808003166 Published online by Cambridge University Press

explores a plausible solution: the B-rewards may represent always giving in to temptation, or symbolize an unattractive character trait that corresponds to such impulsiveness. The negative subjective utility of taking the B-rewards diminishes the overall utility of taking them, such that one rationally pushes through the B-period to the C-period and its rewards. Symbolic utility helps one preserve the full utility of C-period moments.

A different Ainslie-like scenario would invoke symbolic utility to explain the rationality of drastically *discounting* C-period moments. The middle period would now represent a period of professional risk-taking or refusal-to-cheat. What makes it rational for the professional soldier, doctor or other professional to risk and possibly lose his life in the pursuit of professional goals during the middle period, when the higher utilities of the later period are available by shunning risk or cheating? The higher utility of professional conduct during the B-period of pursuing one's career, say, can be accounted for by the weight attached to the subjective utility of doing one's professional duty. But what should be said about those cases in which B-period subjective utility is swamped by C-period utility? One answer involves an unexpected application of Rawls's maxim that utilitarianism fails to take seriously the distinction between persons: if the professional has arrived at a changed self-conception during the B-period, such that his life would be *over* upon failure to do his professional duty, the C-person's utility would not be *his*. The human being would survive as a vehicle for the pleasures of a long life, but the person whose identity is bound up with duty would not. Maximization of decision value leads to distinguishing two people, because of the role of symbolic utility in self-definition, whereas maximization of expected utility counts only one.

APPENDIX: THE REDUCTIVE MODEL

Jeffrey has shown that a description of the consequences of a certain act under a certain condition need be nothing more than a joint description of the act and the conditions. Resnick has argued somewhat less convincingly that the acts themselves in EU's act–outcome pairs might be construed as outcomes. However, the acts themselves in both cases are, in Ramsey's terms, 'ethically neutral'; that is, they don't express symbolic utility. So neither Jeffrey's nor Resnick's suggestion lends support to the reductive model of symbolic utility.

Jeffrey makes his point with an example of 'the right wine'. A dinner guest has forgotten whether chicken or beef is to be served at a dinner party, and consequently he does not know whether to bring red or

https://doi.org/10.1017/S0953820808003166 Published online by Cambridge University Press

white wine. Jeffrey constructs the following consequence matrix for his situation.[15]

	Chicken	**Beef**
White	White wine with chicken	White wine with beef
Red	Red wine with chicken	Red wine with beef

Jeffrey supposes that the dinner guest goes from this consequence matrix to the following desirability matrix.

	Chicken	**Beef**
White	1	−1
Red	0	1

Assuming that the guest regards the two possible conditions as equally likely regardless of whether he brings white wine or red, then the following probability matrix shows the probabilities.

	Chicken	**Beef**
White	.5	.5
Red	.5	.5

Given the numerical probabilities and desirabilities, the desirability of each act can be estimated by multiplying corresponding entries in the probability and desirability matrices and then adding across each row. Dropping the row and column headings, the matrices are:

.5	.5
.5	.5

1	−1
0	1

Multiplying corresponding entries yields a new matrix.

(.5)(1)	(.5)(−1)
(.5)(0)	(.5)(1)

[15] Richard Jeffrey, *The Logic of Decision* (Chicago, 1983).

https://doi.org/10.1017/S0953820808003166 Published online by Cambridge University Press

which resolves to

.5	−.5
0	.5

The desirability of the first act (white) is given by adding across each row.

$$(.5) + (-.5) = 0$$

And similarly for the desirability of the second act (red),

$$0 + .5 = .5$$

So bringing red wine has the higher estimated desirability, and according to Bayesian principles it is the better choice. However, the acts remain 'ethically neutral' in these manipulations. Preference arises because white wine with chicken is the right wine, white wine with beef is the wrong wine, and so forth, as revealed by the desirability matrix. Jeffrey shows that a description of the consequences of a certain act under a certain condition need be nothing more than a joint description of the act and the condition. But such techniques don't promise to reveal the symbolic utility of the action. On the contrary, they assume its neutrality in this regard.

Resnick's proposal redescribes an act so as to include its outcomes. Here is the relevant passage, in which he asks the reader to consider Joan's problem:

> She is pregnant and cannot take care of a baby. She can abort the fetus and thereby avoid having to take care of the baby, or she can have the baby and give it up for adoption. Either course prevents the outcome *Joan takes care of the baby*, but Joan (and we) sense a real difference between the means used to achieve that outcome. There is a simple method for formulating Joan's choice so that it becomes the true dilemma that she sees it to be. We simply include act descriptions in the outcome descriptions. We no longer have a single outcome but two: *Joan has an abortion and does not take care of a baby, and Joan gives her baby up for adoption and does not take care of it*.[16]

Unlike Jeffrey, Resnick begins with an act that is ethically non-neutral because it includes a sub-act that has negative symbolic utility, which stands in a causal relation to a condition of Joan's not taking care of a baby. This relationship is wrapped up in a contrived complex act that inherits the sub-act's negative utility. However, Resnick's proposal

[16] Michael Resnick, *Choices: An Introduction to Decision Theory* (Minneapolis, 1987), p. 16.

https://doi.org/10.1017/S0953820808003166 Published online by Cambridge University Press

does not support the reductive account of symbolic utility (the account that would break it down into expected utility) because the negative symbolic utility of having an abortion – its dilemmatic aspect – is independent of the outcome that Joan does not take care of a baby. For instance, the abortion would be dilemmatic even if she planned to adopt a baby, or even if she intended to take care of *this* baby after having it killed, just in case someone was able to bring it back to life. What's wrong about the act remains with the act. Resnick's analysis does not show how the wrongness is transferred to the act's outcomes.

wcooper@ualberta.ca

https://doi.org/10.1017/S0953820808003166 Published online by Cambridge University Press