

MODELLING THE MORAL DIMENSION OF DECISIONS

BY MARK COLYVAN[†], DAMIAN COX[‡], & KATIE STEELE[†]

In this paper we explore the connections between ethics and decision theory. In particular, we consider the question of whether decision theory carries with it a bias towards consequentialist ethical theories. We argue that although there is a sense in which other ethical theories can be accommodated by standard decision theory, these ethical theories are not naturally suited to the project of maximising expected utility. In particular, we argue that the motivation for the other ethical theories and the psychological processes of the agents who subscribe to those ethical theories are lost or poorly represented in the resulting models.

1. INTRODUCTION

Decision theory has two components: probabilities and utilities. From the formal point of view, these two components play symmetrical roles in decision theory. For each act–state pair (or outcome) we assign a probability and a utility, then we multiply these together and sum the products across each act. The resulting sum we call the *expected utility*. Standard decision theory then tells us to choose the act with the greatest expected utility (if there is such an act). This is all very familiar. What we draw attention to is that, in terms of the formal decision calculus,

[†] Department of Philosophy, University of Sydney, Sydney, NSW, 2006, Australia.

[‡] Faculty of Humanities and Social Sciences, Bond University, Gold Coast, Qld, 4229, Australia.

probabilities and utilities are treated similarly—they are just two real numbers to be multiplied and then added.

The similarities between probabilities and utilities run even deeper. Both are constrained by axiomatic theories: probabilities by the Kolmogorov (1956) axioms and utilities by von Neumann-Morgenstern (1944) utility theory. These axiomatic theories place minimal structural constraints on probabilities and utilities.¹ The axioms ensure that there are no violations of consistency. For example, the Kolmogorov axioms rule out probabilities that don't sum to 1, so if an agent assigns probability p to some proposition Q , then the agent must assign $1-p$ to $\neg Q$. And the von Neumann-Morgenstern axioms insist, for instance, on transitivity of preferences: if an agent prefers A to B and B to C , then the agent ought to prefer A to C . Call an agent whose probability assignments conform to the Kolmogorov axioms a *Kolmogorov-consistent agent*, and call an agent whose utility assignments conform to the von Neumann-Morgenstern utility theory a *von-Neumann-Morgenstern-consistent agent*. It is clear that mere Kolmogorov-consistency and von-Neumann-Morgenstern-consistency are, in general, not enough to ensure that an agent is beyond reproach. Compliance with these conditions does not imply responsible decision-making.

Take probabilities first. Consider an agent who assigns probability 1 to the flat earth theory and probability 0 to every other theory about the shape of the earth. Such an agent is *Kolmogorov*

¹ One might argue about whether the respective axioms are in fact minimal constraints. We won't pursue such issues here. Our interest lies in what other constraints we need in addition to the axioms. Nor will we say much about alternative axiomatisations of probability theory and utility theory. Most of what we say will go through for other axiomatisations, but we stick with these standard axiomatisations to give focus to our discussion.

consistent, but she is a poor or irresponsible *epistemic agent*, for she assigns zero probability to a theory for which there is a great deal of evidence (the roughly-spherical earth theory). Moreover, because she assigns probability 1 to the flat earth theory, no amount of evidence will change this assignment, since updating on new evidence via Bayes's theorem will never lower the probability of maximal-probability propositions. What are the extra constraints we need to ensure that merely Kolmogorov-consistent agents are responsible epistemic agents? Well, this is going to be a complicated epistemological story, and the details will depend on the particular epistemology to which you subscribe.

Now to utilities. Consider an agent who prefers genocide to murder and prefers murder to a walk in the hills. So long as this agent satisfies transitivity (i.e., prefers genocide to a walk in the hills) and other such structural constraints, the agent is a von-Neumann-Morgenstern-consistent agent. But clearly such an agent is a poor *moral agent*. What are the extra constraints we need to ensure that merely von-Neumann-Morgenstern-consistent agents are responsible moral agents? This is going to be a complicated story about ethics, and presumably the details will depend on the particular ethical theory to which you subscribe.

So while decision theory is able to gloss over the details of how probabilities and utilities are assigned (and hence gloss over the thorny issues in epistemology and ethics), a full account of decisions requires both an epistemological and an ethical theory. Moreover, we need the epistemological and ethical theories to be spelled out in ways which enable them to be accommodated in the standard decision-theory framework. Of course a great deal has been written on the epistemic side of this story. In particular, there is a large literature on the question of whether decision theory requires Bayesianism or whether other interpretations of the

probabilities are possible.² But, given the symmetry between the probability side of things and the utility side, it is somewhat surprising that very little has been written about the ethical side of decision theory. This, then, brings us to the central question we will address in this paper: how much freedom do we have in choosing an ethical theory to accompany standard decision theory?

One way of looking at this question is as a search for a philosophical interpretation of utilities. Just as Bayesianism is the obvious candidate for the philosophical interpretation of the probabilities that arise in decision theory (if not *all* probabilities), consequentialism presents itself as the obvious candidate for the philosophical interpretation of the utilities employed by decision theory. But is consequentialism the only game in town, where decision theory is concerned? We investigate whether alternative ethical theories can be at least accommodated in the decision-theory framework. In other words, is the utility function capable of representing the preferences of agents who subscribe to ethical theories other than consequentialism?

Before we move on we should say a little about the significance of what we are trying to do and why we approach it the way we do. In relation to the significance of our project, it might be thought that it is trivial or at least well known that various ethical theories can be modelled in the decision theory framework.³ Even if this is right, it is still interesting to see at least some of the details, rather than rest content with the simple knowledge that it can be done. Moreover, in

² See, for example, Gillies (2000) and Hájek (2003a) on the interpretation issue and Jeffrey (1983) and Kaplan (1996) for the details of decision theory and its role in philosophy (respectively).

³ Indeed, there are some interesting results along these lines. See Oddie and Milne (1991). Others who have considered the issue of the relationship between ethics and decision theory include Broome (1991), Colyvan *et al.* (2001), Dreier (2004), Jackson (2001), Jackson and Smith (to appear), Louise (2004), and Sen (2004).

providing the details, we shed some light on the ethical theories, the decision theory framework, and the relationship between the two. The question arises as to whether it is appropriate or useful to merely accommodate other ethical theories within decision theory or do we do too much violence to the said theories in trying to submit them to this framework? We will return to these issues in the final section.

We set ourselves the task of trying to accommodate the ethical theories in question into *standard* decision theory. Another option might be to play around with the decision theory model to introduce, say, two-dimensional utility functions (e.g., Hájek 2003). While there are no doubt some interesting avenues to be explored here, for the most part we try to restrict ourselves to standard decision theory.⁴ There are several reasons for this but the most significant is simply that this framework is widely used and we would hope that all ethical theories might be able to utilise it. To develop different decision theories for different ethical theories tends to marginalise ethics. Finally, we note that we will occasionally run with very simple versions of the ethical theories in question. For the most part we do this because we want our account to be as general as possible. We do not want to take too many stands on substantial issues (although occasionally this will be unavoidable). So in a nutshell, we want to show the details of how to accommodate fairly generic versions of the major ethical theories into standard decision theory.

⁴ Occasionally we are forced to deviate from standard decision theory, and when we do so we will acknowledge this. Still, our goal will be to stay true to the spirit of standard decision theory, even when we depart from it in detail. Our departures strike us as fairly natural extensions of standard decision theory, although we are aware that not everyone will agree about the naturalness claim.

2. THREE ETHICAL THEORIES

We examine three of the most important types of theory in contemporary philosophical ethics: utilitarianism⁵, deontology and virtue theory. Ethical theories are primarily theories of right action. Although virtue ethics is often taken to be an exception, we treat it here as furnishing a theory of right action. Our challenge is to translate theories of right action into a decision-theoretic framework by representing them as offering diverse accounts of utility. First, however, we outline our general approach to the theories and what we regard as distinctive of each of them.

Utilitarianism provides us with the most natural method of generating utility assignments because it expressly incorporates a theory of utility, one that identifies utility with welfare. Welfare is conceptualized by utilitarians in markedly different ways. For example, classical utilitarianism—the theory associated with Jeremy Bentham and John Stuart Mill—identifies welfare with happiness, where the latter is interpreted as a preponderance of pleasure over pain. Preference-satisfaction utilitarianism, by contrast, identifies welfare with the satisfaction of self-regarding preferences (i.e., with a person's preferences about how their life is to go).⁶ More objective conceptualizations of welfare have also been proposed in which a mixed basket of objective life conditions are said to contribute in some reasonably well defined way to a person's overall welfare. The important point for our purposes is that every utilitarian theory proposes an account

⁵ We recognize that utilitarianism is just one class of consequentialist moral theory. We focus on utilitarianism because it is arguably the dominant consequentialist theory, and serves as a useful point of comparison with the other ethical theories.

⁶ Preference utilitarianism introduces a cardinal measure of social utility because it involves (at least) comparing total numbers of satisfied preferences.

of utility as welfare and introduces a cardinal measure of total utility in a situation. Now, utility assignments of this kind are widely dismissed as too controversial and probably undiscoverable, and for these reasons do not figure much in economic theory and public policy. However, for philosophical purposes, one might nonetheless insist that the notion of a cardinal measure of welfare is coherent and plays a role in determining right action, even though we are usually in a position only to approximately and fallibly identify levels of welfare.

An important feature of utilitarianism is the interpretation of the nature of value it presupposes. For utilitarians, values are attributed to possible states of affairs and furnish reasons to bring them about. Utility is the sole measure of morally relevant value for utilitarians and they are naturally led towards a maximizing principle with respect to it. To act in a way that promotes an outcome other than maximum utility would be to value something other than utility more than one values utility. But since utilitarians recognize no morally relevant value other than utility they generally regard it as a moral error to pursue less than the maximum available utility.⁷

Deontologists differ from utilitarians in that they do not offer a theory of utility at all. Rather, deontology introduces a set of moral constraints upon decision-making. These include prohibitions and obligations that often have the effect of undercutting welfare maximization.⁸ For example, observing a prohibition against targeting civilians in war may prolong a bloody conflict

⁷ For this discussion, we set aside consideration of satisficing versions of utilitarianism, i.e. versions in which moral agents aim for sufficiently good rather than maximally good levels of utility.

⁸ Alongside obligations and prohibitions, deontologists sometimes also posit permissions or prerogatives. We set aside these aspects of deontology here.

so that, on any reasonable assessment of welfare, general welfare is greatly diminished as a result.

Prohibitions and obligations need not be absolute; they may be conditional. For example, a deontologist may allow lying under some circumstances—say when a person is negotiating with you in bad faith—and not others. A deontologist might also take the application of duties to be context-sensitive: the duties relevant in one situation may not be relevant in another. For example, a parent may have duties to their child that a stranger does not have. It is also possible for a deontologist to hold certain duties to be defeasible. This is the view that prohibitions and obligations may be defeated by the threat of certain amounts of disutility. For example, a deontologist might think it permissible to lie or steal when the threat to general well-being, or even to their own survival, is sufficiently high and no better alternatives are available. Prohibitions and obligations may also be ranked, so that, for example, prohibitions against killing outrank prohibitions against stealing, both of which outrank obligations to come to the aid of neighbours in distress.

Deontologists are motivated by a conception of the moral relevance of value that is very different from that of the utilitarian.⁹ Where the utilitarian conceives of judgments of value as supplying reasons to promote particular states of affairs, the deontologist is likely to think of morally significant value judgments as directed towards persons, morally requiring that persons be respected. For deontologists, morally valuing someone entails respecting them, not seeking to

⁹ The following characterization is a simplification. Some deontological approaches rest on direct intuitions about duties and about the priority of the right over the good rather than on accounts of respect. See Ross (1967).

promote their welfare. To respect a person is not to wish to see more of a certain valuable state of affairs, but, at least in part, to accept that we are legitimately restricted in what we may do to or for the person. Deontologists characteristically direct respect towards persons, but other subjects of respect are also possible in deontology. For example, deontological environmental ethicists may value all living things.¹⁰

The third ethical theory under view—virtue ethics—neither offers a theory of utility nor a set of explicit moral constraints upon action. Virtue ethics is first and foremost a re-focusing of moral theory away from the concern to provide an account of correct moral decision-making. The central question for the virtue ethicist is not ‘what should I do?’ but ‘what kind of person should I be?’ and this latter question is not sufficiently well answered by observing that I should be the kind of person who acts rightly. Nonetheless, a number of contemporary virtue ethicists attempt to show how virtue-theoretic considerations contribute directly to the task of moral decision-making. Virtue-theoretic accounts of right action derive an account of right action in one way or another from an independent account of virtue. The idea is to first determine which character traits or motivations are morally admirable (either intrinsically or because they are essential features of a flourishing life or of a morally good life) and use this account to describe how a morally good person would act in various challenging situations. On one natural development of this idea, right action is identified with virtuous agency.¹¹ For such virtue ethicists, to act rightly is to act virtuously and to act virtuously is to manifest a complex inner state: a state involving (at least) morally admirable motivations. An alternative virtue-theoretic approach to right action

¹⁰ See, for example, Paul Taylor (1986).

¹¹ Slote (2001) develops virtue ethics along these lines.

identifies right action with actions that would, characteristically, be performed by a virtuous agent.¹² Such an indirect, or hypothetical, virtue-theoretic account of right action has the advantage of preserving intuitions about the distinction between doing the right thing (acting rightly) and doing the right thing for the right reason and with the right motivation (acting virtuously). By contrast, direct virtue-theoretic accounts—those identifying right action with virtuous action—enjoin moral agents to pursue virtuous action rather than merely conform to the standards of virtuous action.

Virtues interact with each other in complex ways. It is unlikely that any simple ranking of virtues will capture the variable significance of virtues as they apply in complicated and varied circumstances. When does benevolence—a virtue directed at the well-being of strangers—take precedence over the virtue of caring for loved ones? When is the virtue of integrity outranked by the virtues of practicality and willingness to compromise? In answering questions like these, the virtue ethicist may appeal to their version of what it is for the moral agent concerned to live a good life and the role virtues play in this life or they may appeal directly to intuitions about the comparative significance of virtues. There is very little precise work on the problem and this reflects real imprecision in our picture of a good life and uncertainty about key moral intuitions. The important point for our purposes is that we take the virtue ethicist to propose some account—albeit a gappy and imprecise account—of how virtues interact which furnishes them with an account of what it is for an action to be the most virtuous possible action in a situation. For example, on a particular occasion a person might face the choice of acting with general

¹² Rosilind Hursthouse (1991 and 1999) develops such a view.

benevolence or acting with a special concern for loved ones and the virtue ethicists owes us a way of ranking the virtues in this situation (even if they rank the options equally).

In tying right action to virtuous action, either directly or indirectly, the virtue ethicist identifies valued states of affairs that are to be promoted. For the direct virtue ethicist, an agent's own virtuous agency is to be promoted; for the indirect virtue ethicist, an agent's conforming to virtuous action is to be promoted. Thus virtue ethics identifies states of affairs to be promoted, but unlike utilitarian promotion of welfare, virtue-theoretic values are agent-relative (i.e. indexed to agents) and time-relative (i.e. indexed to the immediate choice situation of the agent). Virtue ethicists do not typically hold a moral agent responsible for the impartial promotion of virtuous action; they have it that individual agents have a special responsibility for their own character and they act rightly only when they act virtuously (or act as the virtuous would act).

The three theories outlined here represent fundamentally different approaches to right action. Utilitarians define right action as the achievement of a goal—maximized utility. Virtue ethicists also see right action in terms of the achievement of a goal, this time an immediate and self-oriented goal. In every choice situation, the virtue ethicist's goal is either to manifest virtue in the situation (in direct versions of virtue ethics) or to outwardly conform to virtuous behaviour in the situation (in indirect versions of the theory). Deontologists, by contrast, see right action as a matter of complying with rules of conduct; not promoting the compliance of rules, but actually complying with rules in every choice situation.

3. INTRODUCING ETHICS TO DECISION THEORY

Now we turn to the task of trying to incorporate these three ethical theories into the standard decision theory framework. Our hands are tied somewhat here, because, as we've already pointed out, there is really only one place in decision theory where ethics can enter the picture, and that is as additional constraints on admissible utility functions. And here it might seem that there is only one way to proceed, for the deck seems to be stacked in favour of utilitarianism. Indeed, even the language (“*utility* functions” and “expected *utility*”) suggests that utilitarian ethics is the only real contender here. Accommodating the other two ethical theories is by no means straightforward, but deontology and virtue ethics should not be dismissed simply because decision theory employs a mathematical function called ‘a *utility* function’, and this is suggestive of utilitarianism. As we hope to show, we can indeed give the other two a fair run in the framework of decision theory, at least in terms of getting the “right” answers to decision problems. In an important sense we will be trying to *model* the ethical theories in question, rather than give formal presentations of the theories themselves. This distinction will become important in the final section. First we turn to the task of determining, for each of the three ethical theories, what type of constraints should be placed on the utility function, so that decision calculations yield morally correct actions.

Let's start with utilitarianism. As we've already mentioned, representing utilitarianism in the decision theory framework is fairly straightforward. Decision theory already considers the value of each act-state pair, and this value is measured on an interval (as opposed to a merely ordinal) scale.¹³ The best act is that which has *maximum* expected value or utility. All this sits very well with utilitarianism, but we shouldn't be fooled into thinking that decision theory's utility

¹³ That is, the utility functions employed in decision theory respect distance between the various values.

functions just are the utilitarian's value functions. For a start, nothing in the von Neumann Morgenstern utility theory rules out a utility function that assigns greatest value to the act-state pairs that are associated with the greatest harm. Before decision theory's utility functions can be considered utilitarian, we must add something like the following constraints. An outcome, O_{ij} , is the result of the agent performing action a_i while the world is in state s_j . (We will use this notation for acts, states, and outcomes throughout the rest of the paper.) There will, in general, be a range of possible outcomes of performing a_i . Two constraints on a utilitarian utility function are:

(U1) If O_{ij} involves greater total welfare than O_{kl} , then any admissible utility function u must be such that $u(O_{ij}) > u(O_{kl})$.

(U2) If O_{ij} involves the same total welfare as O_{kl} , then any admissible utility function u must be such that $u(O_{ij}) = u(O_{kl})$.

These constraints are enough to prevent perverse anti-utilitarian functions that assign greater value to the greatest harm to the greatest number. In other words it is sufficient for ruling out perverse unethical functions that value genocide over a walk in the hills. There are still substantial details that need to be dealt with, most of which involve the details of the version of utilitarianism to be represented, and in particular the account of welfare that is in play. We set these aside, though, for there does not seem to be any serious impediment to this kind of approach to the incorporation of utilitarianism into decision theory. Indeed, adding the above constraints (U1 and U2) is very natural, and something like this is likely to have been in the back of the minds of many decision theorists.

Next consider deontology, and how it might be incorporated into decision theory. The first point to observe is that whereas utilitarians evaluate outcomes wholly in terms of their preferred characterisation of welfare represented by an outcome, a deontologist evaluates morally salient outcomes in terms of the actions that produced them. We can say that in both cases, the subject of evaluation is an outcome of an action, but insist that for the utilitarian welfare determines utility assignment and for the deontologist, in certain morally salient situations, the nature of the action determines utility. Also, in a sense we will make clear shortly, we need to assume that deontology is agent and time relative, i.e. only the agent in question's actions at a particular time can have the status of being duties.

As mentioned in our characterisation of deontology above, a deontologist may think of obligations and prohibitions as conditional upon types of situation encountered. For example, a deontologist might consider lying prohibited, except for cases when they think another person is negotiating with them in bad faith. The clearest way to introduce this type of conditionality into our model is at the level of act description. A deontologist may say either that lying is prohibited except in cases of bad faith negotiations or they may refine their description of the relevant act. Under such a description, instead of lying *per se* being prohibited conditionally, lying of a certain kind is prohibited unconditionally. If we allow this way of refining act descriptions, we need not

worry about a deontologist's utility function having to accommodate duties whose claims on the agent are conditional upon these kinds of situational details.¹⁴

Here is a preliminary attempt to describe the constraint prohibitions and obligations make on utility functions, specified in terms of types of acts yielding an outcome:

(D1*) If O_{ij} is the result of an (absolutely) prohibited act, then any admissible utility function u must be such that $u(O_{ij}) = -\infty$.

(D2*) If O_{ij} is the result of an (absolutely) obligatory act, then any admissible utility function u must be such that $u(O_{ij}) = +\infty$.

Given that our project is to supplement the von Neumann-Morgenstern preference axioms with further constraints to model a deontologist's utility function, D1* and D2* clearly won't do because the von Neumann-Morgenstern axioms do not permit outcomes with infinite utility¹⁵. We put D1* and D2* on the table, however, because we are aiming for something in the vicinity of

¹⁴ Conditionality is a special case of context-sensitivity. Duties may not only be conditional upon the satisfaction of background conditions, they may vary in numerous ways from one context to another. For example, rankings of duties may vary from one context to another. We ignore this complication in our model.

¹⁵ The continuity or Archimedean Axiom rules out infinite utilities. Informally, the axiom states that for any three acts p , q , and r where p is preferred to q and q is preferred to r , there is some mixed act comprised of p and r that is indifferent to q . Refer to Resnik (1987, p. 91) for a formal statement of the axiom. It is clear that if act p had infinite utility or act r had negatively infinite utility, then continuity would not hold because no mixture of p and r would correspond to an act with intermediate finite utility.

these constraints.¹⁶ What goes in favour of D1* and D2* is that these constraints make the utility of certain morally salient outcomes dependent on the action that produces them. Moreover, D1* and D2* set obligations (prohibitions) apart from other acts by ensuring that they have absolute maximum (minimum) expected utility.

As well as the problem of infinite utilities, a significant problem with our way of modelling deontology so far is that all prohibited (or obligatory) acts are on a par. Murder, if it is prohibited, is no better or worse than genocide, if this too is prohibited. In effect, we have only modelled absolutely binding prohibitions and obligations and we have not introduced means of comparing the claims of one duty against another. One way of responding to this problem is to assign very large utilities and disutilities in place of infinite utilities. The important thing about duties, from a deontologist's point of view, is that one is rationally bound to them over and above other possibilities of action. An alternative set of deontological constraints, therefore, would be:

(D1) If O_{ij} is the result of a prohibited act, then any admissible utility function u must be such that $u(O_{ij})$ has a disutility greater than any assigned to the outcome of a permitted act.

¹⁶ We note that it is possible to use the extended reals in place of the reals as the range of the utility function. (The extended reals include positive and negative infinity.) This is sometimes assumed to be a minor modification of utility theory but it is not. Unless further adjustments are made, the resulting theory violates the Archimedean axiom (as noted above).

(D2) If O_{ij} is the result of an obligatory act, then any admissible utility function u must be such that $u(O_{ij})$ has a utility greater than any assigned to the outcome of any non-obligatory act.

This allows us to introduce a ranking among obligations and prohibitions. Say there is a ranking of prohibitions P_1, P_2, \dots, P_n , such that P_k outranks P_{k+1} and a ranking of obligations, or positive duties, D_1, D_2, \dots, D_m , such that D_l outranks D_{l+1} , then:

(D3) If O_{ij} is result of a prohibited act of type P_r and O_{kl} is the result of a prohibited act of type P_s , such that P_r outranks P_s , then any admissible utility function u must be such that $u(O_{ij}) < u(O_{kl})$.

(D4) If O_{ij} is the result of an obligatory act of type D_r and O_{kl} is the result of an obligatory act of type D_s , such that D_r outranks D_s , then any admissible utility function u must be such that $u(O_{ij}) > u(O_{kl})$.

One well-known problem with allowing infinite utilities is that they can swamp the probabilities, so that decision theory is rendered useless. For instance, if killing is assigned infinite disutility, then any of my actions that might lead to me killing (no matter how unlikely) might seem to be impermissible. This would clearly be unsatisfactory. The move to using large finite utilities for obligations and small finite utilities for prohibitions might seem to address this problem, but an analogue remains. After all, if the utilities are small enough for prohibitions then these utilities will still swamp *most* probabilities. For example, if we attach a very small utility to killing, then any action of mine (e.g. driving to work)

that has some small but not miniscule probability of leading to me killing in the future will be thus selected against. In effect, we run the risk of ending up with a decision theory that is crippled by swamping effects of the large and small utilities we introduce to capture the deontologist's attitude towards duties and prohibitions.

The swamping problem associated with infinite (or very large) utilities highlights the importance of stipulating that duties be agent and time relative, if deontology is to be accommodated in the decision theory framework.¹⁷ Imposing strict time relativity on the operation of duties might seem intuitively questionable, but this is simply the result of a deontological insistence that duties—moral rules—are to be complied with. Deontological agents do not seek to promote maximal compliance with rules of conduct and they are not prepared to violate rules in order to promote such compliance. Typically, a deontologist would not be prepared to kill just in order to prevent future killings (by themselves or by others). Prohibitions and obligations apply rather to deontological moral agents in their immediate choice situation. Note also that deontologists can capture many intuitions about our responsibility for our future actions by introducing future-oriented duties, such as the duty not to place oneself in moral danger or a prohibition against reckless behaviour.¹⁸

¹⁷ We are assuming here that a deontologist would regard an agent's own future actions as probabilistic, rather than known with certainty, because in the latter case swamping would not be an issue even if duties were only agent relative and not time relative.

¹⁸ It may also be the case that certain actions may be ruled against on consequentialist grounds. For example, my buying a gun would not be prohibited just because it raises the chance of my future self engaging in the prohibited act of killing. Rather, my buying a gun might simply be ruled against because it raises the probability of low utility

The revised deontological constraints on the utility function (D1–D4) are thus a significant improvement on D1* and D2*. But while not *necessarily* von Neumann-Morgenstern inconsistent, D1–D4 are still suggestive of a violation of the continuity axiom. The issue revolves around the status of mixed acts.¹⁹ How does the deontologist rank a mixed act that could yield (with, say, equal probability) either an obligation or a prohibition? Continuity obliges us to accept that such a mixed act has utility close to that of an everyday act like snacking on an apple. This might not ring true, however, to the deontologist. To give another example, is it the case that there exists some weighted combination of eating an apple and torturing innocent children that is indifferent to each act that is ranked between these two? It is not obvious whether the deontologist will answer “yes” or “no” to this question. If it is thought that no mixed act giving any probability to torturing small children, however small the probability, should be ranked alongside, say, eating an apple and paying \$1, then we effectively require a violation of continuity. On the other hand, what *does* the deontologist want to say about such mixed acts? A valuable aspect of the modelling process is that it focuses attention on such questions and thus sheds light on particular details of deontological theory.

In any case, we have not yet provided the means to represent all the demands of deontology. A difficult issue yet unaccounted for is that we want to allow deontologists to recommend *defeasible* obligations and prohibitions, that is, obligations and prohibitions that may be

outcomes such as deaths by shooting. We also note that relativity we highlight does not dispense with *all* swamping problems, just some of the more obvious ones.

overturned by consequentialist considerations when the stakes are sufficiently high. This is perhaps best considered another way in which the description of duties is context-sensitive. For instance, lying under circumstances in which only a modest amount of consequentialist good would otherwise be forsworn is prohibited, whereas lying under circumstances in which excessive consequentialist good would otherwise be forsworn is just an ordinary permissible act.²⁰ The trouble is, if we want to be more precise about the circumstances in which lying is prohibited, then we will need to do the consequentialist utility calculations. But this introduces a strange circularity—it makes the description of an act dependent on the utility of its outcomes, and yet if the act is subsequently described as a duty, then its outcomes in fact have utility (large positive or large negative) associated with that duty-type.

As the reader might suspect, at this point our model gets somewhat complicated. Regardless, we will suggest a way to deal with defeasibility. It is not a matter of adding further utility function constraints to the von Neumann-Morgenstern axioms. To capture defeasibility, we need to assume that the deontologist's preferences are already represented by a utility function (unique up to positive linear transformation), and that they can now make use of the tools offered by such a representation in the course of their moral decision-making.

As stated, the defeasibility of duties means that an agent may have difficulty classifying the available acts in a particular choice situation. Whether or not an act that looks like a duty actually *is* a duty will depend on whether its consequences breach the nominated defeasibility threshold.

¹⁹ We thank David Gray for drawing this issue of mixed acts to our attention.

Now it seems most plausible to say that moral duties are variably defeasible. The circumstances in which a deontologist might be prepared to license the telling of a lie are more widespread than the circumstances in which they would licence the torture of innocents, for example. We suggest that the deontologist select a utility function (from amongst the possible set of utility functions that are each positive linear transformations of each other), and nominate degrees of utility/disutility appropriate to the exercise of every obligation and prohibition—limits beyond which they are prepared to give up the claim of that duty. This could be recorded as a function f from prohibitions and obligations to defeasibility limits. It is a function from act-types to degrees of utility (high positive utility in the case of prohibitions and high disutility in the case of obligations). The idea is that acts are classified as obligations or prohibitions only if their outcomes do not breach the defeasibility limit specific to that duty.

So what is it exactly for a duty to breach its defeasibility limit? It is arguably the potential expected utility of the act, rather than the utility of any of its outcomes in isolation, that should be assessed relative to the defeasibility threshold. And (again arguably) the most plausible model of this makes defeasibility dependent on the nature of the alternative available acts. In this way, defeasibility is not about breaching some absolute utility value (supplied by the function f), but rather is an issue of the harm done or the goods foresworn by acting in the ordinarily obligatory manner. In certain situations, for example, lying may be permissible, not because the consequences of lying exceed some absolute expectation, but because all other available options are just so much worse, from a consequentialist point of view, than lying.

²⁰ Compare with the sort of contextual distinction that was made earlier between ordinary lying and lying under conditions of bad faith.

Here we formalise defeasibility, not in terms of further constraints on the deontologist's utility function, but rather in terms of how an agent should classify acts. Consider defeasible obligations first. Let a^* be an act that is possibly the obligation a_o . The relevant defeasibility limit (given a particular utility representation) is thus given by $f(a_o)$. Let act a_m be the alternative available act in the option set with maximal expected utility.

If, according to initial calculations $EU(a^*) < EU(a_m) - f(a_o)$

then the defeasibility limit is surpassed and there is no obligation a_o , just the regular permissible act a^* within the set of available options. Consider a case where an agent has the opportunity to bring about great good, say by tipping the vote in an important committee meeting, but where he/she must also consider whether to spend the time helping his/her children with their homework, something that in ordinary circumstances is a duty. Well, if the expected utility of going to the committee meeting exceeds the expected utility of upholding the children's homework routine by an amount greater than the defeasibility value for the helping-with-homework duty, then the agent is not duty-bound here. If, on the other hand, the expected utility of going to the meeting is not sufficiently greater than that of upholding the homework routine, the agent in fact here faces a duty. In other words, if the inequality above does not hold, what appeared to be a^* is in fact the obligation a_o and its outcomes will be assigned the appropriate very large utilities.

Defeasible prohibitions work similarly. Let $a\#$ be an act that is possibly the prohibition a_p . The relevant defeasibility limit (given a particular utility function) is thus given by $f(a_p)$. Let act a_m be the alternative available act in the option set with maximal expected utility.

If, according to initial calculations $EU(a\#) > EU(a_m) + f(a_p)$

then the defeasibility limit is surpassed and there is no prohibition a_p , just the regular permissible act $a\#$ within the set of available options. Otherwise (if the inequality above is not satisfied), what looked like act $a\#$ is in fact the prohibition a_p , and its outcomes will have the appropriate very large negative utilities.

Of course, there is scope for disagreement amongst deontologists as to how the defeasibility condition actually works, but we have given here a plausible model that may be adjusted as desired. Of course, the model offered is a significant deviation from the standard decision theoretic framework because it makes the classification of acts a two-step process that can be made precise only by referencing a given utility representation. What is interesting about this part of the model, however, is that it makes decision theoretic tools (however non-standard) available to the deontologist. Note also that in modeling the deontologist's utility function, we must help ourselves to a sufficiently precise characterisation of acts in terms of the duties—prohibitions and obligations—relevant to these acts under the given circumstances, as well as a comprehensive ranking of these prohibitions and obligations.

Next we turn to virtue ethics. This is more difficult still since, as we pointed out in section 2, virtue ethics is concerned as much with the motivations of the agents as with their actions or with

the outcomes of those actions. But we can make some headway here. A concrete example might be helpful at this stage. Consider the choice of whether to donate to charity or not. A consequentialist is only interested in the outcomes of the acts in question—it doesn't matter whether a donation is motivated by generosity, by the desire to be seen to be generous, or by guilt. All that matters, from the perspective of the consequentialist, are the consequences of the act in question.²¹ But the motivations make all the difference in the world for the virtue ethicist. Presumably, the act of charity stemming from generosity is what a virtuous person would aim at doing (at least for direct virtue theorists, who identify right action with virtuous action). This, then, suggests a way of incorporating virtue ethics into decision theory: we discriminate actions a little more finely than normally so that we distinguish actions with different motivations, and assign different utilities according to the virtuousness of the act's motive. If an act has a virtuous motivation, we say that it expresses the virtue in question. Thus, a generous act of charity expresses the virtue of benevolence. A cynical act of charity does not.

A satisfactorily complete virtue theory should provide us with the means of discriminating actions in terms of their expression of virtue. Indirect virtue theories, i.e. those that characterize right action in terms of actions that a hypothetical virtuous agent *would* undertake, will describe the hypothetical expression of virtue. Actions would then be ranked in terms of the extent of their match to the motivations of ideal virtuous agents. Here we consider only direct versions of virtue theory. Though more complex, indirect virtue theories can be accommodated within the general framework we describe. In any situation an agent will confront a finite number of available

²¹ This is a little simplistic. Charity motivated by guilt might have different consequences from charity motivated by generosity, but let's set such complications aside for now.

actions, $a_1, a_2, \dots, a_i, \dots, a_n$. We use our virtue theory to rank these actions in terms of their expression of virtue. Minimal constraints that virtue theories impose upon the relevant utilities are given by V1 and V2 (where we read ' $O_{ij} \equiv O_{hk}$ ' as outcomes O_{ij} and O_{hk} are equivalent, in the sense that they are indistinguishable outcomes in all but their virtue-theoretic motivations).

(V1) If a_i is more virtuous than a_h and $O_{ij} \equiv O_{hk}$, then $u(O_{ij}) > u(O_{hk})$.

(V2) If a_i and a_h are equally virtuous and $O_{ij} \equiv O_{hk}$, then $u(O_{ij}) = u(O_{hk})$.

Such virtue-theoretic constraints on the utility function come down to this: the utility of the outcome of a virtuously motivated act will always exceed the utility of that very outcome produced by less virtuous means. A charitable gift may be a valuable outcome on this way of viewing things, but well-motivated charitable gift-giving invariably possesses higher utility. Equivalent outcomes of equally virtuous actions have identical utility. Say that fairness and generosity are equally virtuous motivations. The outcome of a charitable donation given out of a sense of fairness would then attract the same utility as an identical outcome motivated by generosity. (It is more likely, however, that virtues are not all equal like this. See below.)

V1 and V2 together represent a minimal virtue-theoretic constraint. The resulting utility function will not reflect an intuitively compelling virtue perspective until the level of difference that virtuous motivation brings to assignments of utility becomes very marked. Not all virtues are equally significant and so it is plausible to say that every virtue will make a distinctive

contribution to the utility of the outcomes of actions that express it.²² Moreover, it is highly plausible to say that judgements of virtuous action combine a concern for virtuous motivation with a concern for good outcomes. In all plausible versions of virtue theory, a virtuous action is also a wise or prudent one, one that accommodates thinking about the probabilistic outcomes of actions. Of course, it is the primary business of decision theory to model this kind of rationality and our model of virtue-theoretic decision making ought to make use of this resource. In view of this, we might try to model virtue theory in terms of a virtue-specific multiplicative factor that modifies a prior set of utilities. Thus we introduce a function, V , from motivations to multiplicative factors. Consider a virtuous (or vicious) action a_i , expressing a particular motivation m_i . And consider virtue-neutral action, a_h , expressing motivation m_h . We may then specify the following virtue-theoretic constraint on an agent's utility function.²³

$$(V3) \text{ If } V(m_i) = \zeta, V(m_h) = 1 \text{ and } O_{ij} \equiv O_{hk}, \text{ then } u(O_{ij}) = \zeta u(O_{hk}).$$

According to this constraint, the utility of an outcome of a virtuous (vicious) action is the utility of that outcome produced by virtue-neutral means (i.e. where $V(m) = 1$) multiplied by some factor specific to the virtue (or vice) expressed by the action. (For vicious actions, $0 < V(m) < 1$; for virtuous action, $V(m) > 1$.) Where the values of V are sufficiently large (or small), we will

²² As in the case of deontological theories we discuss above, we ignore the complication of context-sensitivity here. It seems plausible to say that virtues are not ranked absolutely, but are variably appropriate to distinct situations. In one situation, a sense of fairness might be the most appropriate—and thus virtuous—motivation to act upon; in another, a sense of generosity may be of greater moral significance. A fully developed virtue theory should be able to specify what it is that triggers this variability.

²³ To facilitate our presentation of this model, we assign disutilities over the open interval (0,1).

have modelled the decision-making of a moral agent paying significant heed to the motivational character of each of the options they confront. Adjusting the values of V is a matter of determining just how much weight motivational considerations ought to have in determining virtuous action. By modelling virtue's contribution to utility through multiplicative factors rather than, say, through additive ones, we accommodate the intuition that the contribution virtuous or vicious motivation makes to decision making should be proportional to the stakes involved.²⁴ Expression of malice through some relatively trivial act, say spitting in a person's soup, should attract proportionally less disutility than highly significant expression of malice, say falsely informing on a successful rival. It is not only the consequences of acts which should be assigned variable utility, the malice itself appears more or less significant in proportion to the stakes.

There is a subtlety here we need to finesse. Utility functions are unique only up to positive linear transformation and yet our multiplicative factor depends on the particular utility values given. For instance, we could rescale our utility function by multiplying through by 100 and we should get the same decision (i.e., the same action will have the maximum expected utility, if there is such an action, though the value of the expected utilities will be different from those before the transformation). The problem is that if we want our decision to be invariant under such a rescaling (as surely we do), we must choose our multiplicative factor for the particular utility function given—rescale the utility function and we must change the multiplicative factor. This is an inconvenience but a necessary one, given the approach we've adopted to this problem.

²⁴ Though perhaps not everyone shares this intuition. If not, we welcome other approaches (perhaps employing additive factors) in order to address the problem at hand.

V_1 to V_3 modify prior utility functions, but not all prior utility functions are compatible with virtuous agency. It is possible for an agent to bring to the table such perverse prior utilities that otherwise plausible versions of V_3 generate intuitively vicious decisions. For example, a fanatical Nazi might regret the coldness and harshness required to pursue the destruction of European Jewry, and yet hold this goal to be sufficiently important that his strong motivational scruples are overwhelmed. We therefore need to introduce further constraints on the utility functions of virtuous agents. A complete and adequate virtue theory will probably furnish these constraints in terms of the legitimate ends of virtuous agency. This might be characterised in terms of the pursuit of agent-neutral goods, or in neo-Aristotelian versions, as a nested series of agent-specific ends leading to the condition of eudaimonia, or the leading of a good and fulfilled life. Let us describe the former case. One example of this is a version of morality as universal benevolence.²⁵ The primary virtuous motivations here are benevolence and generosity and motivation-independent prior utilities are the familiar utilitarian ones. We introduce two additional constraints on the utility function to model these prior utilities. The extra constraints are as follows:

(V4) Let acts a_i and a_k be virtue-neutral, such that they each express motivation m , where $V(m) = 1$. If O_{ij} involves greater total welfare than O_{kl} , then any admissible utility function u must be such that $u(O_{ij}) > u(O_{kl})$.

²⁵ Michael Slote (2001) attributes such a theory to James Martineau.

(V5) Let acts a_i and a_k be virtue-neutral, such that they each express motivation m , where $V(m) = 1$. If O_{ij} involves the same total welfare as O_{kl} , then any admissible utility function u must be such that $u(O_{ij}) = u(O_{kl})$.

There are still substantial issues of what the various virtues are and how to resolve different virtues into motivational weightings and prior-utility settings. We set such complications aside for now. The details of these interactions is the business of virtue ethics, and it is simply not our task to prejudice the question of how to resolve open or difficult questions within the ethical theories. Our task is merely to show how each theory, *once suitably spelled out by the advocates of the ethical theory in question*, might be modelled in an appropriate fashion in the decision theory framework. So bearing this in mind, we take it that the above axioms constitute a plausible start to the problem at hand.

4. ADEQUACY OF THE MODELS

It is common to distinguish two quite different kinds of model: descriptive models and explanatory models. A descriptive model is a model that's empirically adequate in the sense that it gets the data right (or nearly right). An explanatory model needs to shed light on the underlying reasons for the way the system in question behaves as it does.²⁶ We don't propose that the

²⁶ A couple of examples might help. A purely mathematical description of the growth of a population, in terms of, say, the logistic equation, may be empirically adequate in that such a model makes correct predictions about the abundance of the population in question. But without a story about why the population abundance can be described by the logistic equation, the model fails to be explanatory. On the other hand, an explanatory model might lay bare

distinction between these two types of models is sharp—it certainly is not—but it does provide a useful way to think about the purpose and role of theoretical models.²⁷ With this distinction in mind, let's turn to the adequacy of the three models of ethical decision making we've presented in this paper.

The formal constraints on the utility function that we've proposed above, we take it, amount to reasonable ways of representing the three ethical theories in question. At least the formal constraints are a credible first shot at representing the theories in question. But it is important to note that all we have done is provide a framework that is able to model the preferences of the utilitarian, deontologist, and virtue theorist, as well as the outcomes of their decision making processes; we have not even attempted to model their underlying thought processes and motivations. Nor have we modelled the justifications moral theories furnish agents. Moral theories not only aim at specifying moral behaviour, they aim to supply justifications for moral behaviour. Only our model of utilitarianism furnishes the means of morally justifying actions. According to utilitarianism, a moral agent is always justified in optimally promoting general welfare, which is tracked by expected utility calculations employing an adequate utilitarian utility function. A utilitarian can thus use the fact that the expected utility of a particular action is greatest of all current options to justify their performing it.

the underlying biology and thus be explanatory (a story about carrying capacity, birth and death rates and so on), but may fail to deliver the predictive success of the logistic model.

²⁷ We put aside the issue of the place of normative models. All the models under discussion in this paper are normative (since they involve both ethics and rational decision making), but the descriptive–explanatory distinction is supposed to cut across the normative–descriptive distinction. (It is unfortunate that the word ‘descriptive’ is used in both these contexts—to contrast with both ‘normative’ and ‘explanatory’.)

For deontologists, however, appeal to the effect of the enormous disutility of prohibited options, which are given an arbitrary precision by nominating a specific disutility, has no justificatory power. These precise measures of disutility do not reflect deontological proposals about the precise relative disvalue of a prohibited act, for example. In our model, the disutility of a prohibited act applies only to an agent's current options. But if a person disvalued a prohibited kind of act in general, say lying, and used this valuation as a basis of moral decision-making, the disutility of lying should affect the utility of all outcomes involving lying, not just those involving the agent lying now. Although deontological determinations of right action can be modelled in terms of the pursuit of optimal expected utility, it does not follow that deontologists are motivated to optimise expected utility. Nor are they inclined, or equipped, to justify actions by appealing to the optimisation of expected utility. Deontologists, typically, think of morality as providing a series of constraints on behaviour based on their interpretation of what it is to respect another person, or on intuitions of rightness that are independent of their conception of the good.

Unlike deontologists, virtue theorists share an overall teleological approach with utilitarians. Thus, where deontological interventions on a utility function, in effect, replace normal utility assignments with assignments whose purpose is to ensure compliance or non-compliance with a rule, virtue-theoretic interventions modify the utility function by enhancing the value of virtuously derived outcomes and diminishing the value of viciously derived outcomes. However, a virtue-theorist is ill equipped to justify their decisions in terms of the utilities thus specified. This is because, as with deontology, virtue-theoretic modifications of utility functions apply only to the options currently faced by a decision-maker. They do not describe the kind of agent-neutral value assessments (e.g., malicious actions make the world poorer) that make for plausible

justifications of value promotion. The virtue-theorist may be optimising expected utility under an appropriate description of this utility, but they are not equipped to justify their actions in terms of the promotion of general values expressed by these utilities. Virtue-theoretic justification of action must take a different form: appealing to the importance of self-respect, or the special responsibility each person has for their own character and its expression in action.

On the one hand, the fact that our models are only descriptive should not count against them. It is non-trivial and an interesting fact that we can provide a descriptive model of decision making by virtue theorists and deontologists, in particular. But we need to be careful not to over-interpret the models and thus overstate their significance. For instance, in all three models we've represented the decision-making process in terms of the maximisation of expected utility (with ethical constraints on the utility function). First, we should not read too much into the name 'utility function' or 'expected utility'. These are both just formal features of the model and may have nothing to do with utility in the usual sense. Indeed, this is so even for standard decision theory (without ethical constraints). The utility function is best thought of as an uninterpreted mathematical function constrained by the von Neumann-Morgenstern axioms. Typically the utility function is interpreted in the obvious way, as a measure of agent-neutral and time-insensitive values, but this is a further move—arguably a move away from a descriptive to an explanatory model—and this move can be resisted.

Second, and more importantly, we should not conclude from the fact that all three ethical theories are represented as maximising some quantity, that they are all consequentialism in disguise.²⁸ Virtue theory and, in particular, deontology had to be shoehorned into the consequentialist framework of decision theory. As we've argued, we have not attempted to provide explanatory models, and the explanations of the behaviour of the deontological and virtue agents, and of the justifications available to them, drawn from our models are either misrepresented or opaque. It would thus be a mistake to press further claims about deontology and virtue theory that depend on our having captured the motivations for the theories in question.²⁹ As with all models, it is important to remember that these are just models and that there is danger in reading off too much from the model.³⁰ Having said this, however, there is still a significant issue of how explanatory the models we've presented are. We leave this issue for another occasion.

Finally, we say a few words about why we've approached the task of this paper by placing further constraints on the utility function rather than on preferences. After all, it might be argued that the axiomatisation of preferences is the more fundamental in the von Neumann-Morgenstern theory. Moreover, both the deontologist and the virtue ethicist might complain about the numerical character of the utility function and taking such a rich mathematical structure for granted in representing their ethical theories. The deontologist and the virtue theorist do not

²⁸ Oddie and Milne (1991) and Louise (2004) draw conclusions along these lines from the representability of ethical theories in a consequentialist framework.

²⁹ Such as that the theories are really just consequentialism after all, that they are lacking motivation, or that they have implausible motivations.

³⁰ We do not, for example, conclude that fluids are incompressible because our model of fluid flow assumes this, or that Sydney has no hills because our street directory of Sydney has no hills.

countenance such numerical representations and so we misrepresent them right from the start. They might be more sanguine about taking preferences as basic, since these do not have the numerical character of utility functions. We have a couple of things to say in response to such concerns. First, we admit that placing further constraints on the axioms for preferences may well be a fruitful way to approach this problem. We are not claiming that the approach of this paper is the only way to achieve the reconciliation of ethics and decision theory. Pointing out that there may be other ways to approach the task in question does nothing to undermine our project. Indeed, were such a project to be carried out, it would be fascinating to compare it to the approach we've suggested in this paper. As for the charge that we've misrepresented the deontologist and the virtue ethicist from the start by starting with the utility function, we point out that for these two ethical theories to be reconciled with decision theory somewhere along the line they will need to buy into utility functions. It strikes us as irrelevant whether utility functions are bought at the start or later on.³¹ Finally, we reiterate our earlier remarks about the nature of the models we are proposing here; they are intended to be descriptive, in the sense that they faithfully represent the ethical decisions and not the moral psychology of the agents making the decisions.

In this paper we have shown that, despite initial appearances, deontology and virtue theory can be accommodated in something like the standard decision-theory framework, and thus decision theory need not be thought of as a tool available only to the consequentialist. As it stands, decision theory is silent on ethical matters, but with a little work it can be made to accommodate

³¹ If the worry is that hard-line deontologists and virtue ethicists will not buy into utility functions at *any* stage, then the game is over. The prospect of accommodating such versions of deontology and virtue ethics with decision theory

each of the major ethical theories. While the generality of decision theory is well known, if what we've been arguing for in this paper is correct, there is another way in which decision theory exhibits generality. It is able to accommodate ethical theories without prejudicing the question of which ethical theory ought to be preferred and without prejudicing the question of how the details of these ethical theories ought be spelled out. Decision theory, it would seem, is genuinely neutral with respect to ethical considerations. And this is good news for everyone.³²

References

- Broome, J. 1991. "The Structure of Good: Decision Theory and Ethics" in M. Bacharach and S. Hurly (eds.) *Foundations of Decision Theory*, Blackwell, Oxford.
- Colyvan, M., Regan, H.M. and Ferson, S. 2001. "Is it a Crime to Belong to a Reference Class?", *Journal of Political Philosophy*, 9: 168–181. Reprinted in H.E. Kyburg and M. Thalos (eds.), *Probability is the Very Guide of Life*, Open Court, Chicago, 2003, 331–347.

seems hopeless.

³² We'd like to thank audiences at the 2005 Australasian Association of Philosophy Conference at the University of Sydney, a workshop at the University of Queensland in 2005, a philosophy seminar at the Australian National University in 2005, a philosophy seminar at the University of Colorado at Boulder in 2006, and the 2006 CMU-Pittsburgh graduate conference. We are also grateful to Selim Berker, David Braddon-Mitchell, James Chase, Peter Forrest, David Gray, Alan Hájek, Drew Khlentzos, Julian Lamont, Jennie Louise, Gary Malinas, Graham Oddie, and Martin Rechenauer for very helpful conversations on the issues addressed in this paper or for comments on earlier drafts of the paper.

- Dreier, J. 2004. "Decision Theory and Morality" in A.R. Mele and P. Rawling (eds.), *Oxford Handbook of Rationality*, Oxford University Press, Oxford, chap. 9.
- Gillies, D. 2000. *Philosophical Theories of Probability*, Routledge, London.
- Hájek, A. 2003. "Waging War on Pascal's Wager", *Philosophical Review*, 113: 27–56.
- 2003a. "Interpretations of Probability" *The Stanford Encyclopedia of Philosophy*, (Summer 2003 edition) E.N. Zalta (ed.), URL= <<http://plato.stanford.edu/archives/sum2003/entries/probability-interpret/>>.
- Hursthouse, R. 1991. "Virtue Theory and Abortion", *Philosophy and Public Affairs*, 20: 223–246.
- 1999. *On Virtue Ethics*, Oxford University Press, Oxford.
- Jackson, F. 2001. "How Decision Theory Illuminates Assignments of Moral Responsibility", in N. Naffine, R.J. Owens, and J. Williams (eds.), *Intention in Law and Philosophy* Ashgate, Aldershot, pp. 19–36.
- Jackson, F. and Smith, M. To appear. "Absolutist Moral Theories and Uncertainty".
- Jeffrey, R.C. 1983. *The Logic of Decision*, 2nd edition, University of Chicago Press, Chicago.
- Kant, I. 1996. *Practical Philosophy*, translated and edited by M. Gregor, Cambridge: Cambridge University Press.
- Kaplan, M. 1996. *Decision Theory as Philosophy*, Cambridge University Press, Cambridge.
- Kolmogorov, A.N. 1956. *Foundations of the Theory of Probability*, 2nd English edition, Chelsea Publishing Company (first published in 1933).
- Louise, J. 2004. "Relativity of Value and the Consequentialist Umbrella", *The Philosophical Quarterly*, 54: 518–536.
- Oddie, G. and Milne, P. 1991. "Act and Value: Expectation and the Representability of Moral Theories", *Theoria*, 57: 42–76.
- Resnik, M. 1987. *Choices: an introduction to decision theory*, University of Minnesota Press, Minneapolis.
- Ross, W.D. 1967. *The Right and the Good*, Clarendon Press, Oxford.
- Sen, A. 1997. "Maximization and the Act of Choice", *Econometrica*. 65:745–779.
- Slote, M. 2001. *Morals from Motives*, Oxford University Press, Oxford.

Taylor, P. 1986. *Respect for Nature: A Theory of Environmental Ethics*, Princeton University Press, Princeton.

Von Neumann, J. and Morgenstern, O. 1944. *Theory of Games and Economic Behavior*, Princeton University Press, Princeton.