Ethical Theories and Moral Guidance

PEKKA VÄYRYNEN

University of California, Davis

Let the Guidance Constraint be the following norm for evaluating ethical theories: *Other things being at least roughly equal, ethical theories are better to the extent that they provide adequate moral guidance*. I offer an account of why ethical theories are subject to the Guidance Constraint, if indeed they are. We can explain central facts about adequate moral guidance, and their relevance to ethical theory, by appealing to certain forms of autonomy and fairness. This explanation is better than explanations that feature versions of the principle that 'ought' implies 'can'. In closing, I address the objection that my account is questionable because it makes ethical theories subject not merely to purely theoretical but also to morally substantive norms.

I. INTRODUCTION

What should we think of an ethical theory that places us in the predicament reflected in Russell Banks's short story 'Sarah Cole: A Type of Love Story'?

When you have never done a thing before and that thing is not clearly right or wrong, you frequently do not know if it is a cruel thing, you just go ahead and do it and maybe later you'll be able to determine whether you acted cruelly. That way you'll know if it was right or wrong of you to have done it in the first place; too late, of course, but at least you'll know.[1]

An ethical theory that makes right actions frequently inaccessible to us, or offers us mere hindsight in the face of moral novelty, uncertainty and difficulty, fails adequately to guide action. Is such an ethical theory worse than one that does provide adequate moral guidance?

We can distinguish two motivations for constructing ethical theories.[2] One is *practical:* we might want an ethical theory to guide action.[3] Another is *theoretical* and, specifically, *explanatory*: we might want an ethical theory to explain why certain actions are right and others wrong. To achieve the practical aim, a theory has to identify features of right or obligatory actions which are readily accessible at least to conscientious agents, so that they can determine what is right in particular cases by determining which options have the features in question. To achieve the explanatory aim, the theory has to identify

---

[1] Russell Banks, *Success Stories* (New York, 1986), p. 149.

[2] By 'an ethical theory' I mean a theory about the content of morality, such as a theory of what is morally right.

[3] To simplify discussion, I set aside other distinctive practical functions of normative judgements, such as the critical stance they provide on our immediate desires and aims.

features which *make* actions right or obligatory. These aims are distinct: it is necessary neither that the features that make right actions right are also readily accessible features that help us decide what is right in particular cases, nor that the latter features also are features that make right actions right. Thus it is not *incoherent* to ignore or discount the practical aim in evaluating ethical theories. An explanatory ethical theory may fail to give adequate guidance for acting rightly, and yet give a correct account of what is right. Utilitarian principles, for example, are often presented as standards of rightness but (for familiar reasons) not decision-procedures.[4]

We might, however, think that even theories that do not *aim* to provide adequate moral guidance are nonetheless better to the extent that they also *do* provide adequate guidance, instead of thinking that doing so is no merit at all. Ethical theories are, after all, practical in their subject matter. To treat the claim that an ethical theory gives insufficient guidance for doing what we ought as an *objection* to the theory is to rely on the following norm for evaluating ethical theories:

> *Guidance Constraint:* Other things being at least roughly equal, ethical theories are better to the extent that they provide adequate moral guidance.

My aim is to contribute to our understanding of *why* ethical theories are subject to the Guidance Constraint, *if* indeed they are, rather than to argue *that* they are. I shall largely ignore the sizeable literature on whether there *is* a guidance constraint on ethical theories, for the explanatory question that interests me here is usually neglected in that literature. I find such neglect odd. Whether we are entitled to use the Guidance Constraint in evaluating ethical theories depends on whether it tracks any characteristics about which we should care in an ethical theory.

I begin with some central facts about adequate moral guidance and the relation between ethical theories and moral guidance (§II). I proceed to argue that we can explain those facts on the basis of certain forms of autonomy and fairness (§III). I then argue that this account explains facts about moral guidance better than various versions of the principle that 'ought' implies 'can' (§IV). My overall claim is conditional: *if* ethical theories are subject to the Guidance Constraint, then the best explanation of this fact appeals to the forms of autonomy and fairness in question. In closing, I respond to the objection that my account of the Guidance Constraint is questionable because it makes ethical theories subject not merely to purely theoretical norms but also to morally substantive norms (§V).

---

[4] See e.g. R. Eugene Bales, 'Act Utilitarianism: Account of Right Making Characteristics or Decision-Making Procedure?', *American Philosophical Quarterly* 8 (1971).

II. WHAT IS ADEQUATE MORAL GUIDANCE?

An ethical theory gives adequate moral guidance if it makes reliable strategies for acting well – for doing the right thing for the right reasons in particular situations – available to practical thinking.[5] Accounts of why ethical theories are subject to the Guidance Constraint should explain three central facts about adequate guidance and their relevance to ethical theory.[6] (1) An ethical theory fails to provide adequate moral guidance if it evaluates certain actions as right and others as wrong without providing any sufficiently *reliable* direction for acting in conformity with those evaluations.[7] A theory that is unreliable will all too easily lead us to act wrongly.[8] (2) For an ethical theory to provide adequate guidance, an agent's acceptance of it should reliably lead her to perform right actions *on the basis of* her acceptance of it, and *for* the reasons why (according to the theory) right actions are right.[9] Even if doing the right thing for non-moral reasons (say, to avoid punishment)

---

[5] I owe the expression 'a reliable strategy for doing the right thing for the right reasons' to Sean McKeever and Michael Ridge, 'The Many Moral Particularisms', *Canadian Journal of Philosophy* 35 (2005), p. 86. By 'practical thinking' I mean processes of reasoning – and, more generally, rational transitions in thought which need not proceed explicitly from premises to conclusions – which conclude, if not in action, then at least in a decision (i.e. in the formation of an intention) to act.

[6] Moral theorists who deny that moral principles need be suitable for use in public justification are not thereby enemies of the Guidance Constraint. Because we know more than we can articulate, a reliable strategy for acting well may be available to us even if we are unable to articulate the considerations governing our deliberations in a way that public justification requires. If so, violating the publicity condition does not entail violating the Guidance Constraint.

[7] What it is for a strategy to be sufficiently reliable depends on two further questions. What is a sufficient *degree* of reliability? And what is a sufficient *scope* of reliability (that is, what is the range of possible worlds across which a strategy for acting well must be reliable to whatever is the requisite degree)? These questions lie downstream from my present purposes, but I suspect that the relevant standards of reliability will be analogous to those in epistemology.

[8] More precisely, an ethical theory that provides adequate moral guidance should provide reliable strategies for doing what it evaluates as *objectively* right in the sense of corresponding to what is in fact right in the agent's situation. For ethical theories do not merely tell us what is *subjectively* right in the sense corresponding to what an agent is justified in believing to be right in her situation. (According to some ethical theories, such as decision-theoretic consequentialism, what is objectively right in this sense is, nonetheless, partly a function of the agent's probability function. See e.g. Frank Jackson, 'Decision-Theoretic Consequentialism and the Nearest and Dearest Objection', *Ethics* 101 (1991).) To remain neutral on what is objectively right, I shall treat the Guidance Constraint as requiring an ethical theory to provide reliable strategies for doing what the theory *says* is objectively right. So understood, the constraint allows us to evaluate an ethical theory independently of whether its account of right action is correct.

[9] Or, at least, it should lead her to do so if ethical theory acceptance is sufficiently motivating. I set aside questions about the fit between the motivational demands of morality and the motivational powers of ethical theory acceptance. Thus I remain neutral on the issue whether the acceptance of an ethical theory involves not only certain belief-forming dispositions but also a range of deliberative, motivational and affective dispositions. For discussion, see Pekka Väyrynen, 'Usable Moral Principles', *Challenging Moral Particularism*, ed. M. Lance, M. Potrc, and V. Strahovnik (London, forthcoming).

were a reliable way to act rightly, it would not be a case of *moral* guidance.[10] (3) A reliable strategy for acting well must be *available*, in the sense I explain shortly, for use in agents' practical thinking. If an ethical theory only identifies features of right actions that are either inaccessible even to a conscientious agent or accessible only in hindsight, it provides no *useful* direction for acting well.

In what ways may an ethical theory provide adequate guidance? We saw that no necessary connection exists between the accessible features of right actions, by reference to which we can determine what actions are right in particular circumstances, and the features that make right actions right. So an ethical theory may satisfy the Guidance Constraint even if its standards of right action do not *directly* provide useful direction. For even if those standards are complex or difficult to apply, the theory may satisfy the constraint *indirectly*, by providing surrogate devices that we can follow in some way that is sensitive to their reliability. While such devices may take the form of moral principles, they may instead be guidelines for kinds of thinking, such as analogical reasoning, which need not rely on principles.[11] Strategies for acting well need not consist in moral principles.[12]

Whatever strategies for acting well are like, they need to cover only the circumstances the agent is likely to encounter in order for them to provide reliable guidance. A reliable guide need not cover all cases. It may have some false implications. It may even be purely heuristic. Yet a simplified guide may provide us with more reliable direction than an accurate but complex standard of right action. To illustrate, beings with limited cognitive capacities are prone to mistakes in utility calculations. Utilitarians from Mill on typically respond that, in so far as following commonsense moral precepts approximates following the principle of utility, the former provides us with a reliable and available strategy for maximizing utility. Thus, even if following utilitarian principles is not *itself* a reliable and available strategy for acting well, utilitarianism

---

[10] The Guidance Constraint is neutral with respect to which reasons are the right ones. Two ethical theories might be coextensive, and yet disagree about which features of right actions are those that make them right. To evaluate an ethical theory specifically with respect to the Guidance Constraint is to determine whether it provides adequate guidance for performing those actions that it *evaluates* as right, and performing them for the reasons it *identifies* as the right ones.

[11] For discussion, see e.g. Gerald Dworkin, 'Unprincipled Ethics', *Midwest Studies in Philosophy* 20 (1995).

[12] Thus, moral particularists need not reject the Guidance Constraint. They typically claim that moral principles provide poor moral guidance, but this does not mean that they reject the Guidance Constraint. Similarly, the potential objection that particularist ethical theories fail to provide adequate guidance is plainly distinct from the claim that particularists reject the constraint. See e.g. David McNaughton, *Moral Vision* (Oxford, 1988), pp. 190–3; Jonathan Dancy, *Moral Reasons* (Oxford, 1993), p. 64.

may satisfy the Guidance Constraint via familiar indirection moves. The point generalizes to a host of other ethical theories.[13]

To return to the availability component of adequate moral guidance, let us note that whether a strategy for acting well is available for use in an agent's practical thinking depends on her cognitive capacities.[14] Because agents differ in their cognitive capacities, a guide that provides useful direction for one type of moral agent might be useless (for example, too difficult to apply) for another, or useless in one kind of context but not in another for one and the same agent. So the availability clause states a demand for guides for particular kinds of agents in particular kinds of contexts.[15]

What does this demand amount to? Following a strategy for acting well requires agents to acquire information about their circumstances, and perform inferences and calculations using that information, in their decision-making. For a given strategy to be available to a given kind of agent, the conditions for its use cannot depend on information or inferences and calculations that are relevant to deciding what is right in a particular situation but inaccessible or unmanageable to the agent. Thus the Guidance Constraint implies something like the following availability condition:

> *Cognitive Condition*: For any strategy S for acting well and any type of moral agent A, S is available to A for use in her practical thinking only to the extent that satisfying the conditions for using S in one's practical thinking lies within the limits of A's cognitive capacities.

---

[13] For example, R. M. Hare's distinction between 'ordinary' and 'critical' moral thinking, though developed in the context of utilitarianism, makes a general two-tiered structure available to many other kinds of ethical theories whose standards of right action are difficult to apply directly. Examples would be theories that deploy ideal observer or ideal advisor devices, or the notion of the virtuous person. See R. M. Hare, *Moral Thinking* (Oxford, 1981).

[14] It does not follow from this that the content of standards of right action is itself constrained by the cognitive capacities of the agents bound by those standards. As we shall see in §§III–IV, my account of the Guidance Constraint avoids this controversial implication and this gives it an advantage over a competing account.

[15] See McKeever and Ridge, 'The Many Moral Particularisms', p. 86. This point, together with our earlier point that an ethical theory may satisfy the Guidance Constraint indirectly, allows ethical theories to provide different strategies for differentially competent moral agents. This possibility undermines the following objection (due to an anonymous referee) to the Guidance Constraint: if the same theory must be capable of guiding agents both on the upper and the lower ends of the spectrum of cognitive and moral competence, then those on the upper end might have to use a rougher-grained theory than they are capable of using. But in that case they might have to do worse with respect to acting well than their cognitive capacities allow. This objection fails. For example, if we follow some set of surrogate devices in some way that is sensitive to whether doing so is a good way of putting (say) the principle of utility into practice, we are plausibly said to guide ourselves by the same theory as the more competent agents who are capable of using the utilitarian calculus directly. Yet the latter agents can do as well as they are capable of with respect to acting well.

To get a grip on the Cognitive Condition, note that variations in cognitive performance between people and between different times in a person's life have two importantly different sources. One is that cognitive performance is vulnerable to various distorting, but usually transient, conditions. Examples include the influence of bias, prejudice and morally irrelevant desire, as well as fatigue, listlessness, impatience, anxiety and laziness. The other is that people differ in their cognitive capacities. In assessing whether a strategy for acting well is available to a particular kind of agent, the Cognitive Condition permits us to idealize only the conditions under which those agents exercise their cognitive capacities. Idealizing those very capacities might make satisfying the conditions for the use of the strategy unavailable to the actual agents, given their limited capacities.[16]

As stated, the Cognitive Condition does not say to what kinds of agents an ethical theory that gives adequate moral guidance should make strategies for acting well available. The condition might seem toothless if it did not constrain the range of these agents. For example, the fundamental principle of direct objective act-utilitarianism is directly available to the practical thinking of the exceedingly rare kind of agents who possess the empirical information and the reasoning capacities required for correctly calculating which of the available alternatives the principle requires them to choose.[17] If the Guidance Constraint is to have teeth, we should relativize the Cognitive Condition not (merely) to an ethical theory but (also) to some limits in the people's cognitive capacities.

The relevant cognitive capacities are those of *normal* moral agents. Among actual beings, normal human adults tend to be those who are morally competent to at least a minimally sufficient degree.[18] Thus, no ethical theory provides adequate guidance in actual life unless it makes reliable strategies for acting well available to normal moral agents. But one intuition behind the Guidance Constraint is this: an ethical theory is in one respect worse if it tells us to act in certain ways but provides no strategy that we can use for acting in those ways. Since the relativization I propose captures this intuition, I shall now take it as understood when I speak of the Cognitive Condition.

---

[16] As Richard Foley says in a different context, 'we can legitimately idealize away various distorting conditions, ones that interfere with the operation of these abilities and skills . . . . But we are to avoid idealizations of the abilities and skills themselves.' See Richard Foley, *Working Without a Net* (New York, 1993), p. 160.

[17] See Erik Carlson, 'Deliberation, Foreknowledge, and Morality as a Guide to Action', *Erkenntnis* 57 (2002), p. 74.

[18] Limits of space force me to leave the complicated notion of normality intuitive. I discuss normality, though in a different context, in Pekka Väyrynen, 'Particularism and Default Reasons', *Ethical Theory and Moral Practice* 7 (2004), pp. 58–66.

III. AUTONOMY, FAIRNESS AND MORAL GUIDANCE

Now that we have a decent idea of what adequate moral guidance is, I shall argue that if ethical theories are subject to the Guidance Constraint, the best explanation of this fact features certain forms of autonomy and fairness.

Consider an agent who accepts a given ethical theory and understands at least those aspects of it that are crucial for determining which particular actions are (according to the theory) right and wrong. Such an agent has certain moral ideals: ends, values, concerns, commitments and ideals that she, in virtue of accepting the theory, prefers to guide her decisions and conduct. If the theory fails to make a reliable strategy for acting well available to the agent, she cannot, even in her conscientious moments, adequately guide herself by her moral ideals; her actions will only haphazardly serve those ideals. But if a conscientious morally committed agent cannot guide herself by her moral ideals, she must guide herself by such generally unreliable strategies for acting well as whims, coin tosses, self-interest, etiquette, tradition, religious authority or 'The Ethicist' column. An agent who is guided by such strategies will fail to be self-governing with respect to her own moral ideals.

Self-governance is something we often have in mind when we talk about *autonomy*.[19] Thus, in a clear sense, the inability to guide one's actions on the basis of one's moral ideals undermines the ability of the agents whom it strikes to act autonomously in the light of their moral ideals.[20] Of course, not all values are moral values. An agent can act autonomously in the light of her non-moral values even if the ethical theory she accepts fails to make a reliable strategy for acting well available to her. Even so, an ethical theory provides adequate moral

---

[19] While I focus on autonomy understood as self-governance, people have, of course, characterized autonomy in a wide variety of ways. Gerald Dworkin concludes a survey of characterizations with mild sarcasm: 'About the only features held constant from one author to another are that autonomy is a feature of persons and that it is a desirable quality to have.' See Dworkin, *The Theory and Practice of Autonomy* (Cambridge, 1988), p. 6. Nomy Arpaly lists *eight* different things that have all been called 'autonomy'. See Arpaly, *Unprincipled Virtue* (Oxford, 2003), pp. 118–25. I was first alerted to the bearing of self-governance on the guidance of action by Holly M. Smith, 'Making Moral Decisions', *Noûs* 22 (1988), p. 105, and Alan H. Goldman, 'Rules and Moral Reasoning', *Synthese* 117 (1999), p. 236.

[20] Sometimes by 'autonomy' we mean responsiveness to reasons. See e.g. George Sher, 'Liberal Neutrality and the Value of Autonomy', *Contemporary Political and Social Philosophy*, ed. E. F. Paul, F. D. Miller, Jr., and J. Paul (Cambridge, 1995). While we should not identify self-governance with responsiveness to reasons, an account of the Guidance Constraint that appeals to self-governance is entitled to subsume both under the general rubric of autonomy. An agent who accepts an ethical theory *is* self-governing, so far as her moral ideals go, in so far as her practical thinking responds, in virtue of her acceptance of the theory, to (what, according to the theory, are) the right reasons.

guidance only if agents who accept the theory can act autonomously on the basis of, specifically, their moral values.[21]

The role of self-governance in agency helps to explain why an ethical theory is better to the extent that it makes reliable strategies for acting well available to an agent who accepts it. Self-governance matters to us because (a) it plays an important role in our ability reliably to serve our ends and (b) without that capacity (or its effective operation) we would not be *agents* at all. Regarding (b), self-governance is at least a feature of the effective exercise of the capacities constitutive of agency, if not partly constitutive of agency itself. Regarding (a), an agent is in a position autonomously to serve certain of her ends, namely her moral ideals, only in so far as the ethical theory she accepts provides a strategy for acting well that identifies certain features of right actions: accessible features (say, helping someone in need or avoiding causing pain) such that forming intentions on their basis is a reliable way for an agent to form intentions to perform right actions. An ethical theory that fails to accommodate (a) and (b) fails to protect autonomy understood as self-governance.

A full explanation of facts about adequate moral guidance requires more than the autonomy condition, however. Considerations of autonomy alone are silent on the range of agents to whom an ethical theory should provide adequate moral guidance. But if the Guidance Constraint implies the Cognitive Condition, it requires ethical theories to provide adequate guidance to normal moral agents. We can explain why the Guidance Constraint requires an ethical theory to enable any normal agent who accepts the theory to govern herself on the basis of her moral ideals if we think, further, that the Guidance Constraint rests in part on what I shall (for lack of a better term) call *fairness* in the provision of the opportunity for morally committed moral agents to act well autonomously.

---

[21] These claims are independent of any particular conception of autonomy. But assessing whether an ethical theory provides adequate guidance may have to rely on auxiliary assumptions about autonomy. Suppose, for example, that we reject 'content-neutral' or purely 'structural' conceptions of autonomy, according to which autonomy is compatible with adopting any ideals or values. See e.g. Harry Frankfurt, 'Freedom of the Will and the Concept of a Person', *Journal of Philosophy* 68 (1971); and Dworkin, *Autonomy*. Then our conception of autonomy carries commitments that may restrict the range of ethical theories that we can autonomously accept. For then being autonomous is not compatible with the adoption of just any ideals or values, but either (i) requires the adoption of certain ideals or values, such as perhaps one's own dignity as a person, or (ii) rules out the adoption of certain ideals or values, such as perhaps servility or self-effacement (see e.g. Sigurdur Kristinsson, 'The Limits of Neutrality: Toward a Weakly Substantive Account of Autonomy', *Canadian Journal of Philosophy* 30 (2000)), or at least (iii) carries some normative presuppositions (see e.g. Henry Richardson, 'Autonomy's Many Normative Presuppositions', *American Philosophical Quarterly* 38 (2001)).

This form of fairness is important if an ethical theory should make a successful moral life available to more or less any subject of its moral requirements. Such fairness seems fundamental to morality. Any ethical theory implies that agents who without a good excuse act wrongly are morally blameworthy. A theory that makes no reliable strategy for acting well available even to conscientious normal agents may excuse their failure to act well. But excuses are typically applicable only in the context of wrongdoing. If a life which is rife with wrongdoing but to which little blameworthiness accrues counted as a successful moral life, such a life would be the stuff of minimal achievement. More plausibly, the measure of a successful moral life is the extent to which it consists in morally right actions. Then the idea that an ethical theory should make a successful moral life available to the subjects of its moral requirements requires that, for more or less any normal moral agent, the theory makes some or other reliable strategy for acting well available to her. A theory that meets the fairness condition satisfies this requirement. Thus, fairness in the provision of opportunities for leading a successful moral life helps to explain why an ethical theory should make strategies for acting well available to more or less any normal moral agent.[22]

The fairness condition explains also why the Cognitive Condition is the relevant availability condition in the Guidance Constraint. If an ethical theory should provide adequate moral guidance in order that more or less any normal agent can govern herself by the moral ideals she would have if she accepted the theory, it should provide strategies for acting well that accommodate the cognitive limits of normal agents. To require normal agents to act in certain ways, when the use of the strategies for acting in those ways depend on factors that are cognitively inaccessible to normal agents, is unfair. It is unfair because it makes a successful moral life (one consisting in morally right actions) unavailable even to conscientious normal agents, except by luck or happenstance.[23]

At the same time, the fairness condition explains why the Cognitive Condition allows ethical theories to reflect realistic differences in people's cognitive capacities. For it helps to make sense of what conditions for the use of strategies for acting well are reasonable. Consider, for analogy, that we can reasonably demand certain things of anyone holding a driver's license.[24] Those who get a driver's license

---

[22] For a theory to meet the fairness condition is alone sufficient for it to meet the Guidance Constraint *if* a successful moral life consists in *autonomously* performed right actions. But I need not take a stand on this latter claim.

[23] It also compromises the ability of agents who accept the theory to act autonomously in the light of their moral ideals or values.

[24] I am grateful to Terence Irwin for suggesting this analogy to me.

differ in reaction time, night vision, and the like, but we can reasonably require anyone to meet some threshold in order to qualify. If some can meet it only by taking defensive driving courses or installing some technical device in their vehicle, they are reasonably required to do so. In the moral case, we can reasonably require agents to inquire into the nature of the situation when deciding what to do. The existence of epistemic couch potatoes who care not a whit about the relevant facts or about proportioning their beliefs to their evidence does not dictate what sort of inquiry the use of a strategy for acting well can reasonably demand of agents. Moreover, while people's cognitive performance may fall short of their competence under distorting conditions, we can reasonably require them to do their best to control for such distortions, so as to achieve favourable conditions for judgement. Since my fairness condition proportions the requirement to do one's best to normal agents' competence, it allows that meeting the requirement may be easier for some agents than others.[25] What a strategy for acting well can reasonably require of agents depends on what normal agents can achieve with due care and diligence in exercising their limited cognitive capacities.

Given this background, the autonomy condition explains why an ethical theory does not violate the Guidance Constraint when normal agents' failure to use a strategy for acting well is due to internal cognitive or temperamental interferences with the operation of their cognitive capacities. A normal agent who conscientiously follows a strategy for acting well may fail to choose or act autonomously in the light of her moral ideals if she lacks the mental power to choose the action she judges best or lacks the power to perform bodily movements that, if she were successful, would constitute doing what she has decided to do. Since such failures in the 'internal' conditions of autonomy mark no shortcoming in a strategy for acting well, but only in the agent, they do not show that the strategy undermines normal

[25] Here we can see why fairness provides a better basis than *equality* for the requirement that ethical theories should make a reliable strategy for acting well available to more or less any normal moral agent. Intuitively, the Guidance Constraint permits an ethical theory to provide better moral guidance to some agents than others, so long as it provides adequate moral guidance to more or less everyone. Consider the analogy (suggested to me by Terence Irwin) that we can fairly expect everyone in a class to pass a swimming test, even if it is easier for some people than others, and even if some will do better at it than others, provided that it is open to everyone to do well enough. Similarly, an ethical theory may provide better or easier moral guidance to some persons than others without being unfair, so long as it provides *good enough* guidance to everyone. The Guidance Constraint allows this, since an ethical theory may satisfy the constraint by providing differentially competent moral agents with different strategies for acting well (see n. 15). Even if a given strategy for acting well happens to be more reliable or more available to some agents than others, it may still provide, for more or less any normal agent, some sufficiently reliable strategy for acting well that is available to her.

agents' capacity to act on the basis of their moral ideals.[26] Thus, my double appeal to autonomy and fairness explains why it is that an ethical theory violates the Cognitive Condition only in so far as the reason why normal agents fail to be self-governing is that the conditions for the use of the theory's strategies for acting well lie beyond the limits of such agents' cognitive capacities. Notice that we obtain this result without having to adjust the standards of right action themselves to the cognitive capacities of any agents, normal or otherwise.

IV. THE OUGHT–CAN PRINCIPLE AND MORAL GUIDANCE

So far, I have shown that we can explain central facts about adequate moral guidance, and describe their relevance to ethical theory, in terms of certain forms of autonomy and fairness. I shall now argue that this explanation is better than the natural and tempting alternative that ethical theories are subject to the Guidance Constraint because 'ought' implies 'can'. We can state this principle like this:

(OC)   Necessarily, for any agent A, any action φ and any time t: A ought (morally) to φ at t only if A can φ at t.

My understanding of (OC) is not exotic. I take it that 'ought' expresses objective moral obligation or requirement, 'can' expresses ability plus opportunity of some sort, and the two are related synchronically: an obligation at a given time is related with ability plus opportunity at that time.[27] I also grant that there is some reading of 'can' such that 'ought' implies 'can' by *conceptual* necessity. The attraction of accounting for facts about adequate moral guidance in terms of (OC), so read, is twofold. Unless 'ought' were essentially used to guide agents, it might be difficult to see why 'ought' should apply, as (OC) says it does, only to actions that agents can perform. And if (OC) holds by conceptual necessity, then (OC) lacks the sorts of substantive normative commitments that my own account seems to carry. But, as I shall argue, no version of (OC) both holds by conceptual necessity and explains facts about adequate moral guidance at least as well as my own proposal.

---

[26] Freedom from such interferences as coercion, indoctrination and manipulation are 'external' conditions of autonomy. Assessing an ethical theory in terms of the Guidance Constraint may abstract from external interferences. For more on internal and external conditions of autonomy, see e.g. David Copp, 'The Normativity of Self-Grounded Reason', *Social Philosophy and Policy* 22 (2005), p. 170.

[27] For discussion and massive references to literature, see e.g. Peter B. Vranas, 'I Ought, Therefore I Can', *Philosophical Studies* (forthcoming). For a formulation of (OC) according to which 'ought' implies 'can or could have', see Ishtiyaque Haji, *Deontic Morality and Control* (Cambridge, 2002), pp. 14, 50–1.

According to (OC), what one ought to do depends, by conceptual necessity, on what one can do. But what is the relevant sense of 'can'? That an agent can φ implies at least that φ-ing is *physically* open to her: she must have the requisite physical capacities to φ and be in a position to exercise those capacities. If I have no opportunity to save my fellow martyr from being ravaged by the lions because I am bound to a post, I cannot have a moral obligation to save her. When (OC) is read as the claim that a moral obligation to φ conceptually entails a physical ability plus opportunity to φ, (OC) explains the obvious fact that, in order for an ethical theory to guide an agent in a given situation, it must prescribe only actions that are physically available to the agent in that situation.[28]

If physical ability plus opportunity is *all* we read into 'can', (OC) cannot explain why ethical theories should make reliable strategies for acting well available to normal agents. On this 'fully factual' reading, nothing in (OC) guarantees that I will not be informationally impaired with respect to what I ought to do in particular circumstances. For a simple example, consider *Safe*: the only way to disarm a bomb whose detonation would kill millions of people is by opening the safe and pressing the green button inside.[29] Cases like *Safe* show that the fully factual reading of (OC) fails fully to explain facts about adequate moral guidance. So read, (OC) entails that I ought to open the safe. For although I do not know what I need to do in order to open the safe, I have the physical ability and opportunity to dial any combination, and hence the right combination. Indeed, I ought to open the safe even if these facts are *unknowable* to me. But a theory that assigns (whether correctly or not) such an obligation to me violates the Guidance Constraint. It can make no strategy for doing what I ought available to me. Thus, if we read (OC) so as to allow that right actions may have no features whose presence is reliably accessible to conscientious normal agents, then (OC) cannot explain why the Guidance Constraint should require ethical theories to make reliable strategies for acting well available for use in normal agents' practical thinking.

(OC) can explain why adequate moral guidance cannot depend on factors that a given kind of agent cannot access only if the claim that such an agent can φ implies also that φ-ing is *epistemically* open to her – that certain kinds of information pertaining to the act are available to her. On objective act-utilitarianism, for example, the relevant kind of

---

[28]   Notice that my account explains this fact equally well. If an agent could only serve her ends by performing physically unavailable actions, this would undermine her capacity to act autonomously in the light of her moral ideals.

[29]   I borrow *Safe* from Haji, *Deontic Morality*, p. 16.

information would serve reliably to identify the actions that maximize utility. Thus (OC) can even hope to explain facts about adequate guidance only if we read it as linking 'ought' also to knowledge about how the agent should act in order to do what she ought.[30] This suggests two relevant epistemic readings of 'can':

Actualist reading: A can φ at t only if A knows how to φ at t.

Modal reading: A can φ at t only if A can know how to φ at t.

The corresponding actualist reading of (OC) makes what an agent ought to do depend on what is *known* to her, whereas the corresponding modal reading of (OC) makes it depend on what is *knowable* to her. I shall argue that on neither reading does (OC) both explain facts about adequate moral guidance and hold by conceptual necessity.

The plausibility of the actualist epistemic reading of (OC) turns out to rely on the modal reading. Sometimes an agent does not, at a given time, know how to φ but nonetheless ought to φ, provided she *can* know how to φ. For sometimes an agent who does not know how to φ ought to know better; her ignorance is culpable. People who have little experience with pets ought nonetheless to know that one does not dry a puppy after a bath by throwing him in the microwave. By (OC), one ought to know how to φ only if one can know how to φ. Since this is what the modal epistemic reading of (OC) entails, the actualist reading implicitly relies on the modal reading for its plausibility.

The modal epistemic reading of (OC) entails that the limits of our cognitive capacities constrain the very standards of right action.[31] That is why it promises to explain why providing adequate moral guidance for an agent requires making a reliable strategy for acting well available to her. If an agent ought to do something only if she can (with sufficient reliability) know how to do it, then a reliable strategy for acting well would seem to be available to her.

When we assess the capacity of the modal reading to explain facts about adequate guidance, we run into the complication that claims about what an agent can know are open to many readings. According to the modal reading, is what an agent ought to do constrained by what she can in principle know? What she can know if her cognitive

---

[30] I take no stand on whether knowing how to φ entails or is a species of knowing that one can φ.

[31] James Griffin argues from (OC) to the strong claim that what actions are right is subject to knowledge constraints. See Griffin, *Value Judgement: Improving Our Ethical Beliefs* (Oxford, 1996), pp. 96, 106. According to Griffin, the point generalizes: 'There are no moral norms outside the boundary set by our capacities' (p. 100). For criticisms of such a use of (OC), see e.g. Robert Stern, 'Does "Ought" Imply "Can"? And Did Kant Think It Does?', *Utilitas* 16 (2004).

capacities are as developed as is possible for humans? What she can
know *qua* a conscientious normal agent? What she can know *qua* the
individual agent she is? What she can reasonably be expected to know
given her background information and circumstances? (And so on.) At
least some of these readings are clearly not equivalent. Thus we must
choose among them if we are to rely on a univocal reading of (OC).

  The problem with using any particular modal epistemic reading of
(OC) to explain facts about adequate moral guidance is that different
readings of what one can know have different morally substantive
implications for what one ought to do.[32] Selecting any one of them as
specifying *the* epistemic condition on what one ought to do saddles the
corresponding reading of (OC) with morally substantive implications
concerning what obligations one has.[33] We cannot simply *stipulate*
that the relevant epistemic condition concerns what conscientious
normal agents can know. For, as an intuitive element of the Guidance
Constraint, it is part of our present explanandum (recall §II). Moreover,
such a stipulation would make (OC) entail that the features that make
right actions right must be features to which we have reliable epistemic
access. But we cannot simply stipulate that the practical aim and the
explanatory aim for constructing ethical theories are not distinct after
all.

  If the choice of the relevant reading of 'can' carries morally sub-
stantive implications, it is implausible that the relation which (OC)
asserts between 'ought' and 'can' holds by conceptual necessity. If
some modal epistemic reading of (OC) both held by conceptual neces-
sity and carried substantive moral commitments, it would rule out
on conceptual grounds any ethical theory according to which truths
about how we ought to act are conceptually independent of what we
can know.[34] But if such theories are false, they seem to be substan-
tively rather than conceptually false. Moreover, although conceptual
connections can be opaque and surprising, a modal epistemic reading of
(OC) will assign to 'ought' entailments regarding what is knowable that

---

  [32] See Jonathan Bennett, *The Act Itself* (Oxford, 1995), pp. 53–4.
  [33] Alternatively, we might say that different epistemic readings of (OC) lead to a variety
of 'oughts' (what one ought to do by the lights of what one can know *qua* the individual
agent one is, and so on and so forth), but selecting any one of them as the reading that is
the most relevant to what one ought to do is a morally substantive choice. See Jackson,
'Decision-Theoretic Consequentialism', pp. 471–2.
  [34] Moral realists agree amongst themselves that realism is not the view that all ethical
truths are unknowable. The forms of moral realism alluded to in the text allow that some
ethical truths may be unknowable. Other forms of moral realism may hold that it is
not conceptually possible for ethical truth to transcend our capacity to discover it. One
such realist *may* be Thomas Nagel, but I find his remarks on the issue too ambiguous to
be confident one way or another. See Nagel, *The View from Nowhere* (New York, 1986),
p. 139. Thanks to an anonymous referee for reminding me of Nagel's discussion.

seem too fine-grained to be conceptual. For example, if we took (OC) to entail that, by conceptual necessity, one ought to φ only if conscientious normal agents can know how to φ, then it would be incoherent to claim that one ought to φ only if agents whose cognitive capacities are as developed as is possible for humans can know how to φ. The latter claim may be incorrect, but it surely seems coherent. Finally, consider someone who holds that having an obligation to φ is consistent with not being able to know how to φ, on the grounds that the inability excuses the agent's failure to φ and thereby makes blame inappropriate.[35] If some modal epistemic reading of (OC) were a conceptual truth, this claim would be incoherent. Whether or not the claim is correct, it surely seems coherent.

   I conclude that no reading of (OC) that can explain facts about adequate moral guidance holds by conceptual necessity, and vice versa.[36] To escape my argument, one might seek a reading of (OC) that explains facts about adequate moral guidance and cast it as a *moral* principle that does not hold by conceptual necessity.[37] I cannot rule out this response right off the bat, since my own account requires some morally substantive assumptions. My account is preferable, however. If

[35]   Note that the claim that epistemic inability is an excusing condition because it would be wrong to blame a person for not doing something she cannot know how to do fails to support the modal epistemic reading of (OC). Since claims of blameworthiness concern *agent*-evaluation but (OC) concerns *act*-evaluation, it is possible that an agent ought to φ even if she would not be blameworthy for not φ-ing. For discussion, see Walter Sinnott-Armstrong, ' "Ought" Conversationally Implies "Can" ', *Philosophical Review* 93 (1988), p. 250; and Stern, 'Does "Ought" Imply Can"?', p. 47.

[36]   Other options are to formulate (OC) in terms of *metaphysical entailment* (see e.g. Haji, *Deontic Morality,* pp. 13–4) or *conversational implicature* (see e.g. Sinnott-Armstrong, 'Conversationally'), or as the defeasible meta-theoretical norm that, other things being at least roughly equal, an ethical theory is better in so far as an agent can (in the specified sense) do what the theory says she ought to do (see James Brown, 'Moral Theory and the Ought-Can Principle', *Mind* 86 (1977), pp. 220–2). It is hard to see how any of these formulations would explain facts about adequate moral guidance much better than formulations in terms of conceptual entailment.

[37]   A different possible response is to maintain that (OC) holds by conceptual necessity, but hold that 'ought' conceptually *underdetermines* what must be knowable to the agent in order for it to be true that she can do something. On this view, the morally substantive implications of selecting a given interpretation of knowability as that which partly determines the limits of 'ought' are due not to (OC) but to the comprehensive theory of moral obligation that generates the choice. (This move follows an established methodology: many philosophers reject epistemic readings of (OC) in favor of theories of moral obligation that vindicate only the fully factual interpretation of (OC). See e.g. Haji, *Deontic Morality*, pp. 16–21.) On this view, the content of 'ought' need not make fine-grained distinctions regarding what is knowable. The explanation of facts about adequate moral guidance based on (OC) is more secure to the extent that this is the case, in view of the aforementioned substantive implications of such distinctions for what one ought to do. But to account for those facts by appeal to such a reading of (OC) would allow an ethical theory whose assignment of obligations satisfies (OC) to entail conditions for the use of its strategies for acting well that violate the Cognitive Condition. So this response will not help (OC) adequately to explain facts about adequate moral guidance.

we explain facts about adequate moral guidance without appealing to epistemic readings of (OC), we can separate the controversial issue of whether our cognitive capacities constrain what we ought to do from the uncontroversial point that they constrain what counts as adequate guidance. Moreover, the appeal to an epistemic reading of (OC), understood as a moral principle, is either arbitrary or not basic. Each interpretation of what it is for something to be knowable to an agent yields a distinct moral principle. (OC) can account for the Cognitive Condition only if we select among those principles one on which 'can φ' entails that conscientious normal agents can know how to φ. But, again, we cannot simply stipulate that this is the relevant principle. If there is no independent reason to select that particular principle, using it to explain why ethical theories are subject to the Guidance Constraint is arbitrary. But if there are independent reasons to select that particular principle, we can bypass (OC) and account for facts about adequate moral guidance directly in terms of those reasons. Until we see such independent reasons, and unless they are distinct from and more compelling than the forms of autonomy and fairness to which my account appeals, these forms of autonomy and fairness remain the best explanation of facts about adequate moral guidance and their relevance to ethical theory.

V. THE GUIDANCE CONSTRAINT AND MORAL METHODOLOGY

I shall close with the objection that if the best explanation of why ethical theories are subject to the Guidance Constraint carries morally substantive commitments, the constraint cannot constitute an adequacy condition on ethical theories. The objection stems from the idea that adequacy conditions should hold ethical theories answerable not to moral norms but only to purely theoretical, morally neutral norms. I cannot here address, let alone settle, the general issue of what kinds of norms are in play in the evaluation and selection of ethical theories. My aim is merely to allay worries about holding ethical theories answerable to moral norms in the way that my present proposal does.

Internal consistency, simplicity, unity and explanatory power are plausible candidates for purely theoretical norms for evaluating ethical theories: in so far as these features are advantages in a theory, they are so regardless of the theory's subject matter. My argument in the two preceding sections suggests that we cannot account for the Guidance Constraint in terms of purely theoretical norms.[38] The capacity of

---

[38] The fact that purely theoretical norms, such as simplicity, may conflict with the Guidance Constraint poses no problem, since the former norms may conflict with one

purely theoretical norms to discriminate among competing ethical theories is limited anyway. For example, many philosophers argue that an adequate ethical theory may not allow for situations in which an agent ought ('all things considered') both to do *A* and to do *B* (formally, *O*(*A* & *B*)), but in which she can only do either *A* or *B* but not both. The thought is that genuine moral dilemmas of this kind are impossible because the assumptions that generate their impossibility are impeccable. But a deontic logic that is consistent with the possibility of genuine moral dilemmas can itself be perfectly consistent, and need be no less simple or unified than one that entails the impossibility of moral dilemmas.[39] Reasons to deny or allow the possibility of moral dilemmas apparently have to stem from morally substantive rather than purely theoretical norms.

Given the limited capacity of purely theoretical norms to discriminate among competing ethical theories, it should be no surprise that many norms that we deploy widely in evaluating ethical theories are not morally neutral. Examples would be the familiar norms that, other things being at least roughly equal, an ethical theory is better to the extent that it accords with the firm, independently credible moral intuitions that we have (or would retain) after careful reflection and to the extent that it starts from attractive general beliefs about morality (some of which presumably endorse the Guidance Constraint).[40] These norms are not morally neutral because they reflect substantive moral ideals that enjoy at least provisional acceptance. Purifying our norms for evaluating ethical theories of morally substantive implications would require us to abandon norms that we deploy widely. If we are unwilling to abandon those norms, we must also be open to the idea that if ethical theories are better to the extent that they provide adequate moral guidance, then the best explanation of that fact features some morally substantive ideals.

One might still worry that holding ethical theories answerable to morally substantive norms is unreasonable because it begs the question against certain ethical theories. John Rawls suggests that the worry

another in the same way. If any complications arise here, they will concern the relative weights of different norms for evaluating ethical theories. See n. 43 below.

[39] A pair of such deontic logics, and a similar moral, can be found in John F. Horty, 'Reasoning with Moral Conflicts', *Noûs* 37 (2003). Geoffrey Sayre-McCord argues more generally that 'we can build a deontic logic only by accepting axioms and rules of inference that are incompatible with reasonable moral positions'. See Sayre-McCord, 'Deontic Logic and the Priority of Moral Theory', *Noûs* 20 (1986), p. 179. In debates about moral dilemmas, the most commonly disputed claims are the principle that 'ought' implies 'can' and the 'agglomeration principle' (*OA* & *OB*) → *O*(*A* & *B*).

[40] See e.g. John Rawls, *A Theory of Justice* (Cambridge, MA, 1971), pp. 48–52; and Brad Hooker, 'Intuitions and Moral Theorizing', *Ethical Intuitionism: Re-evaluations*, ed. P. Stratton-Lake (Oxford, 2002), pp. 168–70.

is less serious than it may seem. The merit of proposed adequacy conditions 'depends upon the soundness of the theory that results' and such conditions can be justified 'only by the reasonableness of the theory of which they are a part'.[41] Rawls's idea may to some extent frustrate our desire to establish adequacy conditions on ethical theories *before* particular theories come in, so that we can prevent some from entering. But its application to the Guidance Constraint is instructive.

Philosophers who deny that an ethical theory is any worse for failing to provide adequate moral guidance must also deny that an ethical theory is any worse for implying that what one ought to do is reliably inaccessible to conscientious normal agents. The resulting theory may be tenable, but given how widely it is seen as important for ethical theories to avoid thwarting these practical and epistemic aims, to claim that thwarting those aims is no cost at all in an ethical theory is to give oneself a bullet-infested diet. Such a theory would depart from an overlapping consensus on what aims it is important for ethical theories to serve and what adequacy conditions reasonable ethical theories impose upon themselves. If the practical aim of ethical theory turns out to be important for morally substantive reasons, it does not follow that the Guidance Constraint becomes unreasonable.

These comments reflect the methodological point that the assessment of ethical theories is a holistic affair. The judgement that it is *costly* for a theory to lack some feature typically relies on a large background of auxiliary assumptions that themselves are open to scrutiny. Whether we can reasonably hold ethical theories answerable to a given norm depends on what else we regard as reasonable. Thus I see no good reason for deciding in advance that only morally neutral assumptions can reasonably influence our view of norms for evaluating ethical theories. If my argument in this article is on the right track, exploring our grounds for holding ethical theories answerable to the Guidance Constraint reveals to us certain evaluative commitments, which we can and should subject to further scrutiny.

Let me conclude with a dialectical point. I have argued that the best explanation of why ethical theories are subject to the Guidance Constraint, if indeed they are, features certain forms of autonomy and fairness. If upon reflection we endorse the value of these forms of autonomy and fairness, we are likely to accept that ethical theories in fact are subject to the Guidance Constraint. (But, again, I do not claim to have established this last point here.) If the argument I have given is cogent, the only way to rule out my account of the Guidance Constraint is to argue that the relevant forms of autonomy

---

[41] See Rawls, *A Theory of Justice*, pp. 130–1.

and fairness are unimportant. Though coherent, this debunking option is costly. For example, as we saw in §III, it conflicts with the important role of autonomy in agency and that of fairness in the provision of opportunities for leading a successful moral life. But the debunking option shares some dialectical common ground with my account. The holistic aspect of moral methodology should lead us to expect that the Guidance Constraint is a *defeasible* norm for evaluating ethical theories, just like norms such as simplicity and unity are.[42] Thus a defeasible Guidance Constraint agrees with the debunking option that there can be some theoretical benefits that an ethical theory can gain by violating the constraint and for whose sake it can reasonably sacrifice some autonomy and fairness. On my account of the constraint, just which trade-offs would be reasonable becomes an issue only to those theories that value autonomy and fairness, since only those theories see losses in autonomy and fairness as *costs*. Which trade-offs would be reasonable is a complicated issue that is worth further reflection.[43] But it seems reasonable to allow that a fairly complex theory that rarely undermines autonomy and fairness may well be better than a fairly simple theory that often undermines them.[44] Why should it then be unreasonable to evaluate ethical theories partly in terms of whether they provide adequate moral guidance, if the Guidance Constraint turns out to be best grounded in some such morally substantive ideals as autonomy and fairness?[45]

pekkav@ucdavis.edu

---

[42] An anonymous referee asserted that internal consistency is an indefeasible norm. I need not disagree, since nothing that I say implies that *no* norm for evaluating ethical theories is indefeasible. But in fact the claim is controversial. Some writers argue that it need not be irrational temporarily to tolerate a noticed inconsistency in a set of beliefs when there are reasons to accept each belief, it is unclear which belief one should give up, and one can isolate the effects of the inconsistency on the rest of one's beliefs. See Gilbert Harman, *A Change in View* (Cambridge, MA, 1986). This view agrees, of course, that inconsistent beliefs cannot all be true together. But it denies that consistency is an indefeasible norm of rational belief-revision. As a view about belief-revision, one would expect it to have implications for moral methodology.

[43] Another important issue concerns the relative weights of the different norms for evaluating ethical theories discussed above. My argument does not require me to address this issue. The claim that ethical theories are subject to the Guidance Constraint, as well as accounts of why that is so, are neutral on what weight the constraint has relative to the other norms. On my account, its weight depends on *how* important autonomy and fairness are relative to other ideals.

[44] This is one explanation of why indirect forms of consequentialism typically are more plausible than its direct forms.

[45] I would like to thank Terence Irwin, Antti Kauppinen and several anonymous referees for useful feedback on earlier versions of this article.