Challenging Moral Particularism

Particularism is a justly popular 'cutting-edge' topic in contemporary ethics across the world. Many moral philosophers do not, in fact, support particularism (instead defending 'generalist' theories that rest on particular abstract moral principles), but nearly all would take it to be a position that continues to offer serious lessons and challenges, and that can not be safely ignored.

This collection of new philosophy papers, written by well known philosophers, will find a ready audience within the international academic philosophical community. Given the high standard of the contributions, and that this is a subject where lively debate continues to flourish, it is reasonable to expect that the book will become required reading for professionals and advanced students working in the area.

Matjaž Potrc is Professor of Philosophy at the University of Ljubljana, Slovenia.

Vojko Strahovnik is an Assistant Philosophy Researcher at the Department of Philosophy, University of Ljubljana.

Mark Lance is Professor of Philosophy at Georgetown University, USA.

Routledge Studies in Ethics and Moral Theory

Challenging Moral Particularism

**Edited by
Mark Norris Lance, Matjaž Potrč,
and Vojko Strahovnik**

Routledge
Taylor & Francis Group

NEW YORK AND LONDON

Contents

vi   *Contents*

Acknowledgements

Contributors

Robert Audi is the David E. Gallo chair in Ethics and Professor of Philosophy at University of Notre Dame. He is the author and editor of several books including *Belief, Justification, and Knowledge* (Wadsworth, 1988), *The Structure of Justification* (Cambridge, 1993), *Action, Intention, and Reason* (Cornell, 1993), *Moral Knowledge and Ethical Character* (Oxford, 1997), *Epistemology: A Contemporary Introduction to the Theory of Knowledge* (Routledge, 1998), *The Good in the Right* (Princeton, 2004), *Practical reasoning and Ethical Decision* (Routledge, 2006) and *The Cambridge Dictionary of Philosophy.*

David Bakhurst is John and Ella G. Charlton Professor of Philosophy at Queen's University, Kingston, Ontario. He is the author of *Consciousness and Revolution in Soviet Philosophy* (Cambridge, 1991), and co-editor of *The Social Self* (with C. Sypnowich, Sage, 1995) and *Jerome Bruner: Language, Culture, Self* (with S. Shanker, Sage, 2001). He has also published numerous articles on ethics, philosophical psychology, and Russian thought.

Jonathan Dancy is Professor of Philosophy at University of Reading (UK) and at University of Texas (Austin). He is the author of *An Introduction to Contemporary Epistemology* (Blackwell, 1985)*, Berkeley: An Introduction* (Blackwell, 1987), *Moral Reasons* (Blackwell, 1993)*, Practical Reality* (Clarendon, 2000)*, Ethics Without Principles* (Clarendon, 2004), and numerous articles. He is also the editor of *Perceptual Knowledge, Reading Parfit, Berkeley's Principles of Human Knowledge* and *Three Dialogues between Hylas and Philonous,* and co-editor of *A Companion to Epistemology.*

Brad Hooker is Professor of Philosophy at University of Reading (UK). He is the author of *Ideal Code, Real World: A Rule-Consequentialist Theory of Morality* (Clarendon, 2000), and editor of *Rationality, Rules, and Utility* (Westview, 1993), and *Truth in Ethics* (Blackwell, 1996), *Well-Being and Morality* (with R. Crisp, Oxford, 2000), *Morality, Rules, and Consequences* (with E. Mason and D.E. Miller, Edinburgh UP, 2000; Rowman and Littlefield, 2000), and *Moral Particularism* (with M. Little,

Clarendon, 2000). He has also published many articles, mostly in moral philosophy.

Terry Horgan is Professor of Philosophy at University of Arizona (Tuscon). He is the author of *Connectionism and the Philosophy of Psychology* (with J. Tienson, MIT, 1996), *Austere Realism* (with M. Potrč, MIT, forthcoming), and articles on many philosophical subjects including metaphysics, philosophy of mind, philosophy of psychology, epistemology, philosophy of language, metaethics. He is also editor of numerous collections including *Metaethics after Moore* (with M. Timmons, Oxford, 2006).

Mark Lance is Professor of Philosophy at Georgetown University. He is the author of *The Grammar of Meaning* (with J. O'Leary-Hawthorne, Cambridge, 1997) and of many articles in journals and books covering the topics of philosophy of language, philosophy of mind, metaphysics, epistemology, normativity, and philosophy of logic.

Margaret Little is Professor at Georgetown's Philosophy Department and a Senior Research Scholar at Georgetown's Kennedy Institute of Ethics. She is the editor of *Moral Particularism* (with B. Hooker, Clarendon, 2000) and author of *Intimate Duties: Re-Thinking Abortion, the Law, & Morality* (Clarendon, forthcoming). Her work covers areas of ethics, with particular interests in moral particularism, moral epistemology, motivation, and feminist bioethics.

Sean McKeever is Professor of Philosophy at Davidson College (North Carolina). He is the author of *Principled Ethics: Generalism as a Regulative Ideal*, (with M. Ridge, Oxford, 2006) and of several articles on moral theory, history of ethics, and political philosophy.

David McNaughton is Professor of Philosophy at Florida State University. He is the author of *Moral Vision: An Introduction to Ethics* (Blackwell, 1988), and articles on ethics, on 18th century British moral philosophy, history of philosophy, and on philosophy of religion. He is the founder of *British Society for Ethical Theory.* He is currently working with P. Rawling on a book on deontology.

Nenad Miščević is Professor of Philosophy at Central European University (Budapest) and University of Maribor. He has written and edited several books – including *Rationality and Cognition* (Toronto, 2000), *Nationalism and Ethnic Conflict: Philosophical Perspectives* (ed., Carus, 2000), *Nationalism and Beyond* (CEU, 2001). He has written papers covering philosophy of mind and cognitive science, epistemology, political philosophy, nationalism, metaphysics, and moral philosophy.

Matjaž Potrč is Professor of Philosophy at University of Ljubljana. His books include *Phenomenology and Cognitive Science* (Röll, 1993),

x   *Contributors*

Origins: The Common Sources of the Analytic and Phenomenological Traditions (ed., with T. Horgan and J. Tienson, Memphis, 2002), *Dynamical Philosophy* (ZIFF, 2004), *Practical Contexts* (with V. Strahovnik, Ontos, 2004) and *Austere Realism* (with T. Horgan, MIT, forthcoming). Articles cover metaphysics, moral philosophy, epistemology, philosophy of mind and cognitive science, vagueness, phenomenology, and Slovene philosophical tradition.

Anthony W. Price is Reader in Philosophy at University of London (Birkbeck). His bibliography includes *Mental Conflict* (Routledge, 1995), *Love and Friendship in Plato and Aristotle* (Oxford, 1997), and articles 'Plato: Ethics and Politics', 'On Criticising Values', 'Plato, Zeno, and the Object of Love', and 'The So-Called Logic of Practical Inference'.

Piers Rawling is Professor of Philosophy at Florida State University. He is the editor of *The Oxford Handbook of Rationality* (with A. Mele, Oxford, 2004) and author of articles covering ethics, decision and game theory, logic, philosophy of language, philosophy of science, rationality and reasons. Currently he is working with D. McNaughton on a book on deontology.

Michael Ridge is Professor of Philosophy at University of Edinburgh. He is the author of *Principled Ethics: Generalism as a Regulative Ideal*, (with S. McKeever, Oxford, 2006) and of several articles on moral theory, social and political philosophy, and history of moral philosophy.

Vojko Strahovnik is an Assistant Philosophy Researcher at the Department of Philosophy, University of Ljubljana (Faculty of Arts). He is the author of *Practical Contexts* (with M. Potrč, Ontos, 2004) and the author of several articles on ethics, epistemology and Slovene philosophical tradition.

Pekka Väyrynen is an Assistant Professor of Philosophy at the University of California, Davis. His publications and research interests fall mainly in the areas of ethics, rationality, and epistemology.

5   Usable moral principles[1]

Pekka Väyrynen

1 Introduction

One important strand in moral particularism concerns moral practice. We ought not, particularists maintain, to rely on moral principles in moral thought and judgment because they provide poor guidance for doing the right thing. Another important strand concerns the structure of the moral domain. We ought not, particularists maintain, to see moral theorizing as a project of stating and defending substantive principles concerning the rightness and wrongness of actions, the value of states of affairs, the fairness of societal arrangements, and so on. This is because (depending on the particularist) there are no true moral principles, or we have no good reason to expect there are any, or, even if there are true moral principles, moral facts and distinctions don't depend in any way on there being any. In other work, I defend a generalist account of the structure of the moral domain against the second strand in particularism by defending a novel kind of hedged moral principles that accommodate certain central insights of particularists, but nonetheless support a moderate form of generalism.[2] In this paper, I defend a generalist account of moral guidance against the first strand in particularism and, specifically, its claim of *principle abstinence* (or PA, for short): we ought not to rely on moral principles in moral judgment because they fail to provide adequate moral guidance.[3] My main aim is to show that the kind of hedged principles I defend elsewhere also provide adequate moral guidance, thereby counting as appropriately usable in moral thought. But I also hope that at least the broad outlines of my argument will be found acceptable to generalists more widely.

2 Two arguments for principle abstinence

Ethics aims at action, as the saying goes. Moral theories should enable us to comprehend aspects of ourselves and our world in ways that offer us guidance in our moral lives. Implicit in this thought is a meta-theoretical norm concerning an important and traditional role of moral theories: other things being at least roughly equal, moral theories are better to the extent that they

provide adequate moral guidance. If PA is true, then generalist moral the-
ories would seemingly fail to secure this action-guiding role for moral prin-
ciples, since principles would be poorly suited to enabling us to act rightly
or exhibit moral virtue. I'll begin by distinguishing two sorts of argument
strategy for PA. I'll argue that one is unsound and the other has yet to be
made out in the literature.

The first strategy is to base PA on the theoretical strand in particularism.
Jonathan Dancy, for one, offers this sort of argument: 'Particularism claims
that generalism is the cause of many bad moral decisions. ... Reasons
function in new ways on new occasions, and if we don't recognize this fact
and adapt our practice to it, we will make bad decisions' (1993: 64).[4] This is
to infer PA from *reasons holism*, the general doctrine about reasons that no
necessary connection exists between the property of being a reason and the
property of always being the same kind of reason (either positive or nega-
tive). The core thought is that applying reasons holism to the moral domain
helps to show that there are no (or very few) true general principles that are
well suited to enabling us to act rightly or exhibit moral virtue.[5]

We could construct several arguments for PA that instantiate the above
strategy. But we needn't bother, because the general strategy is unsound. It
is unsound because reasons holism is available to generalists and particu-
larists alike. Certain kinds of (true) substantive moral principles can exist
even if reasons holism is true.[6] It remains unsound even if we weaken par-
ticularism to the claim that, even if there are true principles, moral facts and
distinctions in no way depend on there being any.[7] For this form of parti-
cularism leaves open the possibility that there are moral patterns that can be
accurately captured by principles or generalizations that also are well suited
to enabling us to act rightly and observe moral distinctions, such as that
between right and wrong, in the ways of the practically wise. Moral princi-
ples may, therefore, be able to function as guides even if, contrary to gen-
eralism as I'll understand it, moral facts and distinctions don't depend on
the existence of a comprehensive set of true substantive moral principles.
(By 'substantive moral principle' I mean a (synthetic) proposition that
identifies conditions or properties in virtue of which something has a given
moral property such as rightness, and which are thus explanatory of why it
is right. A set *S* of principles is comprehensive in the relevant sense if, for
any particular moral fact *M*, a principle or a set of principles in *S* is
required for *M* to hold.[8]) As a view in moral metaphysics, generalism
implies no particular view of the relation between moral thought and prin-
ciples. For example, it doesn't imply that the only way people can make
moral judgments at all is by basing them on principles.

The second strategy is to base PA on the idea that, even if moral princi-
ples play a necessary role in accounting for the nature and basis of moral
facts and distinctions, they might have some features other than inaccuracy
or context-insensitivity which make them unsuitable, or at least unnecessary,
for enabling us to act rightly or exhibit moral virtue. One *prima facie* hurdle

for this strategy is that, independently of their role in moral theory, moral principles might be useful practical tools in solving various problems of interpersonal assurance, coordination, and the like.[9] Another is that the most we get from particularists by way of features that allegedly make principles provide poor moral guidance, even if generalism is true, is the claim that principles encourage us to make moral decisions without a sensitive and detailed examination of particular cases.[10] I'll argue in §6 that this particular claim is false because acceptance of moral principles entails a commitment to developing one's moral sensitivity and judgment. But, for all we have seen so far, principles might indeed have features that make them provide poor moral guidance. To assess this possibility, we need to know what constitutes adequate guidance and what it would be for moral principles to provide it.

3 Moral principles and adequate moral guidance

A principle might accurately identify conditions that explain why some such moral property as rightness is instantiated, and so pull its theoretical weight as a standard of right action, without being of much use for agents in their practical thinking. (By 'practical thinking' I mean rational transitions in thought which terminate, if not in action, then at least in a decision to act in a certain way, but which needn't proceed explicitly from premises to conclusions.) Conversely, in order to provide guidance, a generalization needn't identify such conditions with full accuracy or cover all cases. The explanatory and the guiding functions of moral principles are logically distinct. In order for moral principles to function as guides, they need only to provide appropriate guidance to conscientious agents who care about living up to the principles they accept.

Intuitively, moral guidance amounts to offering more than hindsight in the face of moral novelty, uncertainty, and difficulty; it amounts to offering some *strategy* for acting well. More precisely, for a given standard of right action to provide adequate guidance is for it to contribute non-trivially to a reliable overall strategy for doing the right thing for the right reasons that is available to the practical thinking of conscientious, morally committed agents.[11] (I'll abbreviate this idea as 'contributing to a reliable strategy', and 'doing the right thing for the right reasons' as 'acting well', except when greater precision matters.)

We can explicate this idea of what constitutes adequate moral guidance as follows.[12] First, an overall strategy for acting rightly must be reliable to some sufficient degree, because following an unreliable strategy would all too easily direct us to wrong actions. Whatever the sufficient degree of reliability is, the relevant kind of reliability is *conditional reliability:* any strategy for acting well is reliable only to the extent that we operate on an accurate (non-moral) conception of our circumstances.[13] Regarding the scope of reliability, it seems plausible that an overall strategy should exhibit

robust reliability: it should be reliable across some fairly wide range of pos-
sible worlds, and so its reliability should be correspondingly independent of
the world the agent happens to inhabit.[14] Robust reliability allows that just
what generalizations will figure in a reliable overall strategy may vary across
worlds. Someone who adopts a 'shoot first and ask questions later' policy,
as part of their overall strategy for acting well in a world in which adhering
to the policy just happens reliably to make them perform right actions, isn't
deploying a reliable overall strategy unless their acceptance of the policy is
sensitive to the fact that such a world is, contingently, the world that they
happen to inhabit.[15] If they would just as well have adopted this policy had
they lived in a peaceful world where the policy would reliably lead to wrong
actions, then theirs isn't a case of moral guidance in the relevant sense. For,
given the truth of this counterfactual, the policy 'shoot first and ask ques-
tions later' would make no contribution to the reliability of their overall
strategy for acting well. The overall package including the selection and
deployment of guides isn't robustly reliable in this case. The rationale for
the requirement that a guiding generalization must not merely be *part of*
a reliable strategy for acting well but must *contribute* to its reliability is
that a generalization that fails to play the latter role may be an idle wheel in
guidance.

Second, the strategy to whose reliability a guiding generalization must
contribute must be a strategy for doing the right thing *for the right reasons*,
because even if doing what is right for non-moral reasons (say, to avoid
punishment) were a reliable way to act rightly, it wouldn't be a case of *moral*
guidance.[16] A generalization provides appropriate guidance to agents who
accept it as a guide only to the extent that their acceptance of it leads them
to perform right actions if they try to adhere to it.

Third, to assess whether a strategy provides adequate guidance we need
only to consider conscientious, morally committed agents. For unreliability
in the hands of lazy or careless agents indicates no fault *in the strategy*. But
the strategy must be *available* to practical thinking. If the only sort of fac-
tors that contribute to its reliability identified features of right actions which
not even a conscientious agent can access (or can access only in hindsight),
then the strategy would be useless in trying to decide what the right thing to
do is. But a guide that is available to one moral agent might be useless (say,
too complex or difficult to apply) to another, or it might be available only in
some but not all contexts to one and the same agent. So, adequate guidance
requires guides for particular types of agents in particular kinds of contexts.
A strategy *S* for acting well is unavailable to a given type of agent *A* if its
use by *A* requires information or inferences that are unobtainable or infea-
sible to *A*. The following *cognitive condition* explains why *S* cannot be
unmanageable to *A*: *S* is available to *A* only if the conditions for using *S* lie
within *A*'s cognitive ken. If moral principles are to provide adequate guidance
for normal humans, they must contribute to reliable strategies for acting
well which agents with limited cognitive capacities and resources can use.

Fourth, this cognitive condition, and more generally the idea that adequate guidance requires guides for particular types of agents in particular kinds of contexts, might lead us to think that generalist models of moral guidance would fail to apply to most agents, if moral principles could provide accurate standards of right and wrong only by being too complex or difficult to apply for most moral agents. But complexity alone doesn't doom principles to providing poor moral guidance. What adequate guidance requires is that moral principles contribute non-trivially to a reliable and available overall strategy for acting well. A principle may do so either *directly* by figuring in an agent's practical thinking or *indirectl*y by recommending modes of deliberation or helping to identify policies, heuristic guides, or simple rules that are sufficiently robustly reliable. For example, some Kantians think of the Categorical Imperative as a back-grounded regulator of maxims of action.

More familiarly, many utilitarians grant that the principle of utility is too difficult for most ordinary agents to apply directly, because the sorts of evidence and calculations which are required in order to determine what maximizes utility tend to be so complex as to violate the cognitive condition introduced above. But utilitarians often argue that the principle of utility provides indirect guidance by selecting for moral precepts which best enable ordinary agents to maximize utility, even as those precepts sometimes misfire by failing to pick out what is right from the utilitarian point of view.[17] Such precepts are, in effect, a series of heuristics that have some false implications concerning what really matters, namely utility. But they may nonetheless be adequate as guides for acting well.

Given this characterization of what adequate moral guidance is, PA in effect says that moral principles fail to contribute non-trivially to a reliable overall strategy for acting well that is available to the practical thinking of conscientious, morally committed agents. If this claim were true, then moral principles seemingly wouldn't provide adequate moral guidance.[18] What I'll call the Guidance Argument against PA and on behalf of generalist moral guidance can be set out as follows:

(G1)   Generalism can provide adequate moral guidance if moral principles can contribute non-trivially to a reliable and available overall strategy for acting well.

(G2)   For any conscientious and morally committed agent *A*, if *A*'s acceptance of true principles shapes *A*'s responsiveness to the right moral reasons, those principles can contribute non-trivially to some reliable strategy for acting well that is available to *A*.

(G3)   For any conscientious and morally committed agent *A*, *A*'s acceptance of (true) moral principles shapes *A*'s responsiveness to (the right) moral reasons.

(C1)   So, for any such agent, there is some reliable and available strategy for acting well such that the principles which the agent accepts can

contribute non-trivially to its reliability (where the strategy may be different for different agents). [G2, G3]

(C2)   So, generalism can provide adequate moral guidance. PA is false. [G1, C1]

The Guidance Argument is valid. Since (G1) is true by our hypothesis of what constitutes adequate guidance, I'll take (G1) as given. The action is in defending (G2) and (G3), then. I believe that these premises are defensible by many forms of generalism. My defense of (G2) will appeal to an account of moral principles as a kind of hedged principles that tolerate exceptions. But I begin with (G3).

4 Responsiveness to reasons and acceptance of principles

On any model of moral guidance, acting well involves some kind of responsiveness to the right- and wrong-making features, and hence some kind of responsiveness to moral reasons. I'll argue that acceptance of moral principles shapes conscientious, morally committed agents' responsiveness to moral reasons in ways which make moral principles suited to contribute to a reliable strategy for acting well. To defend (G3), we need to know what it is to be responsive to moral reasons and to accept moral principles.

Responsiveness to moral reasons is but a special case of responsiveness to reasons in general. In rough and generic terms, agents who are responsive to reasons (of a given kind) have a relatively stable tendency to form beliefs and intentions (of the relevant kind) that are more or less determinate functions of such characteristics as the contents of the relevant kind of inputs, such as beliefs, desires, intentions, and experiences and other 'seemings' of various sorts. When not mistaken about their reasons, they form beliefs and intentions on the basis of considerations that in fact are reasons. The natural explanation of what in one's psychological makeup underlies such a tendency is a complex disposition to not merely *conform to* but *be guided by* reason. The disposition in question is a disposition to respond, at least within a certain range of circumstances, with those actions and attitudes for which one has reasons (or, at least, reasons that pass some threshold of deliberative significance), and in a way that at least roughly reflects the relative strengths of those reasons. Being responsive to reasons requires such responses only within a certain range of circumstances, however, because we know that interfering factors may 'mask' the disposition (that is, prevent it from being manifested even under its 'triggering' conditions). Such factors include external manipulation of many forms, as well as various cognitive and temperamental factors, such as certain biases, fatigue, listlessness, and depression, which may contingently interfere with the operation of one's cognitive or practical capacities and abilities.

The disposition to be guided by reason is best understood *de re*, as a disposition to respond on the basis of those considerations that have the

property of being a reason, rather than *de dicto*, as a disposition to respond to whatever has the property of being a reason. For example, Olivia, who needs to do well in an exam, may respond to her circumstance simply by taking the necessary means to doing well out of a concern for doing well in the exam rather than a generic concern to follow reasons wherever they lead. In order for Olivia's responses to reflect the strength of some reason R to φ relative to other reasons, she may consider what she thinks of acting on R in view of the other things she is considering, how acting or not acting on R would bear on her other pursuits and to what it would commit her, and monitor her actions in the light of R and of factors and changes relevant to whether φ-ing makes sense in the light of R.[19] Thus Olivia, if reasons-responsive, may well skip a party the night before the exam while quite properly skipping her study and the exam to attend to a family emergency. But if called away on an emergency, she will see the need to notify the professor so as to arrange to take the exam at a later date. Again, interfering factors may mask the disposition. Even if Olivia recognizes her reasons, she may still go to the party if she is weak-willed, or fail to notify the professor if she is distracted or depressed because of the family emergency.

It is important to note that responsiveness to reasons comes in varying degrees of the tightness of fit between recognizing the reasons there are and their relative strengths ('receptivity' to reasons) and translating those reasons into decisions and translating these decisions into behavior ('reactivity' to reasons).[20] For example, an agent may exhibit an intelligible pattern of actual and counterfactual recognition of reasons, but frequently be weak-willed or otherwise unsuccessful in translating their recognition of reasons into decisions or subsequent behavior.

This is important to note, because acceptance of moral principles also comes in varying degrees of strength. Genuinely to accept a set of moral principles is, generically, to have dispositions to respond in certain ways to certain sorts of circumstances and actions in the light of certain features they have, which dispositions underlie one's responses. But what kinds of dispositions, and to what kinds of responses, does one have in virtue of accepting a set of principles?

We can distinguish three types of view on the basis of the three elements of responsiveness to reasons. According to a minimal view, the relevant dispositions consist just in dispositions to form certain moral beliefs. If I accept that lying is *pro tanto* wrong, I will be disposed to form certain beliefs about the morality of actions that I take to constitute lying. These beliefs will be sensitive to my understanding of the context. I will be disposed to regard an act's constituting lying as a weaker reason against doing it when by lying I can protect an innocent person from a murderer, and perhaps to regard an act's constituting lying as no reason at all against doing it when engaged in a game of bluff. The other two views build also certain reactive dispositions into acceptance of moral principles. According to the weaker view, the dispositions that at least morally committed agents

have in virtue of accepting a set of moral principles include dispositions to translate one's moral beliefs into decisions to act in accordance with the principles one accepts. If I accept that lying is *pro tanto* wrong, I am disposed to decide not to perform certain actions so far as their constituting lying goes, and not just because lying would get me into trouble. According to the stronger view, the relevant dispositions also include various motivational, affective, and behavioral tendencies. If I accept that lying is *pro tanto* wrong, I am disposed not to lie even when this would be to my advantage, but to act in an honest and sincere fashion, feel guilt if I lie, and resent lying by others.

Each of these views corresponds to a certain conception of what responsiveness to moral reasons requires. Given that principles identify features that make actions right and wrong, they specify the form and nature of conditions that would license one to treat some considerations as moral reasons. Thus the way in which at least conscientious and morally committed agents are responsive to moral reasons is a function of their receptivity and reactivity to those conditions as grounds of moral reasons. If so, then on the minimal view of what it is to accept a moral principle, such acceptance involves having a disposition to recognize moral reasons, at least within a certain range of circumstances. On the two stronger views, it also involves certain reactive dispositions to translate recognition of reasons into decisions or subsequent behavior (again, within a certain range).

A natural way to describe the way in which the acceptance of a set of principles on *any* of these views makes a difference to the dispositions of conscientious and morally committed agents, is to say that it structures the way in which they are responsive to moral reasons by giving a certain shape to certain aspects of their moral conscience. And were such agents to accept true principles, this would presumably shape the relevant aspects of their moral conscience so that they would be responsive to the right moral reasons. Hence, as long as we read 'acceptance of principles' no more weakly than 'responsiveness to reasons', premise (G3) of the Guidance Argument is very plausible.

One complication I should address is that each of the above views of acceptance of moral principles raises questions about the role that moral principles must play in the practical thinking of agents who accept them if they are to function as guides. According to one view, due to Holly Smith, one uses a moral principle as a guide for making a decision on a particular occasion, just in case one decides to perform an act out of a desire to conform to the principle and a belief that the act conforms, which in turn requires that one 'explicitly represent the principle as the content of a propositional attitude' at that occasion (H.M. Smith 1988: 90–92). I find this view objectionable. What, intuitively, should motivate the agent in a case of adequate guidance by moral principles is a concern for the considerations that the principles identify as moral reasons (a concern that has a certain sort of actual and counterfactual shape). In requiring one instead to make

one's decisions out of a desire to conform to the principle, Smith makes conformity to the principle seem a kind of fetish.[21] The view also seems too strong. Consider, for analogy, that minimal practical rationality plausibly requires conformity to some principle of instrumental rationality. According to Smith's account, one uses that principle as a guide just in case one engages in a type of instrumental reasoning: one uses it just in case one decides to perform an act out of a desire to conform to the principle and a belief that the act conforms. We cannot, however, require that one use the instrumental principle in this way, since one is capable of so using it only if one *already is* minimally practically rational (cf. Dreier 1997: 93–94). But now it is hard to see why using a moral principle as a guide should require that one decide to perform an act out of a desire to conform to the principle and a belief that the act conforms, when using the instrumental principle as a guide seemingly needn't involve any such cognitive structure.[22]

A weaker view that I find more plausible says that using a principle as a guide on a particular occasion requires only that one have it to some extent available for explicit reasoning, not that one explicitly represent it as the content of a propositional attitude at that time. For moral principles may guide responses to particulars even if known only tacitly by those who so deploy them (O'Neill 1996: 86–87; Garfield 2000: 191). So, one may be using a principle as a guide on a particular occasion even if its guiding role on that occasion is tacit or unreflective. Regarding (G3), this view implies that if acceptance of moral principles shapes our responsiveness to moral reasons, then our actual and counterfactual responses can indicate an understanding of principles even if we cannot fully articulate what underlies our responses.[23] Then using a principle as a guide requires only that we be responsive to the features the principle identifies as reasons for certain actions to be right or wrong, and that the principle be to some extent available for explicit reasoning, not that we also explicitly represent the principle as the content of a propositional attitude.[24] A view of this kind allows that acceptance of moral principles may shape our responsiveness to moral reasons in ways that leave room for us to refine our understanding of these principles. It also helps us not to conflate articulation and under-standing. For our understanding, moral or otherwise, isn't exhausted by what we can explicitly articulate (see e.g. Churchland 1996; Raz 1999; Wright 1999). Our reasoning also often reasonably relies on background assumptions that we leave implicit or take for granted (see e.g. Bach 1984).

A further complication I need to address is that any account of moral guidance (generalist or not) will ultimately need to forge a connection between acceptance of moral principles (or other guides) and motivation, since agents who accept some principles will thereby reliably *act* well only to the extent that their acceptance of those principles is sufficiently moti-vating. The strongest type of view of what it is to accept a moral principle forges such a connection by building motivational dispositions into princi-ple acceptance. But (G3) itself doesn't presuppose such a view of the

motivational powers of principle acceptance. Whether such a view is required depends on the motivational demands of morality, which I won't try to descry here. For my purposes, we may work with any interpretation of (G3) on which it is coherent to grant that there are true moral principles but hold that one may be appropriately responsive to the moral reasons which those principles identify without accepting those principles.[25] For then (G3) won't settle the substantive issue whether there are adequate particularist models of moral guidance by conceptual fiat.

5 A model of hedged moral principles

According to premise (G2) of the Guidance Argument, if a conscientious and morally committed agent's acceptance of true principles shapes the agent's responsiveness to the right moral reasons, then those principles can contribute non-trivially to a reliable strategy for acting well that is available to the agent. My defense of (G2) begins with the worry that *any* principle might be such that accepting it *mis*shapes one's responsiveness to moral reasons and so fails to contribute to one's reliability at recognizing (and reacting to) them. No doubt *some* principles are like this, and even conscientious agents might accept principles that misshape their responsiveness to reasons.[26] The worry is whether there might be *no* kind of moral principles whose acceptance can contribute to conscientious and morally committed agents' reliability at detecting the presence of moral reasons.

Particularists often present reasons holism as raising this worry. In §2 I said that holism is available also to generalists because there can be certain kinds of true moral principles even if holism is true. But even if holism leaves room for principles that accurately capture the behavior of moral reasons, its truth would require us to say more about what such principles must be like if they are to contribute to a reliable strategy for acting well. If, as holism holds, there is no necessary connection between the property of being a reason and the property of always being the same kind of reason, then anything that provides a reason for something is, *qua* a reason, a variable reason: in a different context it may be no reason at all, or even an opposite reason. Thus, for any wrong-making feature *F* considered *qua* a reason, there may be contextual conditions which would be 'unsuitable' for something's being *F* to be any reason at all for its being wrong (in which case it is morally neutral, unless further conditions obtain that are suitable for its being *F* to count in favor of its being right). The presence of unsuitable conditions of this kind amounts to the presence of defeaters for *F*'s having any wrong-making force at all, or *undermining defeaters*. An example of an undermining defeater would be that an act's constituting lying doesn't make it at all wrong when playing a game where lying is the point of the contest, or when an honest background agreement to deceive one another is in place.[27] Undermining defeaters are distinct from *overriding defeaters*, whose presence is unsuitable only for something's being *F* to make it wrong

overall. An example of an overriding defeater would be a case where an action is worse for involving a lie but not overall wrong because lying would save an innocent person from Nazi guards at the door. Since an overriding defeater allows that saving the person is nonetheless to some extent wrong in virtue of requiring lying, only capturing the kind of contextual variability of reasons (if any) which is due to undermining defeaters requires holism.

In view of the above, establishing (G2) in a way that accommodates holism seems to require two things. The first is to defend a kind of moral principles that purport to capture the way in which moral reasons behave if holism is true. The second is to show how the acceptance of such principles can contribute to one's reliability at detecting both moral reasons and undermining defeaters. If one is reliable at detecting undermining defeaters, but not because of any contribution from the moral principles one accepts, then the principles wouldn't seem to contribute to one's reliability at recognizing the presence of moral reasons either. For if one is unreliable at detecting undermining defeaters, then one is unreliable at recognizing the presence of moral reasons, and *vice versa*. Thus, if the principles that one accepts didn't contribute to one's reliability at detecting undermining defeaters, then they wouldn't seem to contribute to a reliable a strategy for acting well either.[28]

To build towards the kind of moral principles that achieve these two desiderata, let's note that if holism is true, then cases in which causing pain or breaking a promise isn't at all wrong might seem to make ordinary moral precepts like 'It is wrong to cause pain' and 'One ought to keep one's promises' prone to error. The ways in which many conscientious agents typically rely on such precepts seem, however, to indicate that they regard those precepts as expressing principles that tolerate exceptions and so are structurally more complex than their surface form lets on. For example, many of us judge 'Ravens are black' as true in spite of knowing that there are albino ravens. Just so, many of us would assent to 'It is wrong to cause pain' or 'Pain is bad' in spite of regarding the conditions as unsuitable for pain to have any wrong- or bad-making force at all when causing it is part of a medical procedure that is necessary for saving the patient's life or when it is constitutive of athletic challenge. While the latter case involves consent to pain, we also recognize that consent may not always be an unsuitable condition for pain to be wrong-making. Conditions may well be suitable for pain to be wrong-making when consent is due to manipulation, brainwashing, external conditions that generate adaptive preferences, or the like.[29] We also know of views on punishment, as well as theological views, according to which causing pain may well be right-making when the pain is deserved. Yet, if holism is true, none of this shows that the mere fact that an act would cause pain isn't capable of functioning, in some contexts, as a reason for the act to be wrong.

It seems integral to justification in ethics that when we judge some specific consideration *C* to be (that is, to have the property of being) a reason

86   *Pekka Väyrynen*

for an act's being right or wrong, we should be able to explain why *C* is the kind of reason it is.[30] We routinely accept this explanatory demand both in the case of theoretical reasons for beliefs about non-moral matters and in the case of many non-moral practical reasons, such as mundane instrumental reasons for action. To the extent that non-moral and moral reasons have a unified nature *qua* reasons, that demand should be legitimate in the case of moral reasons as well. It seems especially legitimate in the case of variable reasons. If something is a reason of a certain type in certain situations but not others, surely there should be an explanation of why it is a reason when it is and why it isn't a reason when it isn't. Explaining why a variable reason, when a reason at all, is a reason of a certain type involves explaining why certain conditions but not others count as unsuitable conditions for it to be the type of reason in question.

We can meet this explanatory demand in a deep way that serves moral theorists' aspiration to explain what makes actions right and wrong by specifying what I'll call the *normative basis* of a moral reason. By this I mean some evaluative or deontic condition (property, relation, etc.) the presence of which explains why *C* is a reason (when it is) for an action to be right or wrong. What exactly we regard as the basis of any given moral reason will depend on our substantive moral theory. In the case of pain, familiar proposals from normative ethics include the ideas that if causing pain is wrong-making, this is when and because causing pain *produces something intrinsically bad* or *makes the victim worse off*, when and because it *fails to exhibit the kind of concern or respect which the victim merits*, or when and because the fact that the act causes pain is among the reasons why some more general requirement not to cause pain *cannot be reasonably rejected*. These proposals agree that when causing pain is wrong-making, that moral fact has an explanation in terms of a normative basis, and only disagree on what that basis is. Many explanations of right-making features have a similar relational structure: features are right-making when and because acts having them stand in some such relation as *promoting*, *protecting*, *honoring*, or *respecting* to a positive evaluative or deontic property. This kind of explanations of moral reasons extend to variable reasons and unsuitable conditions associated with them. For example, one might think that pain isn't wrong-making when it is constitutive of athletic challenge, because under that condition pain doesn't make one worse off, or because causing it to one is compatible with one's exhibiting the kind of concern or respect which persons merit. Each idea appeals to one and the same normative basis in its explanation of why pain is wrong-making, when it is, and why some such fact as that pain is constitutive of athletic challenge is an undermining defeater for the fact that the act causes pain to make any contribution to its wrongness.

This shared structure of reasons and their defeaters seems to be reflected in the way in which many agents' judgments about moral reasons are sensitive, to varying degrees, to changes in the features of situations. The way

they treat a principle like 'Causing pain is wrong' reflects a recognition that causing pain isn't wrong-making under certain conditions. But such a recognition doesn't usually lead mature agents to abandon the principle. This suggests that they treat the principle as having a more complex structure than the verbal formulation 'Causing pain is wrong' lets on. If the structure is such as to accommodate reasons holism, and if holism is true, then any apparently simple principles that have such a structure will be less prone to error than they would be if they lacked such a structure.

One way to make moral principles capture this complexity of structure in the behavior of moral reasons, along with the idea that moral reasons have normative bases, is to build a reference to their bases into moral principles. To do this, we need a way of describing normative bases which is suitable for expressing generalizations whose truth tolerates exceptions but remains neutral on the morally substantive question of what these normative bases are. Let *the designated relation* for a property F and a moral property M (such as being right or being wrong) be that relation R, whatever it is, such that x's being F is a reason for x's being M when, and because, x instantiates R. For a given choice of 'F' and 'M,' if R is the designated relation for F and M, then what makes conditions suitable for the fact that x is F to contribute to (or, to be a reason for) its being M is that x instantiates R. Intuitively, R is a relation like promoting or respecting something of moral significance which can explain the reasons provided by F. (Sometimes R may be a monadic relational property.) The designated relation for causing pain (F) and wrongness (M) might be something like *failing to exhibit the kind of concern or respect which persons merit*.

Just *which* relation the designated relation is for any given moral reason is a substantive moral question. But the following kind of principle about causing pain is neutral on that question:

(P)   Any act of causing pain is *pro tanto* wrong in virtue of its causing pain, provided that the act instantiates the designated relation for *causing pain* and *being pro tanto wrong*.

(P) is an instance of a kind of principle that is hedged by reference to the designated relation. For any choice of 'F' and 'M', we can speak of 'the designated relation for F and M'. Thus a hedged moral principle of the form (HP) always seems available for consideration (be it true or false):

(HP)   For any x, if x is F, then x is M in virtue of being F, provided that x instantiates the designated relation for F and M.

(HP) gives a general model of hedged moral principles that purport to identify moral reasons. The crucial explanatory gain of the model is that its appeal to the designated relations helps us answer the question *why* the circumstances that count as reasons and unsuitable conditions are morally

relevant in the ways they are. The model captures generalism: any moral reason's having a normative basis requires the existence of a true moral principle, because according to this model the normative basis of any given moral reason is the relevant designated relation, (the existence of) which entails a (true) principle of the form (HP). The model also explains holism: a feature that a hedged principle identifies as giving moral reasons of a certain type fails to give a reason of that type if conditions are unsuitable for it to do so. And conditions are unsuitable in this respect when, and because, the relevant designated relation fails to be instantiated. (More precisely, the model explains why holism, if true, is true. It doesn't entail holism: principles of the form (HP) allow for the possibility that the relevant designated relation must be instantiated whenever the relevant F is, so that conditions are never unsuitable.)

Although I hope to have conveyed some intuitive and theoretical merits of this model of moral principles, my aim here isn't to show that it has such merits or argue that it supports generalism. My aim is to show how the acceptance of hedged principles can contribute non-trivially to a reliable strategy for acting well that is relevantly available to the agent, even if holism is true.

6 Hedged principles and adequate moral guidance

Hedged moral principles make good sense of our idea of morally committed persons as ones who take certain moral ideals to be centrally relevant to determining what they should do and why, and so would prefer to be guided by them. Hedged principles also make good sense of the idea that an agent who accepts the principle 'Causing pain is wrong' but causes pain to someone for no good moral reason not only commits a particular wrong but also violates an ideal, such as respect for other people or a concern for their well-being, which the agent endorses. Adhering to the principle on that occasion would symbolize the agent's commitment to upholding the ideal in all instances for which the principle stands. Hedged principles make such sense because the designated relations to which they refer involve moral ideals, in which case accepting such a principle commits one to caring about some moral ideal.

For example, persons who accept 'Causing pain is wrong' may do so out of a concern for justifying their actions to others, or respecting them, or not making them worse off, which would involve accepting (P) on the basis of one or another particular substantive view about what the relevant designated relation is. Even if one accepts 'Causing pain is wrong' out of a direct concern for not causing pain to others, or even comes to accept it in one's moral education as an initially simple precept, one would presumably be a defective moral agent if one accepted the principle even if one thought that there was no basis for regarding the property of causing pain as a wrong-making feature. So its acceptance by a conscientious agent who

thinks there is some such basis would seem to indicate a commitment to something like (P).[31]

If the designated relations to which hedged principles refer explain which features are right- and wrong-making and which conditions are unsuitable for them to be so, and if the acceptance of such principles requires some grasp of those relations, then their acceptance shapes one's responsiveness to moral reasons. To illustrate this defense of premise (G2) of the Guidance Argument, we can model the acceptance of hedged principles roughly as involving a commitment to a counterfactual condition.[32] For example, according to a fairly minimal conception of what it is to accept a moral principle, the condition would be something like this: 'I wouldn't take x's being F as a reason for x's being M if x didn't instantiate the designated relation for F and M'.[33] An agent who reliably meets this condition is reliable at detecting the presence of moral reasons and unsuitable conditions. But how exactly can acceptance of hedged principles contribute non-trivially to such reliability?

We can approach this question by asking how best to describe the content of the ability reliably to detect the presence of unsuitable conditions.[34] One way to describe it is to say that it is an ability reliably to apply a list of potentially unsuitable conditions. This view is, however, plausible only to the extent that such a list can be stated in finite and manageable terms. But we cannot simply assume that this is always feasible, and I know of no convincing argument that anything in morality or its action-guiding function requires otherwise. Moreover, grasping such a list would seem to rely at least implicitly on some prior criterion of what makes some condition unsuitable in the first place. Acceptance of hedged principles can make a non-trivial contribution to one's reliability at detecting the presence of unsuitable conditions in part because they supply such a criterion. They imply that unsuitable conditions are those in virtue of which, for the given F and M, something that is F fails to instantiate the designated relation for F and M. (That is, the relevant designated relation provides a condition on which features of situations count as unsuitable for something's being F to contribute to its being M.) What is more, their acceptance can contribute to one's reliability even if we cannot list all potentially unsuitable conditions in finite and manageable terms, since grasping the relevant designated relation doesn't presuppose knowing *just which* specific conditions count as unsuitable. So it seems preferable to describe the content of the ability reliably to detect the presence of unsuitable conditions in terms of grasping the relevant designated relation.

Hedged principles should be able to contribute to a reliable strategy for acting well even if those who accept such principles lack a complete understanding of a full range of correct principles. Just how reliable the acceptance of hedged principles makes one at detecting the presence of moral reasons and unsuitable conditions, and more generally just what role hedged principles play in moral guidance, evidently depends on the degree to which

one grasps the relevant designated relation, which may be limited.[35] For we commonly begin with an inchoate sense of our moral ideals and values. I might accept 'Causing pain is wrong' while lacking a full grasp of the normative basis of a concern for not causing pain. I might at first have in mind only some of its implications or identifying characteristics, such as that, whatever it is, it has to do with well-being. Or I might be unsure whether pain is ever deserved or when it is deserved. Still, if my acceptance of a crude rule like 'Causing pain is wrong' involves the thought that pain's being wrong-making has got something to do with its making people worse off, then my judgments about the moral relevance of pain are guided by a pretty good proxy. I will be systematically picking up on the moral relevance of pain, even if I have in mind no particular list of unsuitable conditions. Thus even a partial grasp of the designated relations can instill modes of deliberation and policies for acting well that are reliable (at least within a certain range) but needn't explicitly feature principles of the form (HP) in their content.

We also seem to be able to refine our grasp of our moral ideals and values, and thereby refine our understanding of hedged principles, in light of what moral experience, thought experiments, and reasoning teach us about their implications. If I think that what makes pain wrong-making has got something to do with well-being, reflection on the effects of pain on the lives of those suffering from it, on desert, and so on, might lead me to think, for example, that well-being matters because of its role in enabling and sustaining autonomy and that causing pain is wrong-making when it undermines one's autonomy in ways one hasn't deserved. Or, for a different sort of example, suppose that consenting to pain isn't an unsuitable condition for causing pain to be wrong-making when it is due to adaptive preferences, and that this is because consent fails to be autonomous when due to such preferences. Then if I haven't realized that what sort of attitudes and behavior an appropriate sort of concern and respect for persons requires, depends on whether consent to pain is due to adaptive preferences, I will be less reliable at detecting the presence of moral reasons and unsuitable conditions as regards pain than if I would be if I had come to that realization.

Assuming, however, that practical wisdom and moral knowledge are possible in the first place, moral experience and inquiry can lead us to realizations of this kind. One way this can happen is via witnessing actual cases and reflecting on examples, such as how pain and adaptive preferences affect a person's life. As we know, these may often be particularly vivid and effective ways of testing and refining one's moral views. More generally, examples of unsuitable conditions can function as clues to what makes conditions unsuitable, and so can help to improve one's sense of what the relevant designated relation is or implies. Their consideration may often be epistemically significant.[36]

But the epistemic significance of examples of unsuitable conditions seems often to depend on auxiliary assumptions concerning the kinds of normative

relations to which hedged principles refer. For example, we seem unlikely to grasp the relevance of adaptive preferences to whether someone's consent to pain makes causing it permissible, unless we rely on some independent, fairly definite sense of why things like causing pain to someone and their consenting to some treatment themselves matter morally. It also seems possible for us to refine our sense of when causing pain is wrong-making by reflecting more abstractly on such elements of the relevant designated relation as, perhaps, the moral standing of persons or sentient creatures and the roles that pain and consent may play in their lives, and then relying on such reflections and our canons of moral and non-moral reasoning to determine whether some particular conditions are unsuitable to make causing pain wrong-making. In either case, we would rely on our initial grasp of our moral ideals so as to develop a more accurate understanding of how those ideals bear on our choices and thereby gain a better basis for judgments about what would satisfy them. Assuming that moral knowledge is possible, we can proceed in such a way that our judgments about moral reasons and unsuitable conditions draw on an increasingly refined moral understanding which more fully reflects the contents of the relevant hedged principles and reaches beyond the simple verbal formulations which we typically give to our principles.[37]

The foregoing suggests that acceptance of hedged moral principles can contribute to one's reliability at detecting the presence of moral reasons at minimum by providing a starting-point and direction for the sorts of moral inquiry that can improve one's reliability at detecting the presence of moral reasons and unsuitable conditions. Given a more complete grasp of the relevant designated relations, it can so contribute by providing an explicit basis for judgments about what considerations function as moral reasons and when conditions are unsuitable for them to do so. We can also draw out the point that a suitable *set* of hedged principles can contribute to one's reliability at discriminating among the alternatives that its members license in particular cases. Agents whose principles require them to donate to charity can see that they may not contribute by mugging the elderly and donating the proceeds. Even if they have moral reason to do something deceitful in a particular case, they can reliably identify as beneath consideration modes of deception that involve injury or self-abasement. If they ram into a garden gnome to avoid running over a child, they will know to offer compensation to its owner and to take their aesthetic judgment that the gnome was just hideous as no reason to withhold compensation.

The foregoing also suggests a generalist response to the argument that moral principles provide adequate guidance at most in a limited range of cases. The argument claims that the simple rules of ordinary morality, such as 'Knowingly causing human death is wrong' or principles condemning actions but permitting omissions, represent heuristics that are well suited to a certain range of problems, but that we shouldn't treat them as freestanding moral principles because they lead to systematic mistakes when

generalized outside that range of problems 'to situations in which their jus-
tifications no longer operate' (Sunstein 2005: 531).[38] Earlier we saw that the
way in which many ordinary moral agents allow many rules they accept to
have defeaters gives us some reason to think that they treat those rules as
having a more complicated structure than their surface form lets on. If we
tacitly understood such rules as having something like the form (HP), that
would entitle us to treat them as freestanding while making us cautious
about over-generalizing them. For even if the way in which we apply our
rules in earlier stages of moral development reflects the use of error-prone
heuristics, a fuller grasp of hedged principles can help us avoid errors
because their built-in justifications (that is, the designated relations) con-
strain their range of application. But, in part because grasping hedged
principles doesn't require grasping any exhaustive list of potential defeaters,
such principles needn't as a rule be so complex or fine-grained that we are
bound to commit more frequent or severe moral errors if we follow them
than if we less reflectively follow simple rules.

   Pulling the foregoing threads together helps us to see how hedged prin-
ciples can contribute to reliable and available strategies for acting well
because of the ways in which a conscientious, morally committed agent's
acceptance of them can shape the agent's responsiveness to moral reasons.
Their acceptance can contribute to such strategies either directly or indir-
ectly, depending in part on the degree to which one grasps the designated
relations and whether one grasps them as such or via their implications or
characterizing features. Either way, the degree to which one is reliable
thanks to one's acceptance of hedged principles can be robust in its scope.
Acceptance of (P), for example, could contribute to being reliable at
detecting the presence of moral reasons, to a degree that is roughly pro-
portional to one's grasp of the relevant designated relation, even if one
encountered a world where (P) is regularly defeated, such as a world of
exercise nuts or one of massive adaptive preference formation.[39] For if
causing pain regularly failed, for whatever contingent reason, to instantiate
the relevant designated relation, then (P) would counsel one to presume that
causing pain isn't wrong-making. Thus acceptance of (P), together with the
appreciation of relevant non-moral information, can instill the kind of sen-
sitivity to morally relevant contingencies which would (at least within a
certain range) reliably lead one to true judgments about the reason-giving
status of causing pain in such a world were one to inhabit it.[40] Furthermore,
hedged principles can contribute to reliable strategies for acting well that
are available for normal human agents with limited cognitive capacities and
resources. For what suffices to make those strategies available is satisfying
such minimal conditions of moral agency as sufficient intelligence and
maturity to develop moral ideals and deploy ordinary moral and non-moral
reasoning to determine what would promote, sustain, or honor them.

   As the last step in defending premise (G2) of the Guidance Argument, the
appeal to hedged principles helps generalists to show how the contribution

of moral principles to reliable strategies for acting well can be *non-trivial*. The issue of non-triviality remains because generalists grant that grasping an appropriate range of principles is neither necessary nor sufficient for having a reliable strategy for acting well. It isn't necessary because, as we have seen, particularists can coherently allow that there are true moral principles but insist that proper responsiveness to the moral reasons that those principles identify only requires responsiveness to moral reasons that needn't be grounded by acceptance of any principles as guides. It isn't sufficient because principles neither apply themselves nor carry rules for their application in their sleeves. Otherwise we would be off to an infinite regress, as each application rule would need a rule for its own application. So we require a sensitivity to particulars even to judge what moral reasons we have in a given case, let alone to judge what our moral duty is.[41] Moreover, even as the acceptable ways of performing our duty are constrained by other moral principles, principles often allow, and sometimes require, varied implementation in the world of varying cases.[42] So, recognizing whether our principles apply, what they require, and how to implement their requirements, calls for sensitivity to the details of the case at hand and judgment.

The above suggests that any reliable strategy for arriving at correct moral conclusions in new cases requires developing a sensitivity and skill of judgment which enable one reliably to judge particular cases aright. If so, and if any such strategy requires agents to do whatever in their power is necessary for using the strategy, then acceptance of moral principles entails a moral commitment to developing sensitivity and judgment. (These might be either parts of reliable strategies for acting well or skills required for a successful use of such strategies.) It is, therefore, false that generalism encourages us to make moral decisions without a sensitive and detailed examination of particular cases. What raises a problem for generalists here is, rather, that the role of sensitivity to particulars and judgment in guiding action seems to be such that possessing them requires something beyond grasping an appropriate range of moral principles (Blum 1994: 39). For if so, and if grasp of an appropriate range of principles is neither necessary nor sufficient for having a reliable strategy for acting well, then generalists need to explain why sensitivity and good judgment merely *supplement* rather than *supplant* principles, making them superfluous for guiding at least virtuous persons.[43]

Appeal to hedged moral principles helps us solve this problem when we consider it from an appropriate developmental perspective. Whatever it takes to get people off to a start with some grasp of moral principles or some degree of moral sensitivity, any degree of the former requires some degree of the latter. For even on a minimal conception, acceptance of moral principles involves some degree of responsiveness to moral reasons, and we need some degree of moral sensitivity to determine what moral reasons we have in particular cases. But, although none of us is born with a full grasp of a set of moral principles, no more is any of us born with a full-blown moral sensitivity. We must allow that each can be improved over time by

education and exercise, that each is nonetheless fallible, that some agents may achieve more development than others, and that even for the best of us such development is constrained by our limited cognitive capacities and resources.

The reason why generalists can deny that sensitivity and judgment supplant principles in guiding action is that grasp of hedged principles can help to improve and refine these skills. What we saw above about the epistemic significance of hedged principles suggests that further reflection on even partially grasped moral ideals that ground the principles that we accept can help bring novel relevant features or combinations of features into salience for us and help us see if they have some novel moral import. For example, reflection on how notions such as autonomy and consent interact can help us see that whether someone's consent to some course of treatment is due to adaptive preferences requires moral attention. Similarly, given even a partial grasp of an ideal like equal respect for persons, reflection on features of persons can help us realize that some of our views on sexual and racial matters, or the ways in which facts about the gender, ethnicity, or religion of the persons affected by our actions influence our decisions, are misguided. For example, someone in charge of personnel decisions who accepts the principle that personnel decisions should be made solely on the basis of professional qualifications might on such reflection come to realize that their decisions have been influenced by mistaken views or bias. Grasp of hedged principles can in this way help us refine our views of which actions count as disrespectful, improperly discriminative, injuring (and so on), and thereby help direct our attention more reliably to the right things.[44]

Since a more refined sensitivity may in turn help us gain a clearer understanding of the moral ideals underlying the principles we accept, improving our grasp of our principles and refining our sensitivities are ongoing processes of moral development which work in tandem and to which my generalist model of moral guidance generates a moral commitment. I suspect there is no level of sensitivity where hedged principles become hindering crutches that we had better discard if we are to improve our reliability at judging particular cases aright. The possibility of novel cases where our old ways of moral attention may not be reliable obtains even in the limiting case of virtuous persons.

7 Some comparisons with particularist guidance

I have defended premise (G2) of the Guidance Argument by arguing that hedged moral principles *can* contribute non-trivially to reliable and available strategies for acting well. But apart from pointing out that their past exercise of sensitivity and judgment may not adequately prepare even virtuous persons for novel circumstances, I have said very little about whether guidance by moral principles is *preferable* to particularist guidance. I would be amiss to neglect this issue, however. For even if moral principles can provide

adequate guidance, particularist guidance might (other things being at least roughly equal) be preferable if, for example, particularism provided strategies for acting well that are more reliable or better available to normal humans than the sorts of strategies to which principles can make a nontrivial contribution. This would still position particularists to argue that, all told, we ought not to rely on moral principles in our practical thinking. A full comparison of generalism and particularism on this score is unfortunately beyond the scope of this paper. For example, I cannot hope to address the complex issues whether principled guidance can counteract more effectively than particularist guidance the morally pernicious effects that such widespread biases as 'framing effects' and special pleading for one's own interests often have on our moral judgments.[45] I also lack the space to argue at any sufficient length that guidance by moral principles is preferable to particularist guidance. My aim is the modest one of identifying some important respects in which my generalist model of moral guidance is at least no worse off than extant particularist models.

In the previous section, I sketched an explanation of how a grasp of hedged moral principles can help us improve our reliability at acting well (including detecting moral reasons and unsuitable conditions, refining our moral sensitivity and judgment, and so on). In appealing to the kinds of dispositions and understanding that are involved even in a partial grasp of such principles, the explanation is a *material* one. Any adequate model of moral guidance should supply some material explanation of being reliable at acting well. For when one reliably acts well, something in one's psychological make-up must underlie and explain one's reliability, and understanding what that is would presumably enable at least some of us ordinary folks to know how to improve ourselves.

It is, however, difficult to find serious particularist attempts to give a material account of the capacities that make virtuous persons reliable at acting well. Dancy, for example, describes virtuous persons' capacities merely *formally* as whatever capacities make them consistently successful: 'To have the relevant sensitivities just is to be able to get things right case by case' (1993: 64). He has little to add by way of a material explanation: 'There is nothing that one brings to the new situation other than a contentless ability to discern what matters where it matters, an ability whose presence in us is explained by our having undergone a successful moral education' (1993: 50).[46] Dancy has equally little to say about how to develop such an ability. Sadly, 'for us it is probably too late. As Aristotle held, moral education is the key; for those who are past educating, there is no real remedy' (Dancy 1993: 64). The narrativist, and intentionally metaphorical, moral epistemology that Dancy sketches is equally silent on the material nature of the capacities that enable one to tell the right kind of story of the situation that captures its moral shape, and those capacities don't reduce to the ability to detect moral reasons and defeaters anyway because of general features of narratives (Dancy 1993: 112–14). At least for

now, then, describing the content of the ability reliably to detect moral rea-
sons and the presence of unsuitable conditions partly in terms of the kinds
of dispositions and understanding involved in the acceptance of hedged
principles gives a deeper and more informative material explanation of
reliability at detecting moral reasons and of how to improve one's reliability.

We might still wonder, though, whether generalist models of moral gui-
dance can be psychologically realistic in certain significant respects. As I
noted in §2, my defense of generalist moral guidance doesn't rely on the
implausible claim that the only way that people can make moral judgments
in the first place is by basing them on moral principles.[47] But we might
nevertheless wonder to what extent generalists can accommodate empirical
evidence from cognitive psychology about how people make moral judg-
ments in actual practice. Agents who take themselves regularly to be guided
by moral principles might be mistaken, in view of the empirical data that we
generally lack reliable access to the actual causes of our judgments in many
areas and our explanations of how we reach our judgments often are *post
hoc* constructions (Nisbett and Wilson 1977). Further empirical data seem
to suggest that moral judgments are typically caused by psychologically
immediate unreflective evaluations (often fueled by stereotypes or emotional
reactions), and that citing moral principles is usually a *post hoc* attempt to
rationalize, and justify to others, judgments made without reliance on
principles (Haidt 2001).[48]

These and other data allow that moral judgment is a kind of cognition
shaped by a moral sensibility, but many cognitive scientists suggest that a
normal human's capacity for moral perception, cognition, and judgment
should be understood specifically in terms of 'pattern recognition' skills or
in terms of neurally stored 'prototypes', 'schemas', or 'exemplars' of cate-
gories like morally significant vs. morally non-significant action and, within
the former, morally bad vs. morally good action, and, within the former,
such specific categories as killing, stealing, lying, and betraying. For some,
the lesson of these models of moral judgment is that neither 'moral experts'
nor even the merely morally competent normal humans are really guided by
principles.[49] Were this last inference sound, it might justify a preference for
particularist guidance.[50]

The inference is dubious, however. Even if our judgments and actions in
concrete situations are directly guided by prototypes or exemplars, the latter
can themselves be shaped by principles. In the terminology of the Guidance
Argument, we can say that even if we understand our sensitivity and capa-
city to respond to moral reasons in terms of prototypes or exemplars, our
sensitivity and responsiveness may themselves be shaped by acceptance of
moral principles in the ways I earlier argued they can be.

Consider someone who defends their judgment of an utterance of a fal-
sehood as wrong by saying that the utterance was a lie. If ordinary judg-
ments are driven by prototypes, schemas and exemplars, then the person's
judgment is a result of locating the utterance, on the basis of its perceived

features, in an abstract multidimensional feature-space stored in the neural network and determining that the utterance has enough of a range of features that the agent associates with a prototypical lie – that is, an abstraction understood as a region in the relevant feature-space which is generated from a range of examples of lies – or with a persona-schema[51] of a liar or a paradigmatic exemplar of a lie. In general, utterances are judged as lies or not on the basis of where they fall in this multidimensional space, and as morally wrong or not on the basis of where they fall relative to the region representing moral wrongness in the multidimensional space representing the category of morally significant actions. Neural networks trained up in this way prime us (as causal antecedents of experience or in some other way) to recognize a movement as aggressive or identify someone as a shifty lawyer or school yard bully on account of their perceived behavior and other contextual clues.[52] The similarity determinations involved in these judgments may be more or less reasoned or articulate, but the bases of such discriminations typically outstrip our capacities for verbal articulation. Any reasoning involved in reaching a judgment is likely to be reasoning by analogy with the relevant prototype, schema, or exemplar in one's neural database.

While the generalist hypothesis that acceptance of moral principles shapes our moral perceptions and responsiveness to moral reasons no doubt requires further empirical investigation, the above kind of picture of moral judgment seems not to undermine that hypothesis. Doubts about the role of principles in shaping, or forming the basis of, moral perception and judgment tend to concern only verbally articulated principles (see Churchland 1996: 106). But, as I argued in §4, principles can guide responses to particulars even if they are only tacitly known, and not explicitly represented as the contents of propositional attitudes, by those who so deploy them. Thus, a verbal articulation of a principle may merely reflect – and perhaps sometimes mark an improvement of – an antecedent grasp of a principle whose typical guiding role is tacit, unreflective, or habitual.[53]

More importantly, principles can guide judgment by shaping the very prototypes, schemas, or exemplars that drive moral judgment (if they do). The role of a prototype, schema, or exemplar is to prime the agent for recognizing various features that are taken to be typical (to various degrees) of members of the category which the prototype, schema, or exemplar represents. It serves this role by making the features in question salient among various cognitive inputs and by making the agent differentially alert to divergences from the constellations of features, by expecting which, the agent approaches new concrete situations. Well constructed moral prototypes that discount misleading or biased similarities can then dispose the agent (at least within a certain range) to respond to morally relevant features and ignore morally irrelevant ones, thereby grounding the capacity to respond to moral reasons. Acceptance of moral principles should then enable us to improve our prototypes by helping us to discount misleading

similarities, assign a higher relevance ranking to features that already are parts of our stereotypes, and extend our prototypes to include further relevant features.

One illustration of the point I am making would be someone whose prototype of a wrongful lie is that of uttering a false sentence that the speaker takes to be false. Acceptance of some principle that makes intentions matter to the moral status of actions might well lead one to extend one's prototype of a lie to include uttering a true sentence that the speaker takes to be false (even if the agent in the end classifies such an utterance less readily as a lie than an utterance of a falsehood the speaker takes to be false). It might also lead one to shape one's prototype of a wrongful lie to assign a higher relevance ranking regarding moral wrongness to whether the speaker intended to utter a falsehood than to whether the utterance was in fact true or false.[54]

Another illustration would be someone whose prototype of a wrongful lie doesn't make much of a moral distinction between lying in a game of bluff and lying in the context of a sales transaction. If one accepts a hedged principle concerning the wrongness of lying, reflection on the normative basis of the principle might well shape one's prototype of a wrongful lie by leading one to exclude lies made in a game of bluff, or more generally in contexts where an honest background agreement to deceive one another is in place, from one's stereotype of a wrongful lie. (Further illustrations can be generated from my discussion in §6 of how improving our grasp of the kinds of designated relations to which hedged principles refer can also help us refine our skills of sensitivity and good judgment.)

I have, of course, not shown that the above speculations on behalf of my generalist model of moral guidance are empirically adequate. My point here is that their plausibility doesn't seem to be undermined by descriptive theories according to which moral judgment is driven by prototypes, schemas, or exemplars, and often involves only *ex post facto* verbal expression of moral principles. To that extent, then, my account seems to fit with a realistic psychology of moral judgment. Notice also that in describing what modes of moral cognition are available to normal humans, descriptive theories of moral judgment constrain (to the extent of their accuracy) any model of adequate moral guidance. Hence the ability of generalist models of moral guidance to accommodate those theories and the empirical evidence supporting them strengthens the case that reliable principled strategies for acting well are available for use by normal humans with limited cognitive capacities and resources. Thus I conclude that evidence from cognitive science gives particularism no edge over generalism in making adequate moral guidance available to ordinary moral agents.

I'll finish by considering two problems Holly Smith raises for the ability of moral principles to serve as adequate guides: the *problem of error* and the *problem of doubt*. The problem of error arises when an act that one believes to be right isn't in fact prescribed by the principle one is using as a guide

(H.M. Smith 1988: 94). An example would be a juror who tries to follow a principle requiring adequate compensation for injured plaintiffs but believes, however reasonably but falsely, that granting a plaintiff $100,000 in damages would adequately compensate him when in fact $500,000 is needed. The problem of doubt arises when 'the decision-maker . . . lacks the empirical premise necessary to connect the principle to any act, and so cannot come to believe of any act that it is prescribed by the principle' (H.M. Smith 1988: 94). For example, the juror may simply not have enough information to determine what would be an adequate compensation. Like-wise, utilitarians might well find themselves in serious doubt about whether to vote for a flat-rate tax or a progressive tax, since they might well be unable to figure out which would maximize overall happiness. For in that case they wouldn't be able to figure out what utilitarianism actually requires them to do (H.M. Smith 1988: 95). My response to these two alleged pro-blems comes in two steps.

The first step is that the problems of error and doubt constitute no pro-blem for the ability of hedged principles as such to function as guides. The problem of error charges in effect that moral principles don't contribute to reliable strategies for acting well because our beliefs about facts that are relevant to the rightness and wrongness of actions are reliably erroneous. But recall from §3 that we can reasonably require only *conditional* reliability from strategies for acting well. If people as a matter of fact often have (reasonable but) false beliefs about what their principles require on a given occasion, this alone shows no fault in a principle that is deployed as a guide. The problem of error threatens generalist moral guidance only to the extent that the conditions for using the relevant strategies for acting well system-atically require beliefs that agents with limited cognitive capacities and resources are unlikely to get reliably right. This might be a sound worry about some particular moral principles.[55] But we can reliably find out whe-ther something causes pain to an agent, whether the pain is constitutive of athletic challenge, whether one consents to being caused pain, whether the consent is due to adaptive preferences, and so on. Thus the problem of error seems to pose no serious problem for using the hedged principle (P), intro-duced in §5, as a guide. And there seems to be no reason to think that (P) is alone among hedged principles in avoiding the problem of error. Similarly, the problem of doubt threatens generalist moral guidance only to the extent that the moral principles one accepts can only contribute to strategies for acting well the use of which requires the sorts of empirical premises of which creatures with limited cognitive capacities and resources are bound to be uncertain to a degree that rules out even partial belief.[56] This might again be a sound worry about some particular principles. But again the empirical premises required for using many hedged principles as guides seem accessible enough even to limited creatures like us.

The second step is that even if the problems of error and doubt did con-stitute a problem for generalists, they would constitute at least an equally

serious problem for particularists. Regarding the problem of error, particularist guidance has no special claim to making agents more reliable at acquiring the true (non-moral) beliefs about their circumstances on which the reliability of moral sensitivity is conditional. For even virtuous persons' cognitive capacities are limited and their sensitivity to the details of their circumstances fallible. So even if there is a problem of error for generalism, there seems to be an equal problem of error for particularism. Similarly, even if there is a problem of doubt for generalism, shouldn't there be equal doubt as to whether one is exercising one's sensitivity on the right things or whether the things that appear salient really should appear salient in one's circumstances? The answer seems to be yes. I conclude that generalists are at least no worse off than particularists regarding the alleged problems of error and doubt.

8 Conclusion

Particularists argue that we ought not to rely on moral principles in moral judgment because principles fail to provide adequate moral guidance. I have argued that this claim of principle abstinence is false. I have done so by defending a generalist model of moral guidance that appeals to a novel kind of hedged moral principles. Such principles can make a non-trivial contribution to reliable strategies for acting well which are available to the practical thinking of normal human agents, even in view of recent evidence from cognitive science concerning how we make moral judgments in actual practice. In particular, accepting and understanding hedged moral principles can make us more reliable at acting well by shaping our responsiveness to the right moral reasons. The conclusion I draw is twofold. First, generalist moral theories can provide adequate moral guidance. Second, they are at least no worse off than particularism regarding the provision of adequate moral guidance. As with sex education, so with moral principles: teaching abstinence isn't the best policy.

Notes

1 This paper derives in large part from my *Ruling Reasons: A Defense of Moral Generalism* (PhD diss., Cornell University, 2002). I am grateful to Terence Irwin for helpful guidance at that stage of the material's gestation. More recently, insightful written comments from Terence Cuneo and especially Sean McKeever helped me to improve the paper in many respects; my warmest thanks to them both.
2 See Väyrynen (2006a) for the argument and (unpublished) for the details of this model of principles.
3 I owe the helpful term 'principle abstinence' to McKeever and Ridge (2005a: 88).
4 Dancy says that a moral principle 'amounts to a reminder of the sort of importance that a property *can* have in suitable circumstances', so that a set of principles can serve as a kind of checklist for morally relevant features (1993: 67). Another writer who bases PA on reasons holism is David McNaughton (1988: 62, 190–93).

5 McNaughton (1988: 190–93), Dancy (1993: 60–62; 2000b: 130–37), and Little (2000: 281) infer particularism from holism.

6 The examples of utilitarian and Kantian principles that presuppose holism presented in McKeever and Ridge (2005b) inspire Dancy (2004: 81) to grant this point. Väyrynen (2006a) defends the point independently.

7 This is the form of particularism that Dancy defends in his most recent writings: 'Moral thought, moral judgement, and the possibility of moral distinctions – none of these depend in any way on the provision of a suitable supply of moral principles. This claim is what I call particularism' (2004: 5). My characterization in the text refers only to moral facts and distinctions, because generalism as I understand it implies no particular psychology of moral thought and judgment, but only that moral judgments depend for their truth on there being true moral principles.

8 Three points of clarification. First, I'll assume that, at least in typical cases, if something provides an explanatory reason why an action is wrong, then it also counts against doing that action in the sense of providing a normative reason not to do it. Second, the language of moral properties, facts, and truths may be construed as minimalistically or, if one prefers, rewritten in terms of moral predicates and sentences. Third, my definitions of generalism and particularism leave logical space for the hybrid view that some moral facts and distinctions depend on moral principles but others don't. But such a fragmented view of the moral domain seems unmotivated until some reason is given to take it seriously, and I'll ignore it henceforth. (A view that adopts principles of *pro tanto* moral reason but eschews principles of overall duty isn't the relevant sort of hybrid view. If moral reasons depend on principles, then so presumably will overall moral duties.)

9 See e.g. Hooker (2000b) and, esp., A.H. Goldman (2002) for discussions of different aspects of this point.

10 See e.g. Dancy (1993: 64) and McNaughton (1988: 190). See also Blum (1994: 39).

11 I owe the term 'reliable strategy for doing the right thing for the right reasons' to McKeever and Ridge (2005a: 86).

12 I provide a fuller presentation of much (but not all) of what follows in the next few paragraphs in Väyrynen (2006b).

13 More prosaically, the idea is that if you feed garbage in, then, no matter how well you process that information, you get garbage out. Compare Alvin Goldman's (1979) discussion of conditionally reliable belief-forming processes.

14 I draw the term 'robust reliability' from Henderson and Horgan (2001), who defend the epistemic significance of robustly reliable belief-forming processes. But see Heller (1995: 505–9) for a weaker robustness requirement.

15 I am indebted to Sean McKeever for this example and part of the point that it illustrates.

16 For another sort of example of non-moral guidance, consider a business executive who says 'I'm making this decision on principle, just to see how it feels' (a cartoon in *The New Yorker*, 10 November 2003, 87).

17 J.S. Mill, for example, argues that ordinary morality progressively captures the 'corollaries from the principle of utility' because these corollaries 'admit of indefinite improvement' and human beings 'have been learning by experience the tendencies of actions' (Mill 2000: 33). See also Sigwick (1907/1981: 199–216) and Hare (1981).

18 Generalism might then also imply that morally committed agents shouldn't try to adhere to the principles they accept.

19 One may be responsive to a reason R to φ without believing *of R* that it is a reason to φ. One needs only to respond in the light of R, in that one's beliefs and intentions reflect, in some not necessarily explicit fashion, R's (subjectively registered) support for φ-ing. Joseph Raz suggests that the 'reason-guided character'

of even actions that I perform without deliberation 'is manifest in the fact that I monitor them, and will abort them if the situation changes, or is revealed to have changed' (1999: 232). In general, competent agents needn't be articulate about the considerations to which they are responding when they are responding to reasons, even when these considerations are articulable. (On the venerable example of chicken-sexing, see Doyle 2000.) And habitual responses that don't arise from explicit reasoning may nonetheless be responses adopted in the light of one's appreciation of one's circumstances, and in that sense for reasons (Mac-Donald 1991: 39; Audi 1997: 147). Yet agents on whose thinking reasons impinge in these ways clearly can understand their import to their actions and attitudes and be guided by that understanding.

20 See e.g. the distinction between weak, moderate, and strong reasons-responsiveness in Fischer and Ravizza (1998).

21 I use the term 'fetish' in broadly the same sense as Michael Smith does when he argues that externalist accounts of moral motivation can only explain the reliable connection between making a moral judgment and being motivated to act accordingly, if they portray good people as being motivated to do what they judge to be right, where the content of this motivation is read *de dicto*, but that being so motivated is a fetish which externalism wrongly elevates into moral virtue (see M. Smith 1994: 74–75). One notable difference is that in Holly Smith's account of using a principle as a guide, the content of the fetishistic motivation is to be read *de re*, as the desire that one conform to a particular principle.

22 Holly Smith's rationale for her view is that moral philosophy is concerned with moral principles understood as objects of evaluation which are to be rejected or adopted in view of considerations for or against them, and that this process is most rational when those principles are entertained in propositional form (1988: 90). But even if Smith is right about what it is for *theory selection* in ethics to be rational, it doesn't follow that ordinary moral *reasoning* can be rational only if it involves selecting among moral theories, or that one's moral reasoning, understood as a kind of mental processing, should mirror the structure of the moral theory that justifies and explains the correctness of the outputs of such processing. Surely an agent can be instrumentally rational without selecting among theories of instrumental rationality.

23 Even so committed a generalist as R.M. Hare (1952: 64) stresses the point. A non-moral example is that one can drive without paying conscious attention and yet proceed on the basis of the law that forbids crossing an intersection against a red traffic light. In this respect, moral cognition may be roughly analogous to cognition in domains such as grammar, commonsense physical intuition, expert medical diagnosis, and musical composition. While cognition may differ in many ways between these domains (for example, some, like grammar, may involve generative systems while others don't), what they have in common is that our capacity to offer complex, subtle, and apparently systematic judgments about particular cases requires little conscious access to the mechanisms or principles underlying these judgments (see e.g. Stich 1993).

24 It isn't idiosyncratic to claim that one can use a principle in one's reasoning even without having the concepts required to formulate it. One can well use Leibniz's Law in one's reasoning so long as one has the concept of identity, even if one doesn't have the meta-logical concepts used in formulating the law, let alone explicit beliefs involving those concepts.

25 Here I grant, in effect, the coherence of denying the converse of (G3). I am indebted to Sean McKeever for pressing me to clarify my view on this matter.

26 For example, they might accept as guides principles that have false implications regarding what moral reasons there are, which affect their reliability at recognizing moral reasons even in the circumstances they are likely to encounter.

27 I owe this style of example to Lance and Little (2006b: 306, 314).

28 By the same token, a full defense of (G2) requires showing how acceptance of moral principles can contribute to one's reliability at judging when overriding defeaters are present. Such a case would need as its basis an account of how the right- and wrong-making features which moral principles identify combine to make actions right or wrong overall. Developing such an account is a task large enough to lie beyond the scope of this paper. What I can realistically hope to achieve is to defend (G2) against the particularist claim of principle abstinence. This aim allows me largely to bracket issues about detecting the presence of overriding defeaters, since particularists no less than generalists need an account of how the right- and wrong-making features of particular actions combine to make them right or wrong overall.

29 We may think of these conditions as 'defeaters for defeaters' or 'meta-defeaters': conditions that defeat the status of consent as a defeater for pain's being wrong-making. Thanks to Neil Levy for drawing my attention to adaptive preferences. Lance and Little (2006b: 319 n.1) attribute the athletic challenge example to Elijah Millgram.

30 In the next few pages I traverse rather quickly a territory that I cover with more argument and detail in Väyrynen (2006a, unpublished). Readers with concerns about the account as it is presented here are advised to consult those papers.

31 Familiar moral precepts are simple on their face at least in part because in moral education it is necessary to start by teaching learners *what* to do, primarily in the kinds of circumstances that they are likely to encounter, and leave it for them to find out later *why* (see Hare 1952: 67). The view I am developing explains this handily, since on this view we can come to understand why we ought to do certain things by grasping the relevant designated relations.

32 I say 'roughly' because we may be unable to describe the circumstances relevant to having a disposition, and therefore the dispositions involved in the acceptance of moral principles, simply in terms of counterfactuals. For example, the relevant counterfactuals might be true in virtue of 'finkish' dispositions, or false despite the presence of the relevant genuine dispositions because of the presence of 'antidotes' to them (see Bird 1998). For a defense of moral commitment as a commitment to a counterfactual condition which isn't sensitive to this point, see Stroud (2001).

33 Even if reliably detecting the presence of moral reasons requires reliably satisfying this condition, the present gloss on what it is to accept a moral principle doesn't beg the question against particularism, since it is possible reliably to satisfy the condition without being committed to it. Also notice that one's commitment may mostly operate as a constraint in the background of one's responses, so that one's responses can exhibit a counterfactual sensitivity to the satisfaction of the condition even if they aren't directly prompted by a desire to satisfy it (see Stroud 2001: 384–85).

34 See Väyrynen (unpublished) for a fuller discussion of this issue and the points in this paragraph.

35 No more than general features of reference are needed to secure the possibility that one may accept a principle of the form (HP) while having only a limited understanding of the designated relation to which it refers, or even while being mistaken about some of the relation's aspects. Although it seems clear that the acceptance of a principle requires some minimal threshold of responsiveness to moral reasons, I won't try to determine how unreliable one must be at detecting the presence of moral reasons and defeaters for us to have warrant for denying that one accepts a principle.

36 Terence Irwin makes the congenial point that generalizations that aren't fully spelled out, such as 'One ought to be helpful without being interfering' and those

that include qualifications like 'too much', are 'useful, but do not provide an effective procedure for just anyone, irrespective of their experience of such situations, to identify the actions that conform to them' (Irwin 2000: 128).

37 My account is fairly ecumenical on whether moral understanding can be improved through practical reasoning, rational intuition, or whatnot. This paragraph aims to put a bit more flesh around remarks such as Roger Crisp's claim that 'no rule can on its own, without some independent understanding of the nature of justice or courage, and some capacity to judge what they require in any particular case, satisfactorily guide action' (2000: 29). Scanlon (1998: 201), Irwin (2000: 120–21), and Nussbaum (2000: 238–39) express similar ideas. See also Väyrynen (unpublished).

38 Sunstein grants that simple rules operating as heuristics may provide the best available form of moral guidance despite the errors to which they lead (2005: 534–35, 541). On moral heuristics, see also McKeever and Ridge (2006: ch. 9).

39 Notice that particularism and generalism as such are silent on how frequently the right- and wrong-making features are defeated. Each allows that conditions might regularly be unsuitable in a given world, for the distribution of the relevant particular non-moral facts in a given world is a contingent matter (see Väyrynen 2004: 66–67).

40 The point would seem to hold *whether or not* such circumstances are ones that the agent, situated as the agent actually is, is likely to encounter. In this respect of robust reliability, hedged moral principles would seem to provide better guidance than the kinds of generalities regarding 'presumptive epistemic warrant' which Margaret Little takes moral principles to express. For she writes: 'The judgment that a given principle such as "lying is wrong" will help rather than mislead a moral novice reflects a judgment about the sorts of context she is *likely to encounter*' (2000: 295, emphasis added; cf. 294).

41 On this point, see not only Aristotle (1985: 1094b15–17, 1109b15), but also Kant (1797: Ak. 390). The representative notion of judgment understands it as 'the ability to evaluate a situation, assess evidence, and come to a reasonable decision without following rules' in a psychological sense of 'follow' (Brown 1988: 137–38). Judgment is widely thought to be indispensable also in scientific problem-solving and theory-selection (see e.g. Brown 1988: *passim*), and Aristotle's discussion of 'equity' appears to emphasize the need for judicial judgment in law (1984: 1374a30–b2; 1985: 1137b12–27). Notably, in these parallel cases principles seem to provide a fruitful framework or guidelines for judgment.

42 This twofold point is developed by Onora O'Neill, who gives 'Good teachers should set work that is adjusted to each child's level of ability' as an example of a principle that requires varied rather than uniform implementation, and notes that each principle 'helps to specify ways in which the others might be appropriately met' (O'Neill 1996: 75, 181).

43 Otherwise it might be that acceptance of principles is a useful ladder in teaching children to develop sensitivity and good judgment, but that for purposes of guidance these principles can be replaced by sensitivity and judgment once the latter are developed. See McNaughton (1988: 202). cf. McNaughton (1988: 62, 190) and Dancy (2005).

44 In this respect, my account of the kinds of guiding role that hedged principles can play is similar to Barbara Herman's account of 'rules of moral salience' (though the accounts differ in other respects). Herman understands such rules as defeasible and revisable interpretations of more fundamental moral conceptions (for Herman, the ideals associated with the Kantian Moral Law) which constitute the structure of one's moral sensitivity, thereby enabling one to recognize those elements of one's circumstances or proposed actions that require moral attention. See Herman (1993: 73–93).

45 In all honesty, I would also have little to add to Sean McKeever and Michael
   Ridge's extended defense of generalist moral guidance against the dangers of
   framing effects and special pleading (McKeever and Ridge 2006: 202–15).

46 Details of Dancy's view also give grounds for complaint. First, Dancy infers that
   the ability is contentless from the claim that the virtuous person is 'not conceived
   of as someone equipped with a full list of moral principles and an ability to
   correctly subsume the new case under the right one' (1993: 50). This inference is
   a *non sequitur:* its premise allows that virtuous persons might have a large, but
   less than full, set of principles (Irwin 2000: 102 n.5), and that their use of those
   principles might not be subsumptive and might require sensitivity and judgment.
   Second, in construing this ability as contentless, Dancy takes himself to be fol-
   lowing John McDowell's view that the virtuous person's conception of how to
   live isn't 'codifiable' (McDowell 1998: 57–58, 66–67). But McDowell denies that
   this ability is contentless: 'What it is for the practical intellect to be as it ought to
   be, and so equipped to get things right in its proper sphere, is a matter of its
   having a certain determinate *non-formal* shape' (1998: 184–85, emphasis added).

47 This claim is ably criticized, primarily on the basis of empirical evidence, by
   Dworkin (1995).

48 It is worth noting that these data are equally problematic (or not) for parti-
   cularists, since presumably they wouldn't recommend that we form unreflective
   snap judgments and then construct *post hoc* rationalizations to support them.
   Also, the claim that Haidt takes his empirical results to support is a descrip-
   tive claim that is consistent with the claim that we can and ought to make
   moral judgments more reflectively on the basis of reasoning or principles (Haidt
   2001: 815).

49 See e.g. Johnson (1993) and Churchland (1996); cf. A.I. Goldman (1993). For a
   dissenting voice concerning the role of principles in certain types of expert cog-
   nition, see the discussion of Kirsh and Maglio (1992) in Clark (1996: 118–20).

50 Dancy (1999a) appeals to connectionism and prototype theory to mount a par-
   allel argument for a particularist view of moral learning. I should also note that,
   with the exception of n. 53 below, the criticisms to follow of this kind of argu-
   ment were developed independently of similar criticisms which McKeever and
   Ridge (2006: 215–22) raise against rejections of generalist models of moral gui-
   dance that are based on evidence from cognitive science.

51 According to the schema theory, moral experience equips us with moral 'scripts'
   and 'personae' that enable us to detect morally relevant features (see Greco 2000:
   241). For example, we may judge that a person is untrustworthy if we 'see' the
   person as closely resembling the persona of a shifty lawyer, or as more like a
   shifty lawyer than Erin Brockovich.

52 It would make no difference to my present aims if we adopted a view of moral
   judgment as akin to pattern recognition that doesn't rely on classification in
   terms of prototypes, but on a direct detection that a certain pattern is present,
   based on features that prompt a recognition of similarity with patterns stored in
   memory (see e.g. Dworkin 1995: 234).

53 Moreover, as McKeever and Ridge (2006: 221) point out, even if Haidt (2001) is
   right that we often verbally express principles only *ex post facto* when we want to
   justify our judgments or actions to ourselves or others, we may thereby, if we are
   sincere, come to accept principles which then come to shape our perceptions and
   acquire a guiding role.

54 Whether this shaping would constitute an *improvement* in the moral prototype is,
   of course, a substantive question, but one that is tangential to my illustrative
   purpose here.

55 For example, a common objection to consequentialist principles is that the diffi-
   culties with acquiring and processing the relevant information about the con-

sequences of our actions raise just this problem. The equally common reply is that familiar indirection moves solve the problem. Which side is right is a complex question on which I take no stand here.

56 Smith fails to note the possibility of a partial degree of belief concerning the relevant empirical premises. See Lockhart (2000) for discussion of the much neglected issue of moral decision-making under uncertainty.