

## Using Moral Principles to Guide Decisions

Forthcoming in *Philosophical Issues*, Vol. 22

A long line of moral philosophers have argued that one, or the principal, role of a moral theory is to provide guidance to agents in deciding what to do. Given this, they have rejected as inadequate moral theories that cannot be used for guidance, however attractive they may be as theoretical accounts of what makes acts right or wrong. Thus, in a classic statement, James Hudson rejects objective utilitarianism, which prescribes maximizing utility, on the grounds that “for human agents the theory is not really “action-guiding”: it *does* say what one should do, but it gives this information in a practically unusable way,”<sup>1</sup> because no one can be confident which action would maximize utility over the very long run.

Utilitarianism is notoriously subject to problems of applicability because of agents’ difficulty in ascertaining which action would maximize utility. However, deontological theories are subject to similar limitations on agents’ knowledge. An agent may be uncertain or mistaken about, or unable to ascertain, whether the statement he is considering making is untrue and so a

---

<sup>1</sup> James Hudson, “Subjectivization in Ethics,” *American Philosophical Quarterly*, Vol. 26, No. 3 (July 1989): 221. Concern about whether agents could use utilitarianism to make decisions is one of the objections to which J.S. Mill addressed himself in *Utilitarianism* (Indianapolis: The Bobbs-Merrill Company, Inc., 1957 (originally published in 1861)), 30-31. For a selection of other examples of philosophers who require moral theories to be usable for guidance, see H. A. Prichard, “Duty and Ignorance of Fact,” in H.A. Prichard, *Moral Obligation and Duty and Interest* (Oxford: Oxford University Press, 1968), 18 – 39 (essay originally published 1932); W.D. Ross, *Foundations of Ethics* (Oxford: Clarendon Press, 1939), Chapter VII; Bart Gruzalski, “Foreseeable Consequence Utilitarianism,” *Australasian Journal of Philosophy* 59 (1981), 163-76; Frank Jackson, “Decision-Theoretic Consequentialism and the Nearest and Dearest Objection,” *Ethics*, Vol. 101, No. 3 (April, 1991), 461-482; Frances Howard-Snyder, “Rejection of Objective Consequentialism,” *Utilitas*, Vol. 9 (1997), 241-248; Mark Timmons, *Conduct and Character*, 6<sup>th</sup> ed. (Wadsworth, Cengage Learning, 2012), 10-11; Elinor Mason, “Consequentialism and the ‘Ought Implies Can’ Principle,” *American Philosophical Quarterly*, Vol. 40, No. 4 (October 2003), 319-331; James Griffin, *Value Judgment* (Oxford: Clarendon Press, 1996), Chapter VII; and R. Jay Wallace, “‘Ought,’ Reasons, and Vice: A Comment on Judith Jarvis Thomson’s *Normativity*,” *Philosophical Studies* (2011) 154: 460.

lie, whether the woman with whom he is considering have intercourse is over eighteen, whether the defendant whom he is about to convict is really guilty, whether the jacket he is about to don in a restaurant is genuinely his, and so forth. Thus it appears that a philosopher who holds that moral theories must be usable for guidance should reject standard deontological views as well as utilitarianism.

Let us call the requirement that an acceptable moral principle must be usable for guiding decisions the Applicability Demand.<sup>2</sup> To understand the Applicability Demand, and to assess whether or not to endorse it, one needs an account of what it is for a theory to be action-guiding in a practically usable way. Little work has been done to articulate the relevant notion of usability. This paper provides such an account. In the course of developing this account, I will argue that we need to distinguish between a *narrow* and a *broad* sense of having the ability to use a principle to guide conduct. Philosophers advocating the importance of a principle's usability must choose which of these two senses articulates the appropriate criterion for moral principles. I will also argue that my account of usability provides a solution to Eugene Bales' insight that there is a sense in which *every* moral principle is usable, however impoverished agents' knowledge of non-normative facts is. Bales notes that although an agent may (for example) be completely ignorant what the utility of her various choices would be, nonetheless she can still derive a prescription from act utilitarianism, namely the prescription "Perform the act that would maximize utility."<sup>3</sup> Clearly such an agent should not count as having the ability to use the utilitarian principle, but we need to be able to say why not. The account I propose will

---

<sup>2</sup> See Mark Timmons, *Conduct and Character*, 10-11.

<sup>3</sup> Eugene Bales, "Act-Utilitarianism: Account of Right-Making Characteristics or Decision-Making Procedure?" *American Philosophical Quarterly*, Vol. VII (July 1971).

contain the tools we need to do this. Actually assessing the Applicability Demand is the project for another occasion.<sup>4</sup>

## 1. Guiding one's conduct by a moral principle

We may start with the question of what is involved in guiding one's behavior by reference to a moral principle.<sup>5</sup> As Kant pointed out long ago, simply behaving in conformity with a principle does not guarantee that you have used it for guidance.<sup>6</sup> You order fish rather than pork at a restaurant, thus conforming to the Levitical law prohibiting consumption of any animal that parts the hoof but fails to chew the cud. But if you are wholly unaware of this law, you have hardly used it for guidance. Nor does knowledge of what the law requires, together with conformity of behavior, show that you have guided your action by it. You might be aware of the dietary law but completely unmoved by it; your only concern in ordering your dinner is to avoid high cholesterol foods. The concurrence between the Biblical command and the avoidance

---

<sup>4</sup> See Holly M. Smith, *Making Morality Work* (in manuscript).

<sup>5</sup> There is a massive contemporary literature on rule-following triggered by Saul Kripke's discussion of Wittgenstein's puzzles about rule-following (Saul Kripke, *Wittgenstein on Rules and Private Language* (Oxford: Blackwell, 1982)). This literature focuses primarily on what *counts* as following one rule rather than another, where an adequate account must (a) account for the possibility of objective error in following the rule, (b) account for the infinite character of rules, and in particular the fact that a rule may theoretically require action under circumstances in which no human agent could ever find herself, and (c) account for the fact that an agent following a rule can justify her action by referring to the rule. (See Frank A. Hindriks, "A Modest Solution to the Problem of Rule-Following," *Philosophical Studies* (2004) 121: 67.) Although these issues are related to those in the present paper, they are largely orthogonal to our concerns here, so I shall not try to address them. Some scholars believe that no successful response has yet been offered to the skeptical challenges about rule-following that Kripke raises in *Wittgenstein on Rules and Private Language*. If and when such a response is found, the account offered in this paper may be translated appropriately to fall under that response. For recent contributions to this debate, see Hannah Ginsborg, "Primitive Normativity and Skepticism about Rules," *The Journal of Philosophy*, Vol. CVIII, No. 5 (May 2011), 227 – 254; and Markus E. Schlosser, "The metaphysics of rule-following," *Philosophical Studies*(2011): 155:345-369.

<sup>6</sup> Immanuel Kant, *Foundations of the Metaphysics of Morals* (Indianapolis: The Bobbs-Merrill Company, Inc., 1959), 13 – 16.

of cholesterol is purely coincidental. These considerations suggest that we can say (roughly) that an agent uses a principle as a guide for making a decision just in case the agent chooses an act out of a desire to conform to the principle, and a belief that the act does conform.<sup>7</sup> Thus, suppose Susan decides to signal a lane change because she desires to follow the highway code, and believes that the highway code requires lane changes to be signaled. She signals her intended lane change. She has used the code to make her decision. We may say roughly, then, that a principle is *usable* by an agent for making a decision just in case it is true that if the agent wanted all-things-considered to act in conformity with the principle, the agent would choose an act out of her desire to conform to the principle and a belief that the act does conform.

We need to make this notion of a principle's being usable by an agent more precise. To see what may be needed, let us first examine the barriers that might prevent a person from using a principle to guide her decision. There could be several. For example, the principle itself may suffer from defects that prevent its practical use, or several principles that together comprise a moral code may jointly suffer such defects. Thus a principle might incorporate concepts that on closer reflection turn out to be incoherent. It is often claimed, for example, that act utilitarianism incorporates the concept of comparing the utility of one person to that of another person, and that this concept is incoherent or can be given no meaning. Or a principle may be so vague or ambiguous that it leaves the moral status of some actions indeterminate. Consider a principle which states that killing persons is wrong, but fails to clarify whether "persons" include human fetuses or not. No one can use this principle (in isolation) in deciding whether it is wrong to obtain an abortion, since no one can tell whether or not abortions are prohibited by it. Similarly,

---

<sup>7</sup> Paul Boghossian argues that the Kripke/Wittgenstein debate should not be construed as a debate about what it is to follow a rule, since the "ordinary concept of following a rule -- as opposed to merely conforming to one -- is the concept of an *intentional* act: it involves the intentional attempt to bring one's behavior in line with the dictates of some grasped rule." See "The Rule-Following Considerations," *Mind*, XCVIII, 392 (October 1989), 516-17.

a single principle (or several principles jointly) might, when applied to a given situation, deliver inconsistent assessments that an agent both ought to do A and ought not to do A. Or a principle might prescribe, as the only permissible choice in a given situation, a type of action that is not available. Thus a personnel policy requiring an employer to hire the job candidate who has better qualifications than any other applicant would be useless in a hiring situation in which there are two equally well-qualified superior candidates. Some of the barriers just described not only prevent the principles affected by them from being used to make (at least certain) decisions, but they also undermine the principles' adequacy as theoretical accounts of right and wrong. For example, suppose a moral code includes two principles, "It is all-things-considered wrong to kill an adult human being," and "It is all-things-considered obligatory to kill one's sister when she has dishonored the family through sexual contact outside of marriage." This moral code delivers inconsistent evaluations of a brother's act of killing his sister for a sexual transgression, because it evaluates the killing as both wrong and obligatory. Such a code is flawed *both* as a theory *and* as a practical guide. Moreover, its defects as a practical guide seem to depend directly on its defects as a theoretical account of right and wrong. Clearly the appropriate response to a code that is flawed in one of these ways is to replace the code with one that is more acceptable as a theoretical account.

But even a principle (or code) that exhibits none of these flaws *qua* theory may still be unusable for actual decision-making because agents lack beliefs about the relevant non-moral facts. To see how this can arise, let us examine more deeply our account of what it is for a principle to be usable for making a decision.

## **2. Two different kinds of usability**

We said that, roughly speaking, that a principle is usable by an agent for making a decision just in case it is true that if the agent wanted all-things-considered to act in conformity with the principle, the agent would choose an act out of her desire to conform to the principle and a belief that the act does conform.

Although this formulation captures an important idea, it requires slightly too much. To require that a principle be “action guiding” suggests that an agent’s use of the principle actually result in the agent’s performing an act under the guidance of the principle. But an agent might be able to use a principle, in the sense that she could derive a prescription from the principle, without then going on to perform the act, or even going on to *choose* to perform that act. Whether or not she takes these further steps depends on other factors, not on her ability to use the principle itself. Thus our rough definition should be refocused to say that a principle is usable by an agent for making a decision just in case it is true that if the agent wanted all-things-considered to derive a prescription from the principle, the agent would do so.

Even this rough account is too simple, since it is ambiguous between two different possibilities. Consider John, who wants to follow the principle “Give an anniversary card to one’s spouse if, and only if, today is one’s wedding anniversary.” John believes that today is their wedding anniversary, and by reference to this principle, decides to give his wife a card. However, John is mistaken: the anniversary is next week. There is an obvious sense in which he has *not* regulated his decision in accordance with his principle—but another obvious sense in which his decision clearly *has* been guided by it. Reflecting on this case, we may draw a distinction between having the ability in the *narrow* sense to use a principle to guide one’s decision, and having the ability in the *broad* sense to use a principle to guide one’s decision. Thus we can say, roughly speaking, that an agent is able to use a principle in the *narrow* sense

just in case he would derive a prescription for action from the principle if he wanted all-things-considered to do so.<sup>8</sup> By contrast, an agent is able to use a principle in the *broad* sense just in case he would derive a prescription for action from the principle if he wanted all-things-considered to do so, and the action for which he would derive a prescription is one that he would be able to perform, and also one that would conform to the principle.<sup>9</sup> Thus John, who decides to give his wife an anniversary card, is able to use his principle in the narrow sense, since he is able to derive a prescription from his principle. However, he is not able to use it in the broad sense, since the action for which he derives the prescription is not one that conforms to the principle. By contrast, Seth wants to follow the same principle, knows that his anniversary is indeed today, and so decides to give his wife a card. Seth has the ability to use this principle in both the narrow sense and in the broad sense.

Thus we should distinguish being a principle's being usable either narrowly or broadly as a decision guide by a given agent, depending on which of these abilities the agent has with respect to the principle. Any agent who has either kind of ability is able to derive a prescription

---

<sup>8</sup> In this discussion I am using the term "prescription" unusually broadly so that it can refer to a prescription to perform some act as obligatory – but also refer to a *proscription* not to perform some act as wrong, or a *permission* to perform some act as morally optional. However, for brevity, I shall focus on the case of prescriptions for obligatory acts.

In ordinary conversation, to say that an act is obligatory is not to say precisely the same thing as to say that it ought to be done all-things-considered. Obligatory acts are a narrower class than acts that ought to be done (it can be true that I ought to donate to famine relief efforts without it being true that I have an obligation to do so), and an obligation (say, to repay my debt) is normally understood as holding even though all-things-considered it ought not to be satisfied because some conflicting moral consideration outweighs it, whereas what ought to be done all-things-considered cannot be outweighed by any conflicting moral consideration. Regrettably English does not provide a common noun for acts that "ought to be done," so for stylistic reasons I shall normally use the term "obligation" for an act that ought to be done, and will refer to such acts as "obligatory." "Duties" are much like "obligations" in these respects.

<sup>9</sup> See Holly M. Smith, "Making Moral Decisions," *NOUS*, Vol. 22 (1988): 91-92, and also "Subjective Rightness," *Social Philosophy and Policy* Vol. 27, No. 2 (Summer 2010), pp. 64-110 for earlier discussions of a closely related distinction.

from the principle (and in consequence can use that principle in deciding what to do). What distinguishes an agent who has the ability to use a principle *merely* in the narrow sense from an agent who has the ability to use a principle in the broad sense is that the latter agent, but not the former, has *true* beliefs about what action is prescribed by the principle: if he follows through in performing the action he believes is prescribed, he will act in a way that conforms to the principle.

Of course, merely *having the narrow or broad ability* to derive a prescription from a principle does not imply that one actually derives any prescription, since one might not exercise that ability. When a person *exercises her narrow ability* to use some principle P as a decision guide, she wants to derive a prescription from P, and this desire leads to her deriving such a prescription.<sup>10</sup> When a person *exercises her broad ability* to use some principle as a decision guide, she wants to derive a prescription from action from P, and this desire leads to her deriving a prescription for action which is such that the person can perform the action whose prescription she derives, and it is true that if she performed it, it would conform to P. As we have remarked, even exercising the narrow or broad ability to use a principle as a decision guide does not necessarily involve the agent's choosing to perform, or performing, the action whose prescription she derives. Exercising either ability simply involves the agent's deriving a prescription from the principle for an action. She might derive a prescription from the principle but fail to follow through, either because she has no allegiance to the principle (she notices that the Levitical dietary law forbids consuming the pork chop in front of her, but cares nothing about this), or because her conflicting interests outweigh her allegiance to the principle (although she is an

---

<sup>10</sup> A more precise characterizing of "exercising an ability" would stipulate that the desire leads to the derivation via a non-deviant path. Since issues about deviant paths (or causal chains) are irrelevant to our main concerns, I shall not try to incorporate this condition in this or subsequent definitions.



observant Jew, the pork chop is too tempting to pass up), or even because she is immediately distracted and turns her attention to another matter. Nonetheless having the broad ability to use a principle as a guide is a distinctly stronger ability than merely having the narrow ability to use the principle as a guide. In the case of broad ability to use a principle, there are constraints on the nature of the act whose prescription the agent derives: it must be one that the agent could perform, and it must be such that if the agent were to perform it, it would actually conform to the principle in question. By contrast, in the case of having the narrow ability to use a principle, the act whose prescription the agent derives might be one she could not perform (perhaps she has been struck with sudden paralysis), or it might not comply with the principle (from the principle that one must pay one's debts, she derives a prescription to send a check to her creditor and dispatches a check, but the check is lost in the mail).

### **3. Immediately helpful descriptions**

To fully clarify the concept of a principle's being usable for making decisions, still more work needs to be done. As Eugene Bales pointed out, in an obvious sense an agent often "can derive" a prescription from a principle, even though the agent cannot derive the prescription for any action under what Bales calls "an immediately helpful description."<sup>11</sup> Thus someone who has no idea what the consequences would be of her various alternatives can still derive a "prescription" from act utilitarianism – she can derive the prescription "Perform the act that would maximize utility." In terms of actually using the principle to guide action, this is no help if she can't identify which act this is. To deal with this problem we need a more complicated

---

<sup>11</sup> Eugene Bales, "Act-Utilitarianism: Account of Right-Making Characteristics or Decision-Making Procedure?" 261.

definition of what it is to be able to use a normative principle. The following definitions outline a strategy for dealing with this problem. I will start by defining what it is for an agent to have the ability in the narrow sense to use a principle as a decision guide, and then use this account to define what it is for a moral principle to be usable in the narrow sense. Later I will introduce accounts of “ability in the broad sense” on the same model.

**1. Ability in the narrow sense to use a moral principle as a decision guide:** An agent *S* is able in the narrow sense at  $t_i$  to use moral principle *P* to decide at  $t_i$  what to do at  $t_j$  *if and only if*:

- (A) there is some (perhaps complex) feature *F* such that *P* prescribes actions that have feature *F*, in virtue of their having *F*,
- (B) *S* believes at  $t_i$  of some act-type *A* that *S* could perform *A* (in the epistemic sense) at  $t_j$ ,
- (C) *S* believes at  $t_i$  that if she performed *A* at  $t_j$ , act *A* would have *F*, and
- (D) if *S* believed that *P* prescribes actions that have feature *F*, in virtue of their having *F*, and if *S* wanted all-things-considered at  $t_i$  to derive a prescription from principle *P* at  $t_i$  for an act performable at  $t_j$ , then her beliefs together with this desire would lead *S* to derive a prescription for *A* from *P* in virtue of her belief that it has *F*.<sup>12</sup>

---

<sup>12</sup> For completeness we would need to state specific variants of Definition 1 setting out whether the “prescribed” acts are obligatory, forbidden, or permissible.

In this and subsequent contexts I understand “want” in a very broad sense, so that it includes not just ordinary desires (e.g., the desire to have an ice cream), but also moral motivations (e.g., respect for duty).

In certain cases whether or not an agent would do something if she wanted to depends on *how much* she would want to do it. Normally, for example, I would not lift the front end of my car even if I wanted all-things-considered to do so. But if my child were being crushed by the car, I might want *hard enough* to lift the car that I would do so (in such cases special physiological factors kick in). In most cases whether or not an agent can apply a moral principle to her decision does not depend on phenomena of this sort, so we do not need to accommodate them in this definition.

We can distinguish between (a) an agent who has (say) the general ability to play the piano, but has no specific ability to play the piano right now, since there is no piano available, and (b) an agent who

With this definition in hand, we can also define its correlate, a principle's arrow usability as an action guide:

**2. Narrow usability as an action guide:** A moral principle P is usable in the **narrow** sense by an agent S to decide at  $t_i$  what to do at  $t_j$  *if and only if* S is able in the narrow sense at  $t_i$  to use principle P as a guide to decide at  $t_i$  what to do at  $t_j$ .

Thus suppose principle P prescribes carrying out one's job responsibilities, and Sally's job responsibilities include opening the safe. Sally believes that she could open the safe, and also believes that her opening the safe would carry out her job responsibilities. If Sally believed that

---

has the specific ability right now to play, since a piano is right in front of her. Similarly an agent might have the general ability to use a principle in the narrow sense, but not have the ability right now to do so, since she doesn't believe of any act that it has F. Definition 1 focuses on the latter sort of specific ability.

There is a possible variant on Definition 1 that would structure clauses (B) and (C) so that they do not require that S actually *have* the requisite beliefs at  $t_i$ , but instead require that if S wanted to derive a prescription from P, she *would then come to have* these beliefs, for example after consulting her memory. I shall not adopt this way of defining "ability to use a moral principle" because of issues about *how long* it would take S to come to have the beliefs in question. She might come to have them instantaneously, or she might come to have them after several minutes of reflection – or even longer. Since many moral decisions must be made on the spot, without the agent's having time for further reflection, it seems most appropriate to require that she have them at the time she is said to have the ability to use the principle. On this view, it might be the case that S does not have the ability at  $t_i$  to use P, but comes to have the ability at slightly later time  $t_j$  to use P.

Note again that this definition, as currently stated, is vulnerable to "deviant causal chain" counterexamples. Since these are immaterial to our main concerns, I shall not complicate the definition in order to forestall such counterexamples.

Definition 1 defines an *epistemic* ability to use a principle as a decision guide, since the agent's ability depends in part on her having the beliefs described in clauses (B) and (C). Although one could define a parallel *non-epistemic* ability to use a principle, the notion of epistemic ability is the most fruitful one for the purposes of this inquiry.

Definition 1 and the allied definitions to follow are all accounts that use subjunctive conditionals to define the concept of "ability." An assortment of serious problems for such definitions has been raised in the literature. Since these problems are not germane to our immediate concerns, I shall not try to address them. For a good substantive review of these issues, see John Maier, "Abilities", *The Stanford Encyclopedia of Philosophy (Fall 2011 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2011/entries/abilities/>.

P prescribes carrying out one's job responsibilities and she wanted to derive a prescription from P for an act performable at t, then her beliefs together with this desire would lead her to derive a prescription to open the safe. Principle P is usable by Sally, in the narrow sense, as a decision guide. This is true even though Sally may be mistaken that she can open the safe (the combination has been changed), or mistaken that opening the safe would carry out her job responsibilities (this responsibility has been assigned to another employee). Still, Sally can *make a decision* by reference to P; she can choose to do what she believes P requires.

It may seem that this definition doesn't succeed in solving Bales' problem of ensuring that an agent can derive a prescription from P in terms of an act under an "immediately helpful description." For suppose a different agent, Molly, believes that P requires her to carry out her job responsibilities, but she has no idea what those responsibilities are, or in particular that they require her to open the safe. We would want to say that P is not usable by Molly even in the narrow sense. Still, it looks as though Definition 1 implies that Molly *does* have the ability to so use principle P (and so P is usable by her), because she believes that there is some act-type (namely, *carrying out her job responsibilities*) that she can perform, and she believes that if she carries out her job responsibilities, her doing so will have feature F, namely the feature of carrying out her job responsibilities. Hence she can derive a prescription from P to carry out her job responsibilities. But this gets her nowhere, since she doesn't know more specifically what her job responsibilities require.

Fortunately, despite appearances, Definition 1 avoids this unwanted implication, because it requires, via clause (B), that Molly believes at  $t_i$  of some act-type A that she could perform A *in the epistemic sense* at  $t_j$ . To see why the unwanted implication is avoided, we need to explain what it is to have the ability to perform an act in the epistemic sense of ability. The

distinction between being able in the epistemic sense to do something, and being able in the non-epistemic sense to do something, is meant to capture the difference between Alice, who is able in the epistemic sense to turn on the light (she is standing in front of the light switch and correctly believes that the switch is connected to the light and would turn the light on if flipped), and Alan, who is able in the non-epistemic sense but not in the epistemic sense to turn on the light (Alan is also standing in front of the light switch and has the physical ability to flip it, but since he doesn't realize that *this* switch is the one connected to the light, he would not flip it if he wanted to turn on the light).<sup>13</sup>

Here is a characterization of epistemic ability, based on Alvin Goldman's definition of this idea:

**3. Epistemic ability to perform an act:** S has the epistemic ability at  $t_i$  to perform act A at  $t_j$  *if and only if*:

- (1) There is a basic act-type  $A^*$  which S truly believes at  $t_i$  to be a basic act-type for her at  $t_j$ ;
- (2) S truly believes at  $t_i$  that she is (or will be) in standard conditions with respect to  $A^*$  at  $t_j$ ; and
- (3) either
  - (a) S truly believes at  $t_i$  that  $A^* = A$ , or

---

<sup>13</sup> There may be an additional complication to spelling out the notion of *epistemic ability to do A* that I shall not try to address. Consider Alex, who, like Alice and Alan, is standing in front of a light switch which he has the non-epistemic ability to flip. However, Alex's identical twin Max is also standing in the room in front of a switch – the switch for an exhaust fan. Alex and Max are standing in a room full of mirrors which create many images of individuals standing in switches. The mirrors are arranged in such a way that Alex cannot tell which images are images of him, and which images are of Max. Although Alex knows he is standing in front of the light switch, he may lack the epistemic ability to turn on the light, because he lacks the ability to *identify himself*. A thorough definition of *epistemic ability* might need to incorporate some treatment of what the literature calls “self-locating beliefs,” but I shall not attempt this. For a survey discussion of self-locating beliefs, see Uriah Kriegel, “Self-Consciousness,” Section b of “Essential Indexicals and De Se Thought,” in James Fieser and Bradley Dowden, eds., *The Internet Encyclopedia of Philosophy* <http://www.iep.utm.edu/self-con/> <http://> ( Last updated: June 16, 2007 | Originally published: June/16/2007) . I am grateful to Alvin Goldman for pointing out this issue.

(b) S truly believes at  $t_i$  that there is a set of conditions  $C^*$  obtaining at  $t_j$  such that her doing  $A^*$  would generate her doing  $A$  at  $t_j$ .<sup>14</sup>

Sally has the epistemic ability to carry out her job responsibilities because (1) there is a basic act-type, namely *moving her fingers in ways XYZ*, which she truly believes to be a basic act-type for her, and (2) she further truly believes that she will be in standard conditions with respect to moving her fingers in ways XYZ, and (3) she further truly believes that if she moves her fingers in ways XYZ, this will generate her opening the safe, and this in turn will generate her carrying out her job responsibilities. In short, she knows *how* to carry out her job responsibilities, and she is physically able to do what she knows how to do.

By contrast Molly does not have the epistemic ability to carry out her job responsibilities. It is true of Molly that (1) there is a basic act-type, namely *moving her fingers in ways XYZ*, which she truly believes to be a basic act-type for her, and (2) she further truly believes that she will be in standard conditions with respect to moving her fingers in ways XYZ. However, it is *false* that (3) she truly believes that if she moves her fingers in ways XYZ, this will generate her opening the safe, and this in turn will generate her carrying out her job responsibilities. Of course Molly believes that carrying out her job responsibilities will carry out her job responsibilities. But there is no *basic act* which she truly believes will generate her carrying out her job responsibilities. Because she does not know how to do this, she can't get started.

---

<sup>14</sup> Alvin I. Goldman, *A Theory of Human Action* (Princeton: Princeton University Press, 1970), 203. "Act-types" are certain kinds of properties an agent may exemplify. *Moving one's finger* would be a basic act-type (for most people), while *pulling the trigger* would be a non-basic act-type whose exemplification would be generated by a person's moving his finger when his finger is positioned on a gun's trigger. Roughly speaking, a person is in standard conditions with respect to an act property just in case (a) there are no external physical constraints making it physically impossible for the person to exemplify the property, and (b) if the property involves a change into some state Z, then the person is not already in Z. See Goldman, *A Theory of Human Action*, 64-65.

Because of this, Molly fails to satisfy clauses (B) and (C) of Definition 1 for **Ability in the narrow sense to use a moral principle as a decision guide**. Molly does not believe of any relevant act-type A both that she could perform A (in the epistemic sense) and that A would have the right-making feature F identified by principle P, namely *carrying out her job responsibilities*. Thus inclusion in Definition 1 of the requirement that an agent believe she has the epistemic ability to perform an action that would fulfill the principle's prescription enables Definition 1 to ascribe ability to use a principle to an agent only when the agent, as Bales puts it, possesses an "immediately helpful description" of an action prescribed by the principle. Molly does not count as having the ability, in the narrow sense, to use her moral principle as a decision guide, and for similar reasons an agent who can derive from utilitarianism the prescription to maximize utility, but who doesn't know which concrete act would do this, does not count as someone who has the ability, in the narrow sense, to use utilitarianism as a decision guide.

Having spelled out a definition of what it is to have the ability in the narrow sense to use a moral principle as a decision guide, let us now define what it is for an agent to be able in the broad sense to use a moral principle. This definition builds on Definition 1 by crucially altering clauses (B) and (C) to stipulate that the agent's beliefs are true:

**4. Ability in the broad sense to use a moral principle as a decision guide:** An agent S is able in the broad sense at  $t_i$  to use moral principle P to decide at  $t_i$  what to do at  $t_j$  *if and only if*:

- (A) there is some (perhaps complex) feature F such that P prescribes actions that have feature F, in virtue of their having F,
- (B) S *truly* believes at  $t_i$  of some act-type A that she could perform A (in the epistemic sense) at  $t_j$ ,
- (C) S *truly* believes at  $t_i$  that if she performed A at  $t_j$ , act A would have F, and

(D) if S believed that P prescribes actions that have feature F, in virtue of their having F, and if S wanted all-things-considered at  $t_i$  to derive a prescription from principle P at  $t_i$  for an act performable at  $t_j$ , then her beliefs together with this desire would lead S to derive a prescription for A from P in virtue of her belief that it has F.<sup>15</sup>

Given this definition, we can also define its correlate, a principle's broad usability as an action guide:

**5. Broad usability as an action guide:** A moral principle P is usable in the **broad** sense by an agent S to decide at  $t_i$  what to do at  $t_j$  *if and only if* S is able in the broad sense at  $t_i$  to use principle P as a guide to decide at  $t_i$  what to do at  $t_j$ .

Thus, as in the first safe-opening case, suppose P prescribes carrying out one's job responsibilities, and Sally's job responsibilities include opening the safe. Let's imagine that Sally truly believes that she could open the safe, and also truly believes that her opening the safe would carry out her job responsibilities. Moreover, if Sally believed that P prescribes carrying out one's job responsibilities and she wanted to derive a prescription from P for an act performable at  $t$ , then her beliefs together with this desire would lead her to derive a prescription to open the safe. Principle P is usable in the broad sense by Sally as a decision guide. If she

---

<sup>15</sup> There is a possible Gettier-like anomaly allowed by Definition 4. Suppose principle P states "A is obligatory if and only if A has F," where  $F = G$  or  $H$ . Perhaps F is *treating one's children appropriately*, and  $G =$  *leaving one's estate to one's children if they have honored their parents* while  $H =$  *disowning one's children if they have failed to honor their parents*. Fred believes that his act A (signing a certain document) has G, and is able to derive a prescription for A from principle P in virtue of his belief that A has G and therefore has F. However, A does not have G, although it does have H (imagine that Fred is being defrauded, since unbeknownst to him, signing the document actually disowns Fred's son – but also imagine that his son has dishonored Fred behind his back). Thus A would conform to principle P, although not in virtue of the precise feature Fred believes it to have. Under Definition 4, Fred would still count as having the ability in the broad sense to use P as a decision guide, even though this ability rests on the accident of A's possessing a feature that Fred does not believe A has. I shall not attempt to revise Definition 4 to avoid this kind of case.



proceeded to derive a prescription and to carry it out, she would succeed in doing what principle P requires her to do. Because P is usable by Sally in the broad sense, we can infer that she is not mistaken either about whether she can open the safe, or about whether it is her job responsibility to open the safe. Because her beliefs about the situation are true, she can both derive a prescription from P and also carry out what P actually requires of her. A third agent Polly, who is in the same situation as Sally, and who has the same beliefs that Sally does, but whose beliefs are false – say, she falsely believes that dialing 60 – 89 – 35 will open the safe – qualifies as someone who can use P in the narrow sense as a decision guide, but she does not qualify as someone who can use P in the broad sense. She fails to satisfy clause (C) of Definition 4.

Given these definitions of what it is for a moral principle to be usable by an agent in either the narrow or broad sense, the advocate of the Applicability Demand must decide whether to interpret the Demand as a requirement for narrow or broad usability. Interpreted merely as a requirement for *narrow* usability, the Demand stipulates only that agents have beliefs that would enable them to derive a prescription from the principle in question. These beliefs may be erroneous (as in Polly's case), but the principle will still count as usable by such an agent. Interpreted as a requirement for *broad* usability, the Demand stipulates not only that agents have relevant beliefs, but also that these beliefs be *true*. On this interpretation of the Applicability Demand, agents must have relevant, accurate beliefs about any potential action for which they might try to derive guidance from the principle. This is obviously a much stronger requirement, and one that will be met by many fewer principles. Advocates of the Applicability Demand have not always been clear about which version of the Demand they wish to defend. Considerations supporting the two versions may derive from different sources. For example, the virtue of the narrow-usability version is that it only endorses a moral principle that enables agents to identify

acts they believe the principle prescribes: it does not leave them in the dark about what choice to make if they want to make a moral choice. The broad-usability version, on the other hand, has two virtues: it both enables agents to identify acts they believe the principle prescribes, and it also ensures that if they choose that action, their action will actually conform to the principle. A moral theorist who is primarily interested in the interior lives of agents may be content with mere narrow-usability, while a moral theorist who is primarily interested in securing actual right conduct from agents may need to adopt broad-usability instead.

#### **4. Issues and responses**

Several issues and implications of these definitions are worth remarking. First, Definitions 2 and 5 characterize a moral principle's usability by an individual agent at some specific time  $t_i$ . Proponents of the Applicability Demand tend to assume (implicitly or explicitly) that an acceptable moral principle must be usable as a decision guide by all moral agents on all occasions on which they have the opportunity to make a moral decision covered by that principle. However, there are weaker possible versions of the Demand: one could require that the principle be usable by the majority of agents, or by each agent on the majority of occasions, or by each agent who has investigated the situation as thoroughly as he ought to have done, and so forth. Whichever version of the Demand is adopted, Definitions 2 and 5 can be adopted for stating it.

Second, Definitions 1 and 4 require S to have certain beliefs about the nature of her options in order to be able to use the principle in question. But what do they imply for cases in which P requires act A in virtue of its having feature F, and S believes of A that it has F, but is unaware of principle P itself, or unaware or uncertain what kinds of actions it prescribes? For

example, an agent might never have heard of utilitarianism, or might be aware of utilitarianism merely as the moral theory that Bentham and Mill advocated, but not know what the content of the principle is, and so not be aware of the kinds of actions utilitarianism prescribes (those that maximize utility). Assuming the various conditions are met, the definitions imply that such a morally ignorant S is able to use P. This seems to be an appropriate feature of one important sense of the term “able to use.” A parallel use of “ability” is our saying that S (who knows how to iron) is able to use the iron in the closet of her hotel room, even though she is unaware of the existence of the iron. If she *were* aware of the iron, she would know how to use it. In expressing this sense of “able to use,” in effect Definitions 1 and 4 set aside *moral* ignorance or uncertainty as potential barriers to the use of a moral principle to guide decisions. Clearly there is another sense of “able to use” that would deny S has the ability to use an iron of whose existence she is unaware, and would also deny that S can use a moral principle with which she is unfamiliar. But this sense does not seem pertinent to the concerns that motivate the Applicability Demand.

Third, the definitions are phrased in terms of an agent’s *beliefs*, not in terms of her *credences*, where an agent’s credence is her degree of belief or subjective probability (which may fall anywhere from 0 to 1.0) for some proposition. Thus the definitions imply that an agent lacks the ability to use her moral principle if, for example, she has some moderately high degree of belief that, say, her opening the safe will carry out her job responsibilities, but does not fully believe or feel certain about this matter. Of course most agents must make decisions in light of less-than-certain credences, so this feature of the definitions entails that most agents are unable to use standard moral principles in making their decisions. This aspect of the definitions is correct. An agent who merely has a moderately high degree of belief that her opening the safe would carry out her job responsibilities does *not* thereby have the ability to use principle P

requiring her to carry out her job responsibilities.<sup>16</sup> Of course, she *does* have the ability to use a different principle such as P\*, which prescribes performing an act if it is highly likely to carry out her job responsibilities. What this shows is that an advocate of the Applicability Demand should recommend abandoning P in favor of P\* or some similar less epistemically demanding variant on P. Indeed, many Applicability Demand supporters have recommended exactly this.<sup>17</sup> The kinds of epistemic deficiency that stand in the way of an agent's being able to use a moral principle include not only failure to have any relevant belief, and failure to have relevant true beliefs, but also failure to have sufficient *certainty* to draw the required conclusions.

Fourth, the definitions do not address the question of whether the agent's relevant beliefs can be *unconscious*. Suppose, for example, that Sally no longer consciously believes that her dialing 60 – 89 – 35 (in her case, the correct combination) would open the safe; possibly she could not retrieve this belief to consciousness even if she tried. She has been opening the safe for so long that when she wants to open it, she simply goes to the safe and goes through the appropriate motions, guided by an unconscious belief that dialing 60 – 89 – 35 would open the

---

<sup>16</sup> Clearly there are many levels of credence (for example, credence in the proposition that A has F) at which an agent is unable to derive a prescription from P, although the agent could certainly derive a *conditional* prescription to the effect that *if A has F, then P prescribes A*. But the question of whether there is some threshold level of credence below 1.0 (say, .9 or above) at which the agent could derive an *unconditional* prescription from P is a knotty one, closely tied to the difficult question of whether there is some threshold level of credence below 1.0 that counts as knowledge. I shall not try to address this question here. One might argue that a form of “pragmatic encroachment” affects the answer to this question: perhaps it is legitimate to derive an unconditional prescription from P when one has a lower level of credence, so long as the stakes are also low; but it is only legitimate to derive a prescription from P when one has a very high level of credence if the stakes are also very high.

<sup>17</sup> H. A. Prichard in “Duty and Ignorance of Fact” and W.D. Ross in *Foundations of Ethics* are only the most dramatic examples of theorists who make this recommendation.

I argue in “Making Moral Decisions” and *Making Morality Work* that an alternative is to retain the Applicability Demand, but view it as satisfied if an agent can use principle P *indirectly* by inferring from P to an auxiliary decision guide (for example, “It would be subjectively right to perform the act that is most likely to carry out one's job responsibilities”) that reflects the values endorsed by P but is itself directly usable (according to Definitions 1 or 4) by the agent because it is less epistemically demanding.

safe. Even though this belief is unconscious, does she count as having the ability to use P in either the narrow or the broad sense? In this case it certainly seems as though she should count as being able to use P, since if she wanted to carry out her job responsibilities, she would do the right thing by opening the safe. The question about unconscious beliefs is a difficult one, but at first glance it appears that we should allow unconscious beliefs to satisfy clauses (B) and (C) in Definitions 1 and 4, at least so long as clause (D) is also true when the unconscious beliefs satisfy clause (B) and (C).

One final note: many philosophers, following Bales, have characterized a moral principle as usable by an agent to make a decision only if there is a “decision procedure” that enables the agent to derive a prescription from the theory. Bales and some subsequent philosophers seem to have assumed that such a “decision procedure” is actually a “mechanical sequence of operations” that the agent could perform in order to arrive at an “overall moral verdict about the action in question.”<sup>18</sup> But a closer examination of what they have in mind by a “mechanical sequence of operations” suggests that the operations in question are operations by which the agent informs herself *whether A has F*, that is, informs herself about the non-moral but morally relevant aspects of her choices: operations by which she discovers what the consequences of her choices would be, what the past and present circumstances are in which these choices would be made, and so forth.<sup>19</sup> And many philosophers object to the idea that there can be any such “mechanical operations” – they note that agents must often call on delicate human skills such as capacity to detect the psychological states of those around them, imagination about the outcomes of certain

---

<sup>18</sup> These quotations are from Samuel Scheffler, *Human Morality* (New York: Oxford University Press, 1992), 39-40.

<sup>19</sup> In their discussion of a demand for “algorithmic decision-procedures” Sean McKeever and Michael Ridge seem to fall prey to this view (*Principled Ethics*, (Oxford: Clarendon Press, 2006), 11).

actions, ability to predict how things will turn out, judgment about how people would react to being treated in certain ways, and so forth. These philosophers value the exercise of such human skills, and balk at any depiction of applying a moral principle that side-steps exercise of these skills in favor of a more mechanical use of *modus ponens*.<sup>20</sup> We should acknowledge that gathering non-moral information relevant to a moral decision can be a sophisticated and complex procedure, not reducible to any purely “mechanical” operations. However, once that information is gathered, there is no moral value lost by an agent’s using *modus ponens* to infer that a given action is prescribed by his moral principle.

Of course many moral philosophers have claimed that sophisticated non-mechanical procedures must also be used in the more narrowly *moral* aspects of deliberation about what to do. Thus H. D. Ross, and following him the particularists, have emphasized the need to weigh or balance conflicting moral considerations in deciding what ought to be done all-things-considered. And they have insisted there is no mechanical way of doing this.<sup>21</sup> But if, at the end of the day, the decision maker arrives at a conclusion that some option ought all-things-considered to be done, he has derived a prescription from his moral theory. For such theories, the feature F of the act in virtue of which it ought all-things-considered to be done might be something like *fulfills the greatest net balance of prima facie duties, or has the greatest balance of favorable moral considerations*. The usability of such a theory to make decisions still fits under Definitions 1 and 4.

---

<sup>20</sup> See Scheffler, *Human Morality*, 38–48.

<sup>21</sup> See, for example, W.D. Ross, *The Right and the Good* (Oxford: Clarendon Press, 1930), and Jonathan Dancy, *Ethics Without Principles* (Oxford: Clarendon Press, 2004)

## 5. Conclusion

Proponents of the Applicability Demand have argued that any acceptable moral principle must meet the test of being usable by human agents – with all their epistemic limitations -- for guiding their decisions. Little rigorous work has been done in explaining exactly what is meant by a principle’s being usable for guiding decisions. In this paper I have proposed accounts of what it is for an agent to have the ability, in either the narrow or the broad sense, to use a moral principle as a decision guide, and corresponding accounts for what it is for a moral principle to be usable in either the narrow or broad sense. Because there are two interpretations of “usability” – narrow usability and broad usability – the Applicability Demand advocate must decide which of these kinds of usability is the one required. I have sketched different kinds of considerations that support one or the other of these two senses of usability as the relevant one. My accounts of an agent’s being able to use a moral principle have been crafted in a manner that enables us to answer Bales’ challenge of how to characterize an “immediately helpful description” of an act – the kind of description that an agent must have of an act in order to qualify as genuinely having the ability to use a moral principle to evaluate that act.

With this groundwork accomplished, we have laid an important piece of the foundation required to begin assessing the tenability of the Applicability Demand.

Holly M. Smith  
Rutgers University