THE GROUNDLESS NORMATIVITY OF INSTRUMENTAL RATIONALITY*

In *The Little Prince*, Antoine de Saint-Exupery\(^1\) has us imagine a monarch who, wanting above all to be obeyed, tailors his commands to the anticipated behavior of his subject. When the little prince yawns, the King commands him not to do so again. The little prince objects—pleading that he cannot help yawning—so the king commands the little prince to yawning. Frightened, the little prince is unable to yawn. The king struggles for a moment and then commands the little prince to “sometimes yawn and sometimes not.” As Saint-Exupery says:

> For what the king fundamentally insisted upon was that his authority should be respected. He tolerated no disobedience. He was an absolute monarch. But, because he was a very good man, he made his orders reasonable (*ibid.*, p. 37).

When the king boasts that he is the absolute ruler over all the universe and that even the stars and the planets obey him, the little prince asks the king to command a sunset. To this seemingly modest request of one with such power, the king replies:

> “If I ordered a general to fly from one flower to another like a butterfly, or to write a tragic drama, or to change himself into a sea bird, and if the general did not carry out the order that he had received, which one of us would be in the wrong?” the king demanded. “The general, or myself?”

\(^*\) A version of this paper was read at a colloquium at Bowling Green State University. I am grateful to the audience for helpful suggestions. I am especially grateful to the participants in the Ohio Reading Group in Ethics—in particular, Justin D’Arms, David Copp, David Sobel, and Sigrún Svavarsdóttir—for careful reading and extremely useful comments.

\(^1\) New York: Harcourt, Brace and World, 1943.
“You,” said the little prince firmly. “Exactly. One must require from each one the duty which each one can perform,” the king went on. “Accepted authority rests first of all on reason. If you ordered your people to go and throw themselves into the sea, they would rise up in revolution. I have the right to require obedience because my orders are reasonable.” “Then my sunset?” the little prince reminded him: for he never forgot a question once he had asked it. “You shall have your sunset. I shall command it. But, according to my science of government, I shall wait until conditions are favorable.” “When will that be?” inquired the little prince. “Hum! Hum!” replied the king; and before saying anything else he consulted a bulky almanac. “Hum! Hum! That will be about—about—that will be this evening about twenty minutes to eight. And you will see how well I am obeyed” (ibid., p. 38).

I. INTRODUCTION

What are we to make of the king and his commands? Were the king to refrain from giving commands only when the subject is incapable of complying, we would think this, in general, admirable and reasonable. To order what one knows to be beyond the capacity of an individual may sometimes have a point, but the point can never be the standard one of trying to bring about compliance. Indeed, were the king also to refrain from commanding those actions which impose too great a cost on his subjects—ordering them into the sea, for example—we might again think this a laudably restrained authority (at least provided nothing of overwhelming importance required such a sacrifice from his subjects). But, of course, the limitations the king places on his commands are, in fact, quite a bit more substantial than these. What the king refuses to do is to command (or at least persist in commanding) any action that will not be performed. As a result, while he is never disobeyed, neither does he (at least in the examples we see) exercise any authority at all.

The commands of the Saint-Exuperian king, while in the form of prescriptions, are too “reasonable” to be regulative. Rather than require that actions be adjusted to his commands, the king carefully tailors his commands to the actions that would be performed anyway. His commands are not normative in any interesting way. “Unreasonability” in the Saint-Exuperian sense is a requirement on a normative theory in order to avoid what might be called normative triviality.

To some, it has appeared that Humean instrumentalism—perhaps the most influential theory of practical rationality—is as bumbling
and ineffective in issuing practical norms as Saint-Exupery’s king. Humean instrumentalism grounds an agent’s reasons for acting on the agent’s desires, values, or some other subjective, contingent, conative states of the agent. The concern is that by doing this, the theory renders impossible the sort of normativity that a theory of practical rationality must have. Humean instrumentalism fails a criterion of adequacy for a theory of practical reasons—that of being “unreasonable” (in the Saint-Exuperian sense).

I argue, on the contrary, that the Humean instrumentalist can be as normatively demanding as it is evident that our judgments about practical reasons require. The critics have a complaint in this regard, but it is a question-begging complaint; it appeals to intuitions that the Humean does not, and ought not, share.

But the allegation of normative triviality is, I believe, better understood as a symptom of a different concern about Humean instrumentalism. What provokes the charge of normative triviality is the concern that Humean instrumentalism has its feet firmly planted in the air—that it offers reasons that are, in the end, ungrounded. Critics have thought that, at best, the Humean could offer an account of the transmission of practical reasons, never of their origin. So long as the Humean tells us that practical reasons are transmitted from ends to means, she is an instrumentalist. For reasons to be transmitted, however, they must exist in the first place. There must be reasons for the ends—at least those which are “ultimate” ends. And, to justify the claim that there are such reasons, the Humean must “go beyond” instrumentalism—she must commit herself to the existence of reasons that are not conditioned on the subjective, contingent, conative states of the agent.

This allegation, I argue, is based on a fundamental misunderstanding of the nature and purpose of Humean instrumentalism. On a proper construal, the theory adroitly side-steps these attacks. We begin, though, with the charge of normative triviality.

II. NORMATIVE TRIVIALITY

It is not completely clear how, precisely, to formulate the requirement that a theory be normatively demanding—that it avoid normative triviality. The story of the little prince might suggest that the

---

2 What I am here referring to as Humean instrumentalism I have elsewhere called, and shall later in this essay call, neo-Humean instrumentalism. The qualification is desirable to contrast it with the views that might plausibly be attributed to Hume himself. For purposes of introducing the issues here, I am setting aside my preferred terminology in order to speak in terms that are more common. In these commonly used terms, Hume seems not to be a Humean.
problem arises when apparent demands are always satisfied; these "commands" never require actions that are, in fact, not done. But obviously, it cannot be a requirement on a normative theory that its prescriptions actually be violated. This would be to make a mistake opposite to that made by Saint-Exupery’s king.

Suppose we require that a (nontrivial) prescription be one that the agent is capable of violating. This seems too restrictive. Surely, I can be guided by a norm, even if I do not have the capacity to act contrary to it. Cases of overdetermination suggest this conclusion. Imagine that I am one of Stanley Milgram’s real subjects in his famous experiments on obedience to authority. Suppose I am told to increase the electrical shock I apply to the supposed subjects but, guided by a norm not to harm innocent people, I refuse. Now, no matter what I had done—press the button or not—I would not have been able to harm the apparent subject. It was not within my capability to do so. It seems to me, however, that I was guided by the norm in question in this case, even though I did not have the capability to violate the norm.

We could say that the minimal demandingness necessary to avoid normative triviality requires that it must be possible to violate a norm.4 A putative norm is normatively trivial when it is not possible for an agent to violate the norm. But this runs into the overdetermination problem, too. It seems to me that, if I do not know that some apparent act description in fact refers to no possible action, I can be "guided by" a norm that condemns such "actions." For example, after listening to a logical proof given by a colloquium speaker, I might—falsely believing the demonstration fallacious but thinking it bad form to refute a guest in public—be guided by a norm not to demonstrate publicly the logical fallacies committed by colloquium speakers, even though it is logically impossible for me to violate this norm in this case.

I do not want to track this problem very far because I do not think precision in the formulation of the nontriviality requirement is necessary for my purposes. It seems to me, though, that the problem with Saint-Exupery’s king’s commands is a “direction of fit” problem. As I suggested before, the king tailors his commands, as much as he is able, to his expectations about the future actions of his subjects—indeed, to his expectations about the actions that his subjects would perform independently of his commands. The explanation of why

---

4 This seems to be the sort of requirement Christine Korsgaard endorses, as discussed below.
the king is never disobeyed runs from his expectations of their actions to the content of his commands. When norms are authoritative, they are not created to match independent expectations about the actions they “govern.”

III. NORMATIVE TRIVIALITY AND SUBJECTIVE EXPECTED UTILITY

Does the requirement that I have called (minimal) demandingness have any bite with respect to theories of practical rationality? It might. Some have suggested that one standard instrumentalist theory, the subjective expected utility (SEU) theory of practical rationality, does not satisfy the requirement of demandingness. The reasoning goes as follows.

SEU holds that one acts rationally when one acts so as to maximize one’s expected utility, where that notion is a measure of the probabilistically weighted utility of the various possible outcomes of the action. On this account, the probabilities, as well as the utilities, are measures of some subjective state of the agent. Probabilities are measures of the agent’s subjective degree of belief and utilities are measures of her subjective preference. But how do we determine an agent’s preferences and degrees of belief?

For expository reasons, let me drop the second part of this question and focus on the measurement of preference. One approach to measuring an agent’s subjective preferences—one that is still popular in some quarters—is based on the doctrine of revealed preference. This doctrine holds that “preference is revealed in choice.” The slogan is misleading. It could refer to a rather modest epistemic thesis: that choice behavior is evidence—perhaps, in many cases, the best evidence we have—of an agent’s preferences. In fact, though, it is a metaphysical thesis: preference just is choice behavior, or, more plausibly, dispositions toward choice behavior.

In a simple dichotomous choice situation, an agent prefers A to B just in case she takes (or would take) A rather than B when offered an exclusive choice. Making certain diachronic consistency assumptions, behavior (or dispositions to behave) over a suitably large number of pairwise choices, allows the attribution of measures of utility that are unique up to a positive linear transformation. And such measures are taken to be strong enough for doing the work of decision theory.

While I think the issue of triviality is ultimately to be cashed out in terms of “direction of fit,” I shall often speak as if the issue is simply whether the norm might require some action that would not be done independently of the norm. This is for ease of presentation only.
Now, if we use utilities, so constructed, as a basis for a normative SEU theory, a problem arises. Amartya Sen⁶ puts it this way:

If you are observed to choose \( x \) rejecting \( y \), you are declared to have "revealed" a preference for \( x \) over \( y \). Your personal utility is then defined as simply a numerical representation of this "preference," assigning a higher utility to a "preferred" alternative. With this set of definitions you can hardly escape maximizing your own utility, except through inconsistency (ibid., p. 322).

But the matter is worse than Sen recognizes. Behavior is a part of nature and nature does not permit contradictions. Unless we make substantive assumptions about an agent’s preference structure—assumptions which cannot be warranted on any purely formal grounds or on strictly behavioral grounds—we cannot convict an agent of inconsistency.

Clearly, what Sen has in mind when he excepts inconsistency is the case of the person with so-called incoherent preferences—for example, a person with a cyclic preference pattern. Such a person prefers \( A \) to \( B \), \( B \) to \( C \), and \( C \) to \( A \). This sort of cyclic preference pattern is usually alleged to be irrational on the grounds that an agent with such a preference pattern can be turned into a money pump—paying some small amount of a good, say $1.00, to make each of the following exchanges: \( B \) for \( C \), \( A \) for \( B \), and \( C \) for \( A \). She winds up "where she began but three dollars poorer."

Such a conviction, though, cannot be maintained on the strict behaviorist grounds offered by the doctrine of revealed preference. A simple example will illustrate well enough. Imagine what I shall call an “All-You-Can-Enjoy Spa,” consisting of a Jacuzzi, a sauna, and a swimming pool. One may, let us suppose, enter any room for the price of a dollar. Now imagine that, tired after a hard day’s work, Arnold pays a dollar to sit in the Jacuzzi for twenty minutes. Then, for a dollar, he moves to the sauna for a few minutes before paying yet another dollar to enter the swimming pool for a quick workout. Finally, tired from his workout, he pays a dollar to enter the Jacuzzi again.

We do not think that Arnold is “back were he started but three dollars poorer”—at least not in any sense that would lead us to accuse him of being a “money pump.” We can perfectly well make sense of the preferences that “rationalize” this behavior. We do not think he has done anything irrational or expressed inconsistent or incoherent preferences. But, behaviorally, there is nothing to distinguish him

from a classical money pump. Our only reason for condemning the classical money pump as displaying practical irrationality is that we balk at the implications of revealed-preference theory. In the stories told about money pumps, we refuse to think that there is an intelligible, coherent set of preferences that rationalize the behavior. But think it we must if we are inferring preferences from behavior as the doctrine commands us. So, perhaps the classical money pump desires to own each of the objects, A, B, and C, for some short period of time. Making the apparently cyclic trades, then, is no more irrational than is Arnold’s moving through the various rooms of the spa.

If we accept the doctrine of revealed preference, taking behavior not merely to be evidence of an agent’s preferences but to be constitutive of the agent’s preferences, the conclusion about the inevitability of maximizing expected utility is stronger than Sen realizes: with this set of definitions, you cannot escape maximizing your own utility, period. So, if we develop an instrumentalist normative theory of practical rationality along the lines of a SEU theory that understands utility as a measure of the agent’s subjective preferences and these preferences to be constituted by the agent’s behavior, our theory will be normatively trivial. It is “reasonable” in Saint-Exupery’s sense, and so, unreasonable as a normative theory.

But it is quite clear where lies the weak point in this cluster of claims—and it is not with the instrumental theory of practical rationality. Obviously, if we are to construct a normative theory of practical rationality along the lines of a SEU theory and we think that our theory must not be normatively trivial, then we cannot understand utility as being constituted by actual behavior. So much the worse for the doctrine of revealed preference as a definition of “utility.”

But this account of “utility” has little to recommend it anyway. Its starkly behaviorist credentials—perhaps attractive to an earlier generation of philosophers—do not hold out much attraction now even if our objective is a predictive or explanatory theory of intentional human behavior. Choice can never “reveal” preference in any situation but the specific, concrete situation in which it is observed. Behavior alone never tells us what are the relevant features of the situation. Because of this, we do not know, on behaviorist grounds alone, how to generalize from observed behavior in order to predict future behavior. Without background, nonbehaviorist assumptions, no future behavior is rendered inconsistent, or even less probable, based on observed past behavior. The revealed preference version of SEU theory is not predictive, unless we make assumptions about agents’ preferences which are not behavioristically based. And in
what significant sense can it be explanatory to account for behavior in terms of utilities that are just summaries of behavior?

The doctrine of revealed preference is problematic all by itself. It is not surprising that, when conjoined to an instrumentalist, preference-based theory of practical rationality, it results in a problematic theory. Indeed, it would be quite remarkable if it did not. To paraphrase J.S. Mill: there is no difficulty in proving any standard of rationality whatever to work ill if we suppose crude behaviorism to be conjoined to it. It seems obvious that one who seeks an adequate instrumentalist, preference-based theory of practical rationality should look to some other account of "preference." But, lately there has been a growing concern that this is not possible. More precisely, the thought that some have promoted is that a satisfactory account of preference will be inconsistent with an instrumentalist, preference-based theory of practical rationality.

Christine Korsgaard is one of the prime movers of this criticism. According to Korsgaard, David Hume, like the revealed-preference theorist, is committed to the view that people cannot violate the instrumental principle. Indeed, I should have said that, on Korsgaard’s reading, Hume, like other revealed-preference theorists, is committed to the view that people cannot violate the instrumental principle, for Korsgaard interprets Hume as holding what amounts to a revealed-preference theory. Let us grant the interpretive point.

7 "Utilitarianism," in Selected Writings of John Stuart Mill, Maurice Cowling, ed. (New York: Mentor, 1968); for the original quote which concerns the need for usable decision procedures in moral practice, see p. 265.


9 I am dubious, though. As we shall see momentarily, Korsgaard thinks the problems arise for Hume because, "Hume identifies a person’s end with what he wants most, and the criterion of what a person wants most appears to be what he actually does" (op. cit., p. 230). In a very famous passage, however, Hume seems to recognize the possibility of failing to do what you want. Here is the familiar passage:

'[T]is only in two senses, that any affection can be call’d unreasonable. First, When a passion, such as hope or fear, grief or joy, despair or security, is founded on the supposition of the existence of objects, which really do not exist. Secondly, When in exerting any passion in action, we chuse means insufficient for the design’d end, and deceive ourselves in our judgment of causes and effects—A Treatise of Human Nature (New York: Oxford, 1968), p. 416.

Now it is, I think, quite inconsistent for someone simultaneously to believe that we may, in action, sometimes choose means insufficient for our designed ends and to believe, as Korsgaard suggests Hume does, that a person’s ends are revealed in
What are its implications? Korsgaard believes that a principle it is impossible to violate cannot be normative; one cannot be guided by such a principle. As was argued above, this is not a satisfactory account of normative triviality. But assuming revisions along the lines suggested there—revisions to which she would probably not object—I agree with Korsgaard on this point.

Korsgaard uses this criterion, which I have called *demandingness*; to pose a problem for neo-Humean instrumentalism. The problem has the form of a dilemma within a dilemma. The first dilemma is this: either the neo-Humean instrumentalist’s theory is normatively trivial, or it draws a distinction between an agent’s ends and what the agent, in fact, does. The second, nested, dilemma is a constructive one seeking to show that whichever of two paths the neo-Humean takes to draw this distinction, he must “go beyond” an instrumentalist theory. So, overall, we are pushed to the conclusion that either neo-Humeanism instrumentalism is normatively trivial or, contrary to advertising, it is committed to its own noninstrumental, unconditional rational requirement—its own “categorical imperative.”

**IV. INSTRUMENTALISM AND NORMATIVE DEMANDS**

Can neo-Humean instrumentalism be demanding? In discussing the potential undemandingness of the SEU theory, we located the source of the problem in the doctrine of revealed preference—a doctrine sometimes phrased so as to appear merely epistemic, but, in fact, metaphysical to the core. Not surprisingly, in her discussion of Hume’s theory, Korsgaard does the same:

The problem is coming from the fact that Hume identifies a person’s *end* with what he *wants most*, and the criterion of what a person wants most appears to be what he actually *does*. The person’s ends are taken to be revealed in his conduct.... So, the problem would be solved if we could make a distinction between a person’s ends and what he actually pursues (op. cit., p. 230).

Can we make the necessary distinction? It would seem obvious that we can, and not merely ad hoc so as to satisfy the criterion of demandingness.

First, what may be a terminological complaint: Korsgaard holds that the distinction we need to make is between a person’s ends and “what he actually *pursues*” (emphasis added). But the doctrine of revealed preference says nothing about what a person pursues (in any
natural sense of ‘pursues’). Rather, the doctrine of revealed preference holds that preference is revealed in conduct. While we could deny a distinction between what a person pursues and what she does, I think we should not. When talking with her overbearing mother, Amelia repeatedly and predictably says things that provoke conflict. This is part of her conduct. But she may not intend to provoke conflict; she may not desire conflict; she may not pursue conflict. Amelia does provoke conflict with her mother—and she does so as the result of her voluntary, intentional actions—but she does not pursue conflict with her mother.  

If not everything that an agent does is something the agent pursues, then Korsgaard’s argument does not show that, in order to satisfy the requirement of demandingness, the neo-Humean instrumentalist needs a distinction between an agent’s ends and what she pursues. Neo-Humean instrumentalism will be at least minimally demanding if it can draw a distinction between mere conduct and an agent’s ends, even if it can draw no distinction between an agent’s ends and what she pursues. So, let us suppose that conflict with her mother is something that Amelia not only does not pursue, it is something she desires not occur. Then, the neo-Humean instrumentalist can say that Amelia has a reason to refrain from making the inflammatory comments. Neo-Humean advice is corrective; it is demanding. This does not require a gap between Amelia’s ends and what she pursues; it is enough that there be a gap between her ends and what she does.

I doubt that this is demanding enough. But, before moving to this point, let us take stock. Provided the neo-Humean instrumentalist can distinguish between the actions required by the instrumentalist principle in service of an agent’s ends and her actual behavior, he can satisfy the requirement of demandingness, as we have explicated it. That is to say, it is simply not true, as Korsgaard says, that the sort of position we are now considering “exclude[s] the possibility that we could be guided by the instrumental principle” because “anything [we] do counts as following it” (op. cit., p. 229).

But is this demanding enough? Let us say that a principle is demanding with respect to actions if it could, under some conditions, prescribe actions that are not undertaken. A neo-Humean instrumentalism that distinguishes between an agent’s pursuits and what

---

10 In some cases of this sort, we might correctly attribute an unconscious motivation to Amelia. But such an attribution will be implausible in others. All that is required for my claim to be true is that there are some cases in which one brings about a result by one’s voluntary, intentional actions but does not pursue that result.
she actually does is demanding with respect to actions even if it cannot distinguish between an agent’s ends and what she actually pursues. For all that I have said though, such a theory cannot be demanding with respect to pursuits. And, one might plausibly claim, an adequate theory of practical rationality must be demanding with respect to pursuits, as well.

I agree. An adequate theory of practical rationality should not only allow space between those actions we are rationally committed to by our ends and those actions we actually perform. It should also allow space between those actions we are rationally committed to by our ends and those actions we are rationally committed to by our pursuits. In other words, it should allow for the rational criticism of (at least some of) our pursuits.

Can a neo-Humean instrumentalist do this? Most assuredly. For many of our pursuits do not further our ends. Remember Amelia. Let us suppose now that what she is pursuing is harmony with her overbearing mother. But she pursues this in the belief that harmonious relations will eventually lead her mother to quit being so overbearing and, as a result, their relationship will develop into a healthier one. What she wants most is to have this sort of healthy relationship with her mother. In light of this, she sets about to control her tendency to say things that provoke conflict. She pursues harmony in the relationship. In fact, though, harmonious relations with her mother will never result in a healthy relationship. Instead, the only way to bring about the sort of change Amelia most wants in the relationship is for her to stand up for herself forcefully against her mother’s intrusions in her life. Now, Amelia’s pursuit of harmony does not further her ends. What is rationally advisable given her ends is for her to stand up for herself against her mother’s intrusions.11

11 If Amelia does not believe that her ends would be better promoted by giving up the pursuit of harmony and standing up for herself, then her failure to change her pursuits does not reflect a failure in her practical rationality. If she has no reason to believe that her ends would be better served by changing her pursuits and some reason to believe that they would not, then her failure to change her pursuits does not reflect on her rationality at all. Still, the advisable course of action in light of her ends is to change her pursuits. There is no reason that a neo-Humean cannot recognize this as a legitimate and interesting evaluation of actions from Amelia’s evaluative point of view. This sort of evaluation, while it steps outside of Amelia’s doxastic and epistemic situation, does not adopt a nonsubjective evaluative perspective; it evaluates actions, as any neo-Humean theory would, from the evaluative perspective of the agent. When Amelia chooses actions, she is aiming to choose actions that do promote her ends—not actions that she believes promote her ends or actions that it would be reasonable to believe would promote her ends given her epistemic situation. When Amelia solicits advice from others about how to act, she is seeking guidance about what will further her ends, not about what it is reasonable
Here, we have a gap between the actions normatively required by Amelia’s ends and actions normatively required by at least some of what she pursues in the mistaken belief that it will secure her ends. Amelia should have different pursuits.

How can this distinction between ends and pursuits be drawn? One plausible thought is that a person’s ends are those states of affairs the person values in themselves. Pursuits include everything that the person sets about to make true. Neo-Humean instrumentalism can be demanding with respect to pursuits by evaluating them in light of what the agent values in itself. Perhaps even this is not demanding enough. Some would hold that an adequate normative theory of practical rationality must be demanding with respect to an agent’s ends.

Before evaluating the plausibility of this requirement and the ability, if any, of the neo-Humean instrumentalist to meet it, it is worth noting that the ground has shifted. We have, by now, moved from an almost completely uncontroversial criterion of adequacy for a normative theory of practical rationality—nontriviality—to one that is very nearly question begging. Many neo-Humean instrumentalists would claim that it is precisely a tenet of their theory that an agent’s ultimate ends—the things she values in themselves—are not subject to rational appraisal. Insistence that a normative theory of practical rationality allow us to subject an agent’s ultimate ends to rational criticism does not provide a criterion of adequacy for such a theory; it merely states a rejection of one prominent family of theories.

While I do not find the requirement that a normative theory of practical rationality be demanding with respect to ends a very useful one in resolving disagreement about these theories, I do think the neo-Humean can meet even this requirement, at least as the requirement has been formulated. There may, after all, be straightforward instrumental reasons for acquiring or extinguishing ultimate ends. Amelia desires, for its own sake, to have a healthy relationship with her mother. Achieving this end, however, would require Amelia to devote so much of her attention and effort to this task that she would not be able to achieve any of the other ends she has. Now, by itself, this does not “speak against” the end of having a healthy relationship with her mother. Such an end is simply one that Amelia cannot afford—or is unwilling—to pay for, given the costs involved. Amelia to believe will do so given her beliefs or given beliefs she should have in her epistemic situation. I discuss these issues in “Irrational Desires,” Philosophical Studies, lxii (1991): 23-44, especially p. 25.
finds, however, that so long as she intrinsically values (has as an end) a healthy relationship with her mother, she keeps devoting time and effort to bringing about this state of affairs, even though she knows it is hopeless. She “chuse[s] means insufficient for the design’d end.” She would achieve those of her ends that she can achieve more fully if she did not intrinsically value a healthy relationship with her mother. In this case, instrumental rationality, as understood by the neo-Humean, recommends that Amelia extinguish (to the degree that she can without incurring too high a cost) her intrinsic desire to have a healthy relationship with her mother. Ultimate ends can be instrumentally irrational in the sense that one has reason to extinguish them. In this sense, the neo-Humean can subject even an agent’s ultimate ends to rational scrutiny.\(^\text{12}\) What the neo-Humean instrumentalist is unable to do is to find an agent’s ultimate ends intrinsically rational or irrational.

Let us call a principle *instrumentally demanding with respect to ends* when it holds, based partly on instrumental considerations, that an agent’s ultimate ends may be different from those she should, rationally, have. A theory that, based solely on intrinsic considerations, holds that an agent’s ends may not be those she should rationally have is *intrinsically demanding with respect to ends*. Consider these properties as putative criteria of adequacy for a normative theory of rational choice. Now, my claim is that even neo-Humean instrumentalism satisfies the first proposed criterion. And, further, I claim that the second criterion is completely question-begging. It asks us to accept as a criterion of adequacy of a normative theory of practical rationality that a central tenet of the neo-Humean theory be denied. Hardly an Archimedean leverage point!

The Neo-Humean instrumentalist theory, unmodified, satisfies the noncontroversial requirements of demandingness provided it is not hobbled with a very implausible and outdated behavioristic account of an agent’s ends. Indeed, the theory can satisfy far more stringent and controversial requirements of demandingness—including demandingness with respect to pursuits and instrumental demandingness with respect to ends. So, what is the problem?

V. KORSGAARD’S DILEMMA

Korsgaard believes that neo-Humean instrumentalists face two dilemmas, one nested in the other. The first is between normative triviality and distinguishing between an agent’s ends and what she actually

\(^{12}\) See my “Irrational Desires” (pp. 23–26) for a fuller discussion of this issue with a somewhat different focus.
does. What I have argued is that the neo-Humean instrumentalist (and, indeed, I think Hume, himself), rather obviously and easily, takes the second horn. But, on the tip of this horn waits another dilemma. Korsgaard presents this second dilemma by offering the neo-Humean instrumentalist two ways to draw the distinction necessary for avoiding normative triviality. (But Humeans should beware of Kantians bearing gifts.)

Here is how Korsgaard presents the second dilemma:

[T]he problem would be solved if we could make a distinction between a person’s ends and what he actually pursues. Two ways suggest themselves: we could make a distinction between actual desire and rational desire, and say that a person’s ends are not merely what he wants, but what he has reason to want. Or, we could make a more psychological distinction between what a person thinks he wants or locally wants and what he “really wants” (op. cit., p. 230).

The first approach seems, rather obviously, to grant a role to rational appraisal that defenders of instrumentalist conceptions of practical rationality have rejected; it seems to entail that there can be intrinsic rational appraisal of an agent’s ultimate aims. But what I have offered above in defending the demandingness of neo-Humean instrumentalism can be seen as a version of Korsgaard’s second approach to avoiding normative triviality. And, I believe, the sort of theory I sketched is what most neo-Humeans have intended all along. So, what sharp point does Korsgaard have for those who take the second horn of her second dilemma?

She says:

[T]he second option...also takes us beyond instrumental rationality, although this may not be immediately obvious. If we are going to appeal to “real” desires as a basis for making claims about whether people are acting rationally or not, we will have to argue that a person ought to pursue what he really wants rather than what he is in fact going to pursue. That is, we will have to accord to these “real” desires some normative force. It must be something like a requirement of reason that you should do what you “really want,” even when you are tempted not to. And then, again, we will have gone beyond instrumental rationality after all (op. cit., p. 230).

This charge is initially puzzling. The alleged step beyond instrumental rationality—the assertion that “a person ought to pursue what he really wants”—is just a reassertion of the neo-Humean instrumental
principle. As David Lewis\textsuperscript{13} says in another context, “what is already there cannot be added” (ibid., p. 534); one does not take a step beyond neo-Humean instrumentalism by repeating the thesis of neo-Humean instrumentalism.

But Korsgaard’s complaint is that the neo-Humean instrumentalist, no less than her opponents, is committed to a substantive principle of rationality which generates reasons that are not contingent on the agent’s aims and desires—one which is, in this sense, “categorical.” The emphasis by neo-Humean instrumentalists on means/ends reasoning focuses attention on what I have elsewhere called (following Shelly Kagan\textsuperscript{14}) pure instrumentalism—the thesis that reasons are communicated across causal, criterial, and mereological relations. And this thesis is noncontroversial. But neo-Humean instrumentalism is also defined by what I follow Stephen Darwall\textsuperscript{15} in calling the desire-based reasons thesis: reasons (in the sense relevant to the rational advisability of action and the appraisal of agents as rational) are grounded, ultimately, in the subjective, contingent, conative states of the agent—in her intrinsic valuings, for example.\textsuperscript{16} Of course, one can, and many do, accept pure instrumentalism without accepting the desire-based reasons thesis. The desire-based reasons thesis is not an instrumental principle and so, neo-Humean instrumentalism (which includes both principles) “goes beyond” a purely instrumental principle.

Roughly, the idea seems to be this. Everyone should admit that a person has reason to undertake the means to those ends she has reason to bring about. That is just the thesis of pure instrumentalism. But the desire-based reasons thesis asserts that there is a reason for an agent to bring about the ends that she has (what she intrinsically values, for example). And this reason does not, and cannot, derive from these things being necessary means to our ends.

\textsuperscript{13} “Levi against U-Maximization,” this JOURNAL, LXX, 9 (September 1983): 531-34.
\textsuperscript{14} “On Instrumental Reasons,” presented at the American Philosophical Association, Central Division Meeting, April 29, 1989.
\textsuperscript{15} Impartial Reason (Ithaca: Cornell, 1983), p. 27.
\textsuperscript{16} That an individual’s intrinsic valuings are subjective, contingent, conative states of the agent is not completely uncontroversial. I develop and defend the claim that they are such states in “Values and Desires” (typescript). While I favor a form of neo-Humean instrumentalism that grounds an agent’s reasons on her intrinsic valuings, for present purposes this can be treated as a mere example. Those who challenge the subjective, contingent, conative status of an individual’s intrinsic valuings may substitute another, less controversial, mental state such as preference or desire.
Korsgaard confuses the reader about her real complaint by suggesting that the feature she alludes to here arises because the neo-Humean makes a distinction between an agent’s ends and what he is going to do. If I am right that her concern is that the principle of pure instrumentalism is only about the transmission of reasons and some further normative principle is needed to ground instrumental reasoning, the “further normative principle” would be just as necessary for those who draw no such distinction. Let us call this the grounding problem.

VI. QUINN’S QUESTION

Korsgaard is not alone in worrying about whether neo-Humean instrumentalism can yield reasons that are grounded in a way that is necessary if they are to be normative. Warren Quinn (op. cit.) shares the concern.

Suppose, with Quinn, we take neo-Humean instrumentalist theories to attempt to ground judgments of practical rationality ultimately on the agent’s pro- and con-attitudes and that we understand these to be “functional states that dispose us to act.” Quinn asks “whether pro- and con-attitudes conceived as functional states that dispose us to act have any power to rationalize those acts” (op. cit., p. 236). To answer this question, Quinn imagines a peculiar functional state that disposes its possessor to turn on radios she sees to be turned off. And he says: “I cannot see how this bizarre functional state in itself gives me even a prima facie reason to turn on radios, even those I can see to be available for cost-free on-turning” (op. cit., p. 237).

An answer to Quinn will wait. Here, I want to accept his implicit criterion for a normative theory of practical rationality: conformance with the principle must seem (at least on reflection) to rationalize the actions.

Quinn’s concern, remember, was how “pro- and con-attitudes conceived as functional states that dispose us to act have any power to rationalize those acts.” He did not see how a mere disposition to perform an action could even tend to show that the agent had a reason to perform that action.

I think there is a fairly obvious response to Quinn’s concern. It is this: if pro- and con-attitudes are conceived only as functional states that dispose us to act, they do not have any power to rationalize those actions. Quinn is right to say that a mere disposition to turn on radios in one’s vicinity does not give one “even a prima facie reason to turn on radios.” But a mere disposition to perform an action is clearly not constitutive of having a pro-attitude toward that action. A person’s disposition to fall asleep at department meetings, get tongue-tied
when talking with a celebrity, or display anger when talking to an ex-spouse hardly shows that the person intrinsically values these actions or any state of affairs to which she believes them to be instrumental. For all Quinn’s examples show, the problem is not that certain pro-attitudes cannot rationalize actions but, rather, that these pro-attitudes cannot be adequately understood as merely dispositions to act or be motivated to act—an analysis that, to my knowledge, has not been proffered by any defender of a neo-Humean instrumentalist position.

Interestingly, in setting up the problem—just a paragraph before he actually raises it—Quinn recognizes three separate times that the neo-Humean instrumentalist thinks of the functional state of disposing one to perform an action as only one (if a central one), among several, elements of a pro-attitude. And, so, the question naturally arises why Quinn thinks he has raised a difficult problem for the neo-Humean. Does not the question of whether a functional state that disposes us to act could have any power to rationalize the act obviously depend on what other characteristics that functional state has?

What follows is quite speculative, of course, but I suspect that Quinn thought that nothing one could plausibly add to the mere disposition to act in order to get a full description of a pro-attitude would do any more to rationalize the act. Infuse a pro-attitude with as much functional meaning as you want—inject into it as much qualitative content as you will—pro-attitudes will not rationalize actions. They cannot ground reasons for acting. For example, treat a pro-attitude toward some state of affairs not only as including a disposition to bring about that state of affairs, but also a disposition to feel regret at missed opportunities to do so and a positive affective mental state toward that state of affairs. We can still wonder whether there is anything in all of this that rationalizes the act in question—whether the attitude toward the state of affairs in question grounds a reason to bring it about.

If this is the thought that lies behind Quinn’s concern, then I think it is just the worry Korsgaard has. Transmitting reasons from ends to means is noncontroversial. But how does the neo-Humean instrumentalist “get reasons going” in the first place? It must be by unwittingly and illicitly appealing to a noninstrumental normative principle to ground practical reason—something like: one ought, rationally, to bring about one’s ends.
VII. NORMATIVITY ON A NON-NORMATIVE FOUNDATION

The problem. The worry, then, is fairly simply put. The principle of pure instrumentalism ensures the transmission of reasons from ends to means, but the neo-Humean, no less than anyone else, needs a substantive and noninstrumental principle of rationality to give rational endorsement to the ends—to ground the chain of reasons. No reasons will be transmitted to the means unless there is a reason for the ends. This objection really has nothing to do with demand-  

Now, it might be thought, we have a puzzle for the neo-Humean instrumentalist. (If I am right, this is really what lies behind the concerns about normative triviality.) Suppose he makes the following claims:

(a) Whatever is endorsed by practical reason is so endorsed by virtue of being instrumental to an agent’s ends.
(b) Ends cannot convey reasons unless there are reasons for the ends.
(c) Ends are beyond intrinsic rational evaluation.

He is committed, then, to the conclusion that nothing can be endorsed by practical reason. If, on the other hand, he claims that practical reason endorses an agent’s ends (no matter how defined), then, no less than any other theory of practical rationality, he owes us a justification of this clearly noninstrumental principle.

How might the neo-Humean instrumentalist reply? My answer will begin far from theories of practical rationality, but not quite as far away as the story of The Little Prince took us.

Excursus into legal philosophy. Through the middle of the twentieth century, there was a debate among legal positivists. Unlike their natural-law opponents, legal positivists saw law as being solely a product of human activity. Natural-law theorists saw legal validity as requiring some sort of conformance of human dictates with the requirements of a natural law, which exists independently of human activity. Positivists denied this. Instead of subjecting statutes and orders to an external test of legal validity, they employed a “pedigree test” of legal validity: a statute is valid law, for example, if it is validly derived from (or enacted in accordance with) a more basic valid law. Of course, the pedigree cannot go back forever. To avoid regress, this chain has to end in a fundamental law.

Hans Kelsen called this fundamental law the “Grundnorm.” The Grundnorm exists because it is actually obeyed. But existence is not validity. So, the question arises, is this fundamental law valid or not?

If not, how could it confer validity on the “derived” laws? If so, from whence did its validity come? The pedigree test gives only a mechanism to “transmit” validity to subordinate norms. To confer validity on the Grundnorm would be to “go beyond” the pedigree test. But, if there could be an independent source of legal validity, why does it apply (only) to the fundamental norm that is actually effective in society? Why does this source not confer validity directly on each particular law? Why does it not confer validity on some other basic norm that satisfies a requirement different from social effectiveness? Why not to a norm that is morally demanded? Or to a norm that every rational being would will to be a universal law of nature insofar as reason had control of his will?

Kelsen’s answer was that the validity of the Grundnorm is “presupposed” (ibid., pp. 201-05)—an answer which, to quote Bertrand Russell18 speaking of another matter, has all “the advantages of theft over honest toil” (ibid., p. 71). Kelsen’s “answer” addresses none of the concerns we just raised about the validity of the Grundnorm, and so, as Russell urges his readers, “Let us...proceed with our honest toil” (ibid.).

H. L. A. Hart19 shared Kelsen’s legal positivism, but offered a different account of the validity of “derived” laws. They did not have to receive their validity by being “derived” from a valid law—as if validity poured from the basic rules to the derived rules through the conduit of valid derivation. Rather, the basic rule of law, which Hart called “the ultimate rule of recognition,” defines “validity” within the legal system in question. Given this role, the question of the validity of the ultimate rule of recognition does not arise. It involves a sort of “category mistake” to think of the ultimate rule of recognition as being valid or invalid. Hart says:

We need the word ‘validity’, and commonly only use it, to answer questions which arise within a system of rules where the status of a rule as a member of the system depends on its satisfying certain criteria provided by the rule of recognition. No such question can arise as to the validity of the very rule of recognition which provides the criteria; it can neither be valid nor invalid but is simply accepted as appropriate for use in this way. To express this simple fact by saying darkly that its validity is ‘assumed but cannot be demonstrated’, is like saying that we assume, but can never demonstrate, that the standard metre bar in Paris which is the ultimate test of the correctness of all measurement in metres, is itself correct (ibid., pp. 105-06).

The result is that rules of law may be valid in virtue of their relation to a fundamental rule, even though this fundamental rule is neither valid nor invalid. Its status as law is secured by a social fact: its general acceptance in a population, among other things. Unlike other rules of law, the ultimate rule of recognition is not law because it is valid; it is law because it is accepted. This is problematic if one thinks of the resultant validity, legal validity, as unconditional—if one thinks of the normativity of law as being unqualified. A classical natural-law theorist, for example, will worry about how a fact about social acceptance can ground the validity of a norm of law.

But, on Hart’s view, laws are valid only relative to a system of law. To call a statute, for example, valid law is to say that it is properly created within some (usually implicitly indicated) legal system. This is what Kelsen misses. He has only half digested the implications of legal positivism. Recognizing that law is a product solely of human activity, he still looks for a concept of validity that can be applied from some external perspective so as not to make validity a “system bound” concept. He “presupposes” the validity of the Grundnorm because he does not fully appreciate that all we seek to say when we talk of the validity of a law is its conformance with the basic norms of the legal system. That is what “legal validity” means.

Application to reasons for acting. Now, what application do the lessons of this excursion into the recent history of legal philosophy have for the normativity of instrumental reasons and the viability of neo-Humean instrumentalism as a theory of practical rationality? As a transition, let us begin by talking about legal reasons for acting.

When we say that a person has a legal reason to perform an action, we justify this by appeal to, for example, specific statutes or court orders that the person has a legal reason to obey. If called on to justify the claim that the person has a legal reason to obey this specific statute or court order, we will cite more basic legal norms which, we argue, the person has a legal reason to obey and which endorse this statute or court decision. But the (legal) reason giving comes to an end at some point. And, in Hart’s view, that point is at a complex social fact about rules that are in effect and the attitudes people have toward those rules.

To be sure, this makes legal reasons much less impressive than they might otherwise seem. One might very well recognize that he, or others, have a legal reason to do something but correctly think this conclusion is completely unimportant in deciding what to do. And, this is part of the message of legal positivism, too. Making legal validity and the existence of legal reasons depend on a logically and morally arbitrary social fact is unacceptable if we want claims of legal
validity and what one has legal reasons to do to have all of the
grandeur that the natural-law theorist invests in them. But this is a
reason to reject a hybrid theory that takes the account of legal validity
and legal reasons from the positivist and the significance of legal
validity and legal reasons from the natural-law theorist. It is not a
mark against the legal positivist.

Finally, now, to apply these insights to the issue of the normativity
of instrumental reason and the adequacy of neo-Humean instrument-
alism, let me begin by noting a rich and important dissimilarity
between the two issues. Recall that Hart says: “we need the word
‘validity’, and commonly only use it, to answer questions which arise
within a system of rules.” We can say no such thing about the concept
of a reason for acting. Indeed, part of the problem with reconciling
the competing theories of reasons for acting is that we have in mind
too many different ideas and too many different uses for the con-
cept.20 But we can say this: one thing we seek to do when talking
about reasons for an agent to act is to evaluate actions from within the

20 When we talk about reasons for acting, we are often concerned with the actual
reasoning process employed by the agent—not the issue of what action is rationally
advisable from any standpoint. This is, of course, parallel to discussions of theoret-
ical rationality. Our concerns are often with the agent’s belief-forming processes
and the evaluation of an agent’s dispositions to believe—not with the question of
what beliefs are rationally recommended. Similarly, when we talk about reasons for
acting, we are often concerned with the agent’s intention-forming processes and
the evaluation of the agent’s dispositions to act.

But, even where we are concerned with the evaluation of the actions themselves,
we have a variety of interests. At times, our interest is in endorsing an action sans
phrase. With respect to such evaluation, I see no reason to assume the evaluative
standpoint of the agent. We are also concerned, however, to evaluate actions from
the evaluative standpoint of the agent. Even when we do this, we have a variety of
concerns.

Sometimes we are concerned to evaluate actions in relation to the agent’s desires
(preferences, values, or some such state) and the agent’s actual beliefs. We want to
know if the action is the one that would be rationally advisable if the world were as
the agent believed it to be—regardless of whether the agent’s beliefs were true, or
even reasonable. An agent who performs actions that would not achieve her ends
were the world as she believes it to be makes a special kind of practical error.

Sometimes we are concerned to evaluate actions in relation to the agent’s desires
(preferences, values, and the like) and the beliefs it would be reasonable for the
agent to have. Here, we seek to determine whether the action was rationally
advisable given the epistemic situation in which the agent found herself. Agents
who perform actions that would not achieve their ends were the world the way it is
rational for them to believe it to be, may be making the same practical mistake
noted above, or they may be making an epistemic mistake that has practical
implications for them.

Sometimes, I contend, we are concerned to evaluate action in relation to the
agent’s desires (preferences, values, and the like) and the actual state of the world.
Here we are concerned with the advisability of the action from the agent’s evalua-
tive standpoint. This is what I call rational advisability.
agent's system of norms—in relation to what the agent intrinsically values; one thing we seek to do when we talk about the practical rationality of an agent is to evaluate the coherence of her pursuits and intentions from within the agent’s system of norms—in relation to the agent’s values. It is this sense of there being a reason for an agent to perform an action (of the action being rationally advisable), and this sense of an agent being practically rational, that the neo-Humean instrumentalist seeks to explicate.

Of course, the neo-Humean instrumentalist must engage in the same “lowering of expectations” that the legal positivist must. Saying that there is a reason in this sense for an agent to perform an action does not mean anything more than that this action best promotes the agent’s ends. It certainly does not, by itself, support what Allan Gibbard\(^2\) has called “a flat, flavorless endorsement” (ibid., p. 7).

Ought people to take efficient means to their ends? If the ‘ought’ is one of rational advisability, the neo-Humean instrumentalist says, then, (trivially) they ought. If the ‘ought’ is an expression of the speaker’s acceptance of norms that recommend his performing the action, there is no reason to think that he ought. I would, for example, urge those whose fundamental values are those of the National Socialist Party to act contrary to the reasons they have in the neo-Humean instrumentalist sense of reasons.

VIII. NEO-HUMEANISM UNDERSTOOD ARIGHT

Neo-Humeans do not believe—or, at any rate, ought not to believe—that practical rationality is an objective matter. To be sure, there may be a fact of the matter about what action is rationally recommended for an agent. This fact, though, depends crucially on subjective states of that agent. On my view, it is the agent’s values that play this role in determining what is rationally advisable for the agent. Other neo-Humeans focus on desires or preferences. This is an internecine dispute; it is not important here. All neo-Humeans take the rational advisability to depend, ultimately, on subjective, contingent, conative states of the agent in questions.

But no neo-Humean should think that, at bottom, actions are rationally advisable because they are instruments to rationally advisable ends (aims, projects, the satisfaction of desires or preferences, or the promotion of values). Certainly, actions can be rationally advisable for this reason, but this cannot be what rational advisability is at bottom. For at bottom, there is a fact—a brute fact—about the

---

agent's subjective, contingent, conative states. And the critics of neo-Humeanism are right to say that, on the neo-Humean view, this fact is not intrinsically rationally appraisable. Their mistake, the neo-Humean contends, is in thinking that it must be.

Critics of the neo-Humean position have implicitly accepted a view of rational advisability which is analogous to the view of legal validity which was accepted by critics of legal positivism. This view holds that, while an action’s rational advisability might depend on the action having a proper relation to some end or goal, this relationship could confer rational advisability on the action only if the end or goal were rationally advisable. Similarly, the critic of legal positivism holds that, while a rule's legal validity might well be dependent on its being validly enacted in accordance with some more basic legal rule, this relationship could confer validity only if the more basic rule were, itself, legally valid. At first blush, this seems eminently plausible—indeed, almost inescapable. But escape it we must if we are to be neo-Humeans.

Kelsen, for all his brilliance in developing the modern form of legal positivism, could not shake the intuition that the basic rule of law had to be legally valid in order for other legal rules to be legally valid in virtue of some relation to it. Because his conception of legal validity was still clouded by intuitions based on a natural-law tradition, he was forced to resort to presupposing the legal validity of the basic rule of law. The Kelsenian picture must be rejected if we are to understand neo-Humeanism aright.

The neo-Humean adopts a view of how an agent’s fundamental ends determine the rational advisability of her actions that is more similar to Hart’s view of how the basic legal rule determines the legal validity of other rules of a legal system. For the neo-Humean, the agent’s ultimate ends (I would say, “her intrinsic values”) are neither rationally advisable nor rationally inadvisable, in themselves. They are, rather, the brute facts about the agent’s psychology in virtue of a relationship to which policies, plans, and actions can be rationally advisable or inadvisable. The agent’s ultimate ends confer this status on policies, plans, and actions not because these ends have some normative standing in themselves. They do this because the particular property of being rationally advisable just is the property of being properly related to these brute facts (just as the property of being legally required just is the property of being properly related to some complex brute fact about a society).

The theory of rational advisability offered by the neo-Humean does not give rational advisability all the status some would want in such a concept. It certainly does not proffer a simple, unqualified answer to
the question of what to do. That is fine. We advise people from many different perspectives. Sometimes our advice explicitly indicates the perspective; sometimes the indication is contextual. Sometimes our advice is based on norms that we, ourselves, adopt; sometimes it is based on norms we apply without endorsement. Rational advisability is advisability from just one perspective. Because this perspective is constructed from the agent’s own values, it is one to which the agent has a special relationship. Still, there are, as the critic of the neo-Humean position wants to insist, many ways in which we might want to evaluate actions other than from the normative perspective of the agent. The neo-Humean should not deny this. All she must deny is that these other sorts of evaluation constitute evaluations of the rational advisability of actions.

IX. CONCLUSION

On the neo-Humean’s view, judgments of rational advisability do not address the question of how to act, simpliciter. Still, they address questions which are important and in which we are interested. They address the questions of what actions are recommended by an agent’s ends and when an agent has the intentions and pursuits appropriate to her ends. One is free to think these questions uninteresting. Disinterest cannot be refuted. But, for those interested in these questions, the neo-Humean instrumentalist offers answers that do not require anything “beyond” instrumentalism. It accomplishes this by construing instrumental rationality as having a sort of “groundless normativity.”

DONALD C. HUBIN

Ohio State University

22 I defend the claim that the neo-Humean can defend a special motivational status for claims of rational advisability in “What’s Special about Humeanism,” Noûs, xxxiii, 1 (1999): 30-45.