



THE HUMEAN THEORY OF PRACTICAL IRRATIONALITY

BY NEIL SINHABABU

JOURNAL OF ETHICS & SOCIAL PHILOSOPHY

VOL. 6, No. 1 | NOVEMBER 2011

URL: WWW.JESP.ORG

COPYRIGHT © NEIL SINHABABU 2011

The Humean Theory of Practical Irrationality

Neil Sinhababu

IN “THE NORMATIVITY OF INSTRUMENTAL REASON,” Christine Korsgaard presents a problem for those who accept similarly structured Humean views of both action and rationality.¹ (I will call the conjunction of views she criticizes the double-Humean view.) Korsgaard contends that the double-Humean view implies the impossibility of irrational action, as it claims that we can only perform the actions that it deems rational. First I will develop Humean views of rationality and action so as to display the force of Korsgaard’s objection. Then I will respond by showing how double-Humeans can develop their view to account for just as much practical irrationality as there is.

Double-Humeans can answer Korsgaard’s objection if their views of action and rationality measure agents’ actual desires differently. What determines what the agent does should be the motivational forces that desires produce in the agent at the moment when she decides to act. That is when her desires play their causal role in determining action. What determines what it is rational to do should be the agent’s dispositional desire strengths. Our normative intuitions about rationality concern these states. Since the action that desire motivates us most strongly to do at the moment of action may not be the action that would best satisfy our dispositional desires, irrational action is possible. This way of filling out the double-Humean view is true to the best reasons for accepting both of the theses that compose it, and lets us better understand the nature of irrational action.

1. The Double-Humean View

Humeans about motivation hold that desire is necessary for motivation, while Humeans about reasons hold that our reasons for action are grounded in our desires. These theories have not always been elaborated so as to clearly answer questions about which action someone will perform or which action it is rational to perform. I will elaborate Humean theories about motivation and reasons in a way that answers these questions, as doing so helps to display the force of Korsgaard’s objection.

According to the Humean theory of motivation, the only states of mind necessary for motivating action are a desire and a belief that an action will bring about the desired state.² This is a nonnormative theory describing the

¹ Christine Korsgaard, “The Normativity of Instrumental Reason,” in G. Cullity and B. Gaut, eds., *Ethics and Practical Reason* (Oxford: Clarendon Press, 1997).

² For similar formulations of the Humean theory of motivation, see: Michael Smith, *The Moral Problem* (Oxford: Blackwell, 1994); Ralph Wedgwood, “Practical Reason and Desire,” *Australasian Journal of Philosophy* 80 (2002): 345-58; Neil Sinhababu, “The Humean Theory of Motivation Reformulated and Defended,” *Philosophical Review* 118 (2009): 465-500.

psychological states that cause action. If we add up the motivational forces favoring or disfavoring each action, each force being the product of a desire with a particular strength and a belief about how acting will affect the probability of desire satisfaction, we arrive at the following theory about what agents will do:

Humean theory of action: Agents do whatever maximizes expected desire satisfaction.³

The notion of “expected desire satisfaction” here is familiar from decision theory, and can be characterized as a subjective-probability-weighted average of the extent to which various outcomes would satisfy the agent’s desires.

According to the Humean theory of practical reason, agents have reason to act insofar as the action promotes the satisfaction of their desires.⁴ This is a normative theory in which reasons give actions *pro tanto* justification. Weighing and adding up the reasons as we previously did with the motivational forces, in terms of desire strengths and changes in probabilities of desire satisfaction, we can produce the following theory about what it is rational for agents to do:

Humean theory of rationality: It is rational for agents to do whatever maximizes expected desire satisfaction.

Henceforth I will take double-Humeans to be committed to the conjunction of the Humean theory of action and the Humean theory of rationality. I beg the indulgence of readers who see better options for Humeans regarding what agents will do and which action is rational, as my purpose in developing Humean views on these two questions is mainly to demonstrate the force of Korsgaard’s objection. As long as one’s answer to “What will agents do?” is the same as “What is it rational for agents to do?” – as is the case with the two Humean theories laid out above – the objection will apply.

2. Korsgaard’s Objection

Korsgaard provides an instance of a general argument against accepting accounts of motivation and rationality as neatly matched as the two

³ This and the next formulation treat inaction as a way of doing something.

⁴ For similar formulations of the Humean theory of rationality, see: Smith 1994; Wedgwood 2002; Mark Schroeder, *Slaves of the Passions* (Oxford: Oxford University Press, 2007). I formulate both Humean theories in terms of “expected desire satisfaction” rather than “what would maximize desire satisfaction if the agent’s beliefs were true,” to better handle cases where (for example) one pays \$1 for a 1 percent chance of getting \$500. Expected value formulations rightly present this gamble as rational. But formulations involving the truth of one’s beliefs might not present the gamble as rational, since one does not believe that one will get \$500.

components of the double-Humean view. She objects that such accounts make practical irrationality impossible:

The problem is coming from the fact that Hume identifies a person's end as what he wants most, and the criterion of what the person wants most appears to be what he actually does. The person's ends are taken to be revealed in his conduct. If we don't make a distinction between what a person's end is and what he actually pursues, it will be impossible to find a case in which he violates the instrumental principle (1997: 230).

Korsgaard spells out the "instrumental principle" as follows in a later work: "[P]ractical rationality requires us to take the means to our ends."⁵ She adds that "some philosophers regard the ends in question as things desired" (2008: 5). This group would seem to include Hume and contemporary Humeans about reasons. One might understand the Humean theory of rationality as an elaboration of the instrumental principle that delivers precise answers as to which action is rational.⁶ Treating the performance of an action as a means to satisfying our desires, an account of rationality based on the instrumental principle would present the action that provides the greatest expected desire satisfaction as the rational one.

Korsgaard's objection to the double-Humean view is that it makes practical irrationality impossible by giving matching accounts of rationality and motivation. If a theory of motivation says that one will perform a particular action, and a theory of rationality says that it is rational for one to perform that action, the conjunction of the theories will imply that one acts rationally. According to Korsgaard, Hume not only says that "people don't in fact ever violate the instrumental principle. He is actually committed to the view that people cannot violate it" (1997: 228). To avoid this consequence, we might reject at least one component of the double-Humean view.

The sort of practical irrationality the double-Humean needs to explain is not merely a matter of doing the wrong thing because of false belief, like drinking gasoline because you mistakenly believe that it is gin.⁷ In these cases, we might say that there was a reason for you to do otherwise, but that by itself does not make you irrational. As Korsgaard says, "the possibility of mistake is not in general very interesting" (1997: 227). She explains why in an earlier paper: "judgments of irrationality, whether of belief or action, are, strictly speaking, relative to the subject's beliefs. Conclusions drawn from

⁵ Christine Korsgaard, *The Constitution of Agency* (Oxford: Oxford University Press, 2008), p. 5.

⁶ As stated, the instrumental principle does not tell us how we ought to choose between two incompatible ends, or how to incorporate differing subjective probabilities of satisfaction. The Humean theory of rationality answers these questions.

⁷ Humeans about rationality are well positioned to explain how agents can do things they lack reasons to do because of mistaken belief. This is different from accounting for irrationality. See David Sobel, "Subjective Accounts of Reasons for Action," *Ethics* 111 (2001): 461-92, which deals with mistaken belief in a comprehensive way.

mistaken premises are not *irrational*.”⁸ One need not develop the double-Humean view in any especially sophisticated way in order to account for the failure to do what there is most reason for you to do, at least when we are talking about the kind of reasons that are tied to actual (and not expected) desire satisfaction.

The trouble is to account for the sort of failure to act rationally that can occur even when the agent has no relevant false beliefs. Korsgaard offers an example:

Howard, who is in his thirties, needs medical treatment: specifically, he must have a course of injections, now, if he is going to live past fifty. But Howard declines to have this treatment, because he has a horror of injections. Let me just stipulate that, were it not for his horror of injections, Howard would have the treatment. It's not that he really secretly wants to die young anyway, or anything fancy like that. Howard's horror of injections is really what is motivating him. ... [L]et's again stipulate that he has not miscalculated or made a mistake. He sees that, if he were governed by considerations of prudence, he would have the injections: he agrees that a long and happy life is a greater good than avoiding the injections. But he still declines to have them: he chooses "his own acknowledged lesser good" (1997: 227).

Howard knows the consequences of his actions, even as he rejects the injections that will save his life. In refusing the injections, Howard acts irrationally. If double-Humeans cannot account for his irrationality or the irrational behavior of agents in so many similar cases, their theory will be deeply implausible. The rest of this paper will develop both sides of the double-Humean theory so as to account for the irrationality of Howard and people like him.

3. The Humean Theory of Action and Immediate Motivational Forces

In filling out the Humean theory of action, we should bear in mind how the motivational force our desires produce can vary. For example, a desire's motivational force can be amplified by vivid sensory or imaginative representations of its object. So when we are determining what the agent will do in a given situation, we should look at the motivational forces her desires produce after the vividness of these representations has affected them.

Hume's psychological insights help us develop the theories that he inspired about motivation and action. After distinguishing between calm and violent passions, he describes how vivid images of desired objects affect an agent's motivational states. He notes that by "varying the situation of the object" we can "change the calm and violent passions into one another" (2.3.4).⁹ For example, one can change a calm passion into a violent one by

⁸ Christine Korsgaard, "Skepticism about Practical Reason," *Journal of Philosophy* 83 (1986): 5-25.

⁹ David Hume, *A Treatise of Human Nature*, 2nd edition, L. A. Selby-Bigge and P. H. Nidditch,

bringing its object closer to the agent. Hume explains why “The same good, when near, will cause a violent passion, which, when remote, produces only a calm one” (2.3.4) – because we imagine nearby goods more vividly:

There is an easy reason, why every thing contiguous to us, either in space or time, shou’d be conceiv’d with a peculiar force and vivacity, and excel every other object, in its influence on the imagination. Ourself is intimately present to us, and whatever is related to self must partake of that quality. But where an object is so far remov’d as to have lost the advantage of this relation ... its idea becomes still fainter and more obscure (2:2:7).

The closer something is to us in space or time, the more vividly we will imagine it, and the more violent our passions for it will be. An increase in violence consists in a magnification of the passion’s effects, including its motivational force as well as its felt emotional intensity.¹⁰ Hume explains our tendency to discount future goods in terms of the greater motivational force produced by our more violent passions for nearby goods, which he explains in terms of their greater imaginative vividness.

Vivid sensory or imaginative representations will increase the violence of a desire. Because of this, the motivational force a desire can produce may vary dramatically over time, even without any permanent changes in the agent’s psychological dispositions.¹¹ My desire for a chocolaty dessert may increase in violence as I see the delicious thing brought out on the dessert tray. My desire to lie around in bed is more violent when I feel the physical comforts of my warm bed all around me than it is after I rise. Howard, in Korsgaard’s case, likely imagines the injections (which as Hume would note, are nearer to him in space and time) much more vividly than the distant benefits of being a happy 60-year-old, and his aversion to the injections thus becomes violent.

When agents are presented with unequally vivid representations of the objects of their desires, they may act irrationally. Perhaps my desire for chocolate, intensified by the sight of the delicious dessert, will cause me to eat it, even though it wrecks my diet. When I lie too long in bed, the comforts of bed are much more vivid to me than the benefits of getting work done that day. Howard, with his aversion to injections intensified by his

eds. (Oxford: Clarendon Press, 1978).

¹⁰ Calm passions can still overpower violent passions and determine action. The motivational strength of a passion will rise when it becomes violent, but may still remain below that of a much stronger passion that remains calm. Then the calm passion can overpower the violent passion. This is what happens in resolute action.

¹¹ I invoke immediate motivational force rather than simply using the occurrent/dispositional distinction and talking about the “occurrent strength” of desire because desire has many functional outputs (motivating action, producing pleasure and displeasure, and directing attention), some of which are more affected than others by vivid representations. Thus I cannot appeal to a unified notion of “occurrent strength” covering all these outputs. See Sinhababu 2009 for more discussion of how vivid representations affect desire.

proximity to them, has greater motivation to pursue what he acknowledges as his lesser good. In these cases, the benefits of irrational action are presented more vividly than the costs, so agents' desires for the benefits flare up out of proportion to their desires to avoid the costs. This makes the agents act irrationally and incur costs that they would have avoided if they had considered everything with equally vividness.

Filling out the Humean theory of action so that "desire" is understood as "desire, accounting for the motivational effects of the agent's sensory or imaginative representations" is true to the theoretical grounds for the view. Whether because of the teleological structure of action explanations or in pursuit of the best explanation of how we think and feel as we deliberate and act, Humeans about motivation hold that desire is necessary for action, and that belief is insufficient. The proposed elaboration of the Humean theory of motivation into a view about which action we perform is true to both the letter and the spirit of this principle. Desire operates in a way that falls squarely within the Humean theory of motivation, and the states that intensify desire – vivid sensory or imaginative representations of its objects – are not beliefs.

4. The Humean Theory of Rationality and Dispositional Desire Strengths

Now I will consider the appropriate conception of desire strengths for the Humean theory of rationality. The right conception is not that of the motivational force that desires are producing at the moment of action, which can be heavily affected by what agents perceive or imagine at the time. Instead, "desire" in the Humean theory of rationality should refer to the agents' dispositional desires. Dispositional desire strengths provide a baseline to which vivid images of the objects add in generating the immediate motivational force of the desires. Humeans about rationality should use this antecedent dispositional baseline, and not the balance of motivational forces at the moment of action, to determine what it is rational to do.

Dispositional mental properties have been the subject of much work in the philosophy of mind.¹² They are a fundamental part of our ordinary mental-state attributions. For example, we can say that Aisling desires Ireland to be free more than she desires any amount of money, even if she is busy minding her store and not thinking about Irish freedom. Even when she is in

¹² For a classic treatment of mental dispositions in philosophical psychology, see Gilbert Ryle, *The Concept of Mind* (London: Hucheson, 1949). Views opposed to Ryle's, like those in D. M. Armstrong, *Belief, Truth, and Knowledge* (Cambridge: Cambridge University Press, 1973), will allow my argument to go forward as long as they allow both for occurrent desires and stable non-occurrent desires of some sort. To see the notion of dispositional desire put to good use, see David Boonin, *A Defense of Abortion* (Cambridge: Cambridge University Press, 2003).

a dreamless sleep, with no occurrent mental states at all, we can say that she desires the freedom of Ireland more than money. Her dispositional desire strengths make the claim true. Other mental states, including belief, operate similarly. We can ascribe beliefs concerning political matters to her when she is asleep or busy with other things. In doing so, we ascribe dispositional beliefs.

While an agent whose dispositional desire for A is stronger than her dispositional desire for B will usually be disposed to choose A over B, this will not always be the case. We should not confuse having a stronger dispositional desire for A than for B with a disposition to choose A in all circumstances. In some circumstances, vivid representations of B may amplify the force of the desire for B, but not that of the desire for A, leading the agent to pursue B rather than A. Ordinary desire attributions take this into account. As someone trying to quit smoking is asleep, we might say, “He wants to quit more than he wants to keep smoking. That’s what he says when you ask him, and it’s why he stays away from smoky places. As he knows, the smoky air will inflame his desire for another cigarette, and then he’ll smoke again.” A good psychological model for this kind of behavior would attribute to the smoker a stronger dispositional desire to quit than to smoke, and then allow vivid sensory representations of cigarette smoke to amplify the motivational force of the desire to smoke. This would explain both why he would smoke in a smoky environment, and why, knowing this, he chooses to stay away from smoky places. An interesting feature of this desire attribution is that while the smoker is in some sense more disposed to smoke than to abstain when in smoky places, we still treat his desire to abstain as the stronger dispositional desire. Ordinary mental-state attributions thus seem to distinguish dispositional desires from overall dispositions to perform some action.

While it is hard to give a good philosophical account of dispositional mental states, because of difficulties with properly spelling out the conditions under which they manifest and because of the conditional fallacy, the notion of dispositional desire is an important part of ordinary folk psychology.¹³ Describing the sleeping person as wanting to quit more than he wants to smoke is a fully intelligible desire attribution, even when we add that being in a smoky place might make him decide to smoke. We often make similar desire attributions about dear but distant friends, talking about what they want most while being unaware of what they are immediately motivated to do at the moment. We can do so while being fully aware that vivid representations of other things they desire (or other situational factors) might lead them to pursue these other things instead. Humeans about

¹³ For a discussion of the conditional fallacy, see R. K. Shope, “The Conditional Fallacy in Contemporary Philosophy,” *Journal of Philosophy* 75 (1978): 397-413; and C. B. Martin, “Dispositions and Conditionals,” *The Philosophical Quarterly* 44 (1994): 1-8.

rationality should invoke the desire strengths being compared in such cases. Those who would object that there is something problematic about the very notion of such dispositional desire strengths run the risk of treating some of our ordinary folk-psychological talk as nonsense.

Dispositional desire strengths allow us to build good explanations of motivation, justifying us in positing them as real psychological entities. Consider the previously offered cases where vivid representations of chocolate, warmth, injections and smoking increase the motivational force of desires for or against them. Each case involves vivid representations increasing the motivational force of desire beyond a pre-existing baseline. This baseline is the dispositional strength of the desire. To put the point more generally, how an agent is motivated is explained by factors including (1) the agent's means-end beliefs, (2) what is vividly represented and (3) some fixed mental characteristic of the agent that explains why vivid representations and means-end beliefs concerning some things have motivational effects. On a Humean view, dispositional desire strengths are this third factor. Information about how to meet old friends would move me; information about how to acquire chewing gum would move me less; information about how to acquire Styrofoam would not move me at all. The strengths of my dispositional desires concerning these things explain these counterfactuals. We should say that dispositional desire strengths remain stable as means-end beliefs and vivid images come and go, since the counterfactuals they support about the motivational effects of various means-end beliefs and vivid images are stable. Even if I am currently seeking chewing gum, I would abandon this pursuit if I believed that doing so would allow me to meet a long-lost friend. If dispositional desires explain how strongly various things motivate us, we have reason to think they are real, and we may invoke them in psychological and normative theories.

Dispositional desire strength is superior to the immediate motivational force created by desire in fitting our intuitions about rational action. Changes in our situation that leave our relevant beliefs and the dispositional strengths of our desires constant while changing the immediate motivational force that our desires produce do not generally change our reasons for doing various things. The recovering smoker is more motivated to smoke when placed in a smoky bar than he is when jogging, but this does not mean he has greater reason to smoke. Intuitively, he has just as little reason to smoke, even as his motivation increases. While it may be more understandable for him to smoke when stuck in a bar than suddenly to veer into a store while jogging, buy cigarettes and smoke them, this just means that it is a form of irrationality that better fits our understanding of how people's motivations change with the circumstances.

Here I am just defining the "right answer" – the rational thing to do – as the option that will maximize expected dispositional desire satisfaction. I am not laying out a decision procedure constitutive of rational deliberation, or

claiming that any particular mental events of sensory or imaginative representation must figure in deliberation for one to act rationally.¹⁴ One can act rationally after representing all options vividly or representing all options nonvividly. One can even act rationally when representing the object of one's weaker dispositional desire more vividly than the object of one's stronger dispositional desire, as long as the stronger dispositional desire wins. (Perhaps in these cases one could be described as irrationally reluctant to pursue one's greater good. All I claim is that the action would be rational.) There may also be ways to miss out on maximal expected dispositional desire satisfaction, and thus be practically irrational, without any difference in the vividness of representations. Depression may prevent someone from acting on her strongest dispositional desire. Perhaps in some cases of impulsive action, weaker dispositional desires motivate action before stronger dispositional desires can exert their influence. I focus on unequal vividness because it presents us with clear, everyday examples of how we can fail to maximize expected dispositional desire satisfaction.

Now I hope it is clear how the double-Humean view allows irrational action. The Humean theory of action should say that agents do whatever their desires produce the greatest motivation for them to do at the moment when they engage in action, while the Humean theory of rationality should say that rational actions are those that maximize expected satisfaction of the agent's desires construed dispositionally, while irrational actions fail to do so. Since more vividly representing the object of a weaker dispositional desire may bring it about that the strongest dispositional desire does not exert the greatest motivational force, agents can act irrationally.

This double-Humean view accords with our pretheoretical judgments about how common rational action is. It accounts for the truth of our pretheoretical belief that people act irrationally only in a minority of cases, since the mental states that justify action – dispositional desires and means-end beliefs – play a substantial role in explaining it. Our dispositional desire strengths are a large factor in determining the immediate motivational force produced by our desires as we act, so there is reason to expect our actions to maximize expected dispositional desire satisfaction most of the time. But there is room for irrationality – for example, when we have unequally vivid images of the objects of desire, making a weaker desire especially violent and thus more motivationally forceful than its dispositional strength would suggest. Thus the double-Humean view explains why people usually, though not always, act rationally.

¹⁴ For impressive arguments that even akratic agents can act rationally, see Nomy Arpaly, *Unprincipled Virtue: An Inquiry into Moral Agency* (Oxford: Oxford University Press, 2003). Arpaly's arguments incline me to think that no particular decision procedure is necessary for rational action.

5. The Spirit of the Humean Theory of Rationality

Korsgaard seems to hold that proposals like mine are not true to the spirit of the Humean theory of rationality. Her objection to such proposals seems to be that the reasons for accepting the theory prevent its defenders from distinguishing the rational action from the one we actually do. I concede that double-Humeans will no longer be able to use a traditional argument for their view of rationality, which grounds rationality in motivational force, if they do as I suggest. But this argument does not capture what is really appealing about the Humean theory of rationality, and defenders of the theory can happily set it aside.

Korsgaard writes that “the problem would be solved if we could make a distinction between a person’s ends and what he actually pursues” and suggests two ways to do this (1997: 230). The first is to “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” (ibid). Korsgaard is right that specifying what one has reason to desire and what one does not have reason to desire would bring in additional principles that would violate the spirit of the Humean theory of rationality. As David Sobel writes, “If someone added to instrumentalism an account of an agent’s true ends such that necessarily all agents had the end of being moral, then the distinctive traditional subjectivism of instrumentalism would be lost” (2001: 486). My solution incorporates no additional principles distinguishing rational and irrational desire, preserving the subjectivism that is essential to the Humean theory of rationality and escaping this objection.

The second way she describes includes my proposal: “[W]e could make a more psychological distinction between what a person thinks he wants or locally wants and what he ‘really wants.’” While the question of what someone “thinks he wants” plays no role in my proposal, one could describe the distinction I am using as between what someone “locally wants” and what he wants in a nonlocalized way. According to Korsgaard, this second proposal is not available to double-Humeans. It “takes us beyond instrumental rationality, although this may not be immediately obvious” (1997: 230). This is because:

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 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 [ibid].¹⁵

¹⁵ The phrase “what you ‘really want’” is not good for picking out dispositional desires in contrast to the immediate motivational force that they produce. “Really want” seems to be

Korsgaard's objection to this move seems to be that it requires us to attribute some normative force to psychological states that are unsuccessful in causing action, while withholding the same force to the psychological states that actually cause action. My solution does this. But it is not clear why it takes us beyond the Humean theory of rationality, which demands that we treat rationality as a matter of maximizing expected desire satisfaction. What prevents "desire satisfaction" here from being the satisfaction of dispositional desires? Korsgaard seems to think that Humeans are required to identify the rational action with the one actually performed. But why should any Humean have this commitment?

Humeans may incur this commitment if they rely on what Mark Schroeder calls the Classical Argument, which adds existence internalism about reasons to the Humean theory of motivation to get the Humean theory of practical reason. As Schroeder summarizes the argument, "If having a reason requires motivation, and motivation requires a desire, then having a reason requires having a desire" (2007: 7).¹⁶ Korsgaard presents a version of this argument as her opponents' reason for rejecting categorical imperatives: "Since the idea of being motivated by belief alone seems mysterious, the suspicion arises that categorical imperatives cannot meet the internalism requirement, and they are supposed to be especially problematic" (1997: 220). While Schroeder's version of the Classical Argument concerns what it is to have a reason and not which action is rational, it can be straightforwardly extended to deliver a Humean account of rationality. If (1) its being rational to do something requires that one be most motivated to do it, and (2) being most motivated to do something requires that one desire it most, then (3) it will always be rational to do what one desires most.

The first premise of this argument, which could be called rationality/motivation existence internalism,¹⁷ forces Humeans to identify rationality with motivation, much as Korsgaard says. It leaves them unable to explain irrationality. If one always does what one is most motivated to do – and it is unclear what else being most motivated could mean – rationality/motivation existence internalism will imply that one cannot fail to act rationally. When we loosen the tie between rationality and motivation by grounding rationality in dispositional desires, we deny the first premise of the Classical Argument for the Humean theory of rationality.¹⁸ But if we reject

ambiguous between the entities on both sides of the distinction. If there is a sense in which I do not really want the chocolate dessert that I know will wreck my diet, surely there is another sense in which I do.

¹⁶ Also see Wedgwood 2002 for criticism of this argument.

¹⁷ This would follow and extend the terminology in Stephen Darwall, "Reasons, Motives, and the Demands of Morality: An Introduction," in Stephen Darwall, Allan Gibbard and Peter Railton, eds., *Moral Discourse and Practice* (Oxford: Oxford University Press, 1997).

¹⁸ For the most influential paper incorporating a version of the Classical Argument, see

this argument, do we have any reason to accept this theory? If not, Korsgaard will be right that the Humean theory of rationality cannot be grounded in mental states that fail to produce action, on pain of giving up the reason to accept it.

Fortunately, better arguments are available to Humeans about rationality. As Schroeder writes, “I don’t think that the Classical Argument is a particularly good argument for the Humean Theory of Reasons, and I certainly don’t think that it is what really gives the Humean theory its deep philosophical appeal” (7). He argues for the Humean Theory of Reasons on the grounds that it provides a reduction of the norms of practical rationality that fits our intuitions about these norms, setting aside cases in which our intuitions are likely to be unreliable. In particular, it accounts for our intuitions about “why it is that some reasons are reasons for everyone, while other are reasons for only some” (212). Using dispositional desires to determine what reasons we have would preserve this attractive feature of the theory, and the corresponding Humean view about rationality should be correspondingly attractive.

It is beyond the scope of this paper to argue that the Humean theory of rationality best accounts for our intuitions, or to examine the arguments that others have offered. But it is within the scope of this paper to call attention to this way of defending it, urge double-Humeans to argue for it on these grounds and suggest that they respond to Korsgaard by setting the Classical Argument aside.¹⁹ This advice is continuous with the argument for formulating the Humean theory of rationality so that it concerns dispositional desires rather than immediate motivational forces. This formulation better fits our intuitions about rationality. It need not coincide with the way actions are motivated.

6. Conclusion

Korsgaard criticizes “a theory very fashionable in the social scientific and economic literature, sometimes called the self-interest or economic theory of rationality” according to which “it is rational for each person to pursue his overall good: to act on some variant of the principle of prudence” (1997:

Bernard Williams, “Internal and External Reasons,” in his *Moral Luck* (Cambridge: Cambridge University Press, 1981). As Williams distinguishes what we are motivated to do from what we might be motivated to do if we deliberated well, he still might be able to make room for irrationality.

¹⁹ For another reading of Korsgaard’s argument, see Stephanie Beardman, “The Special Status of Instrumental Reasons,” *Philosophical Studies* 134 (2007): 255-87. Beardman’s reading has Korsgaard attributing an excessively strict principle of constitutiveness to her opponents. My version of the Humean theory of rationality escapes this version of Korsgaard’s objection as well, as it does not rely on the problematic principle that Beardman presents. Beardman’s point that “the structure of desire and related states, in a realistic psychology, will allow for the complexities of human perversity” is one that this paper develops in detail.

230). Here I agree with her. Economists are often less careful than philosophers about untangling normative and descriptive issues. Furthermore, their models of how people plan for the future often ignore the psychological point that things sensed or imagined more vividly will increase the violence of our passions and be pursued with greater vigor.²⁰ Where they insist on these models, they may be subject to Korsgaard's criticism, and cannot use the response that I have offered to double-Humeans. One hopes that the increasing popularity of psychologically informed economics will lead to more sophisticated models of the mental states involved in action.

Philosophers inclined towards the double-Humean view, however, need not wait on these developments. We are already in position to develop broadly desire-based accounts of action and practical rationality that leave room for irrational action. We can use the immediate motivational force produced by desires at the moment of action to determine what an agent will do, and use dispositional desires to determine what an agent ought to do. This will require reconceiving the reasons for accepting the Humean theory of rationality: rather than defending the view on the basis of the Classical Argument, we should argue that it best accounts for our intuitions about practical rationality. If these arguments are successful, we will be left with theories of action and rationality that account for just as much irrationality as there is.

Neil Sinhababu*
National University of Singapore
Department of Philosophy
neiladri@gmail.com

²⁰ The Ricardian Equivalence hypothesis is one influential position in economics that would be imperiled by more attention to the way that vivid images increase the violence of our passions. Defenders of Ricardian Equivalence hold that governments cannot successfully stimulate the economy through deficit spending, because instead of spending their stimulus money as intended, people will save all of it to defray future taxes that will be levied to make up the deficits. But if we assume that people have more violent passions for nearby things, since they have more vivid images of these things, we can see why they will stimulate the economy by spending money, even at the expense of their future security. The goods they achieve by saving are further away and less vividly represented, producing less violent passions; while the goods they achieve by spending are nearer and more vividly represented, producing more violent passions that will have greater influence in driving their behavior. For an influential defense of Ricardian Equivalence, see Robert Barro, "Are Government Bonds Net Wealth?" *The Journal of Political Economy* 82:6 (1974): 1095-1117.

* This paper was greatly improved by feedback from audiences at the 2008 AAP conference, the University of Southern California, the University of Miami, the University of Illinois, Oriol College Oxford and the University of Michigan. In particular, I would like to thank David Dick and Mike Titelbaum for getting me to think about the issue, and Michael Smith, Mark Schroeder, Rachael Briggs, Kenny Easwaran, Ben Blumson and Nick Stang for helpful comments.