Moral Metaphysics or Moral Psychology?

Adrian Piper's *Rationality and the Structure of the Self*

Paul Guyer

1. Piper's Ambition

Adrian Piper's *Rationality and the Structure of the Self* (second edition, 2013)[1] is a monumental work in meta-ethics and moral psychology, inspired by Kant but dealing decisively with the history of a considerable portion of twentieth-century moral theory along the way. The work consists of two volumes, the first a critique of a "Humean" approach to its subjects and the second the defense of a "Kantian" approach. These terms are placed within scare-quotes to indicate even though Piper's work is thoroughly informed by detailed knowledge of both Hume and Kant, the work is far from being an historical work; the "Humean" model criticized in Volume I is the "Belief-Desire Model" that all practical reasoning begins with preferences not set by reason itself and the "Utility-Maximizing Model" that reason functions purely instrumentally in determining how best to realize the goals set by such non-rational preferences, while the "Kantian" view is that reason itself sets the overriding ends of practical reasoning.

Piper surveys numerous versions of "Humeanism," including not only the paradigmatic version of Richard Brandt but also, no doubt controversially, the "instrumentalism" of none other than John Rawls; I will not use my space to go through the details of her treatment of "Humeanism," but the gist of her criticism is that any purely preference-based conception of practical rationality, even one such as Rawls's model of "reflective endorsement" or Frankfurt's model of second-order preference, is going to leave the coherence of any temporally-extended manifold of an individual's choices at the mercy of some preference that can itself always be changed, and thus allows for no realization of a stable, unified self acting over time. As she summarizes her criticism before the decisive chapter of Volume II, there is "nothing inherent in the unreconstructed utility maximization model of rationality that requires a rational agent to be psychologically consistent, and no resources within the conventional constraints of this model for

inferring from mere *psychological* inconsistency any violation" of any norm, whether interpersonal, as we would assume moral norms to be, or even intrapersonal, as we might assume prudential norms or counsels to be. One might think that "utility theory" is merely a "truism about always doing what we most want to do," but Piper's criticism is that without the framework for structured selfhood that only a "Kantian" model of the self offers, there is not even a "we" or an "I" that can have a coherent conception of what it most wants to do (II:163). The "Humean" self is indeed just a bundle, not only of impressions and ideas but also of wants with no consistency constraints. This may not seem fair to the historical Hume, or at least the mature Hume of the *Enquiry concerning the Principles of Morals*, for whom "tranquillity" was a chief moral good, in the form of a quality immediately agreeable to oneself, the achievement of which would presumably require the satisfaction of strong psychological consistency constraints.[2] But it seems a fair criticism of much "Humeanism."

If it is Hume's bundle theory of the self that underlies the problems of the "Humean" model of practical rationality (though ignored by Hume himself in his own practical ethics),[3] it is Kant's conception of the unity of apperception that drives Piper's model of "Kantian" rationality. This is important, because unlike other versions of contemporary "Kantian constructivism" to which Piper's approach might be compared, her argument does not depend upon any special conception of *practical* reason, but on a conception of *reason* or *rationality* as such comprising most explicitly a conception of consistency that is derived from logic but applied to choice. This is a difference with, for example, Christine Korsgaard's versions of constructivism, either her earlier version in *The Sources of Normativity*, where commitment to the moral law is argued to be the necessary condition of having a *practical* identity, or her later version in *Self-Constitution*, where the argument turns on a conception of what it is to act on *a reason.*[4] Piper's view is rather that commitment to morality is implicated as the condition for including intentions and choices in a self that is unified by the conditions of rationality as such.[5] In other words, namely Kant's words, the guiding-thread of Piper's approach is "that the critique of a pure practical reason, if it is to be carried through completely, be able at the same time to present the unity of practical with

speculative reason in a common principle, since there can, in the end, be only one and the same reason, which must be distinguished merely in its application."[6]   To be sure, for Kant there is a special issue about the possibility of practical reason, namely whether the application of the one and only reason that there is to the realm of practice, that is, to our choices, can be *efficacious*, or that we can act in accordance with the dictates of reason even when our empirical circumstances seems to suggest that we cannot. This is the challenge of showing that "pure reason of itself alone suffices to determine the will" rather than that "it can be a determining ground of the will only as empirically conditioned," the challenge that leads Kant so say that "It is therefore incumbent upon the *Critique of Practical Reason* as such to prevent empirically conditioned reason from presuming that it, alone and exclusively, furnishes the determining ground of the will."[7] Of course Piper does not want to follow Kant in supposing that the efficacy or practicality of pure reason can be secured only by establishing the theoretical *possibility* of a free will at the noumenal level by means of the argument that spatio-temporal causality is only an artifact of the phenomenal appearance of reality to us and then the *actuality* of the freedom to be rational by inference from our undeniable obligation to act in accordance with the moral law.[8] But one of the most valuable parts of her work is her empirical account of how reason can in fact establish a firm foothold among our desires, one that is firm enough so that even when we are tempted to violate its norms we still seek to preserve the appearance (in an ordinary sense) of rationality, or what Piper diagnoses as "pseudorationality." Piper does not help herself to Kant's own assurance that pure reason can always be practical no matter what, but she does show in detail how we can learn to guide our choices by reason and how we attempt to preserve the appearance of reason even when we do not really want to be rational.

Piper's recognition and diagnosis of pseudorationality is important because it shows, in spite of some language that might suggest otherwise, that she does not succumb to the fantasy of assuming that it is a metaphysical necessity that we must choose rationally if we are to choose at all, as in my opinion Christine Korsgaard does, who can therefore explain moral failure only as defective reasoning about what morality requires rather than as a direct choice not to be moral.[9]

Rather, she recognizes that human beings have often have a genuine choice whether *to be rational*, and they do not always choose to be rational, although they are disposed to preserve the *appearance* of rationality not only to others but to themselves as well.   I will suggest that a chief virtue of Piper's work is her illuminating account of rationality and irrationality in practice, which I believe is more psychological than metaphysical. I will conclude by briefly raising the question whether she has exploited Kant's own conception of reason as fully as she might have for her purposes.

2. Piper's Model of the Rational Self

Piper's formal account of the rational structure of the self is presented in Chapters II through IV of Volume II of *Rationality and the Structure of the Self*, while Chapters V through VIII develop her moral psychology, Chapters V and VI offering her positive account of how reason influences action and VII and VIII her analysis of pseudorationality. Chapter IX lies at the intersection of meta-ethics and moral psychology with an account of how imperatives arise from the application of norms to our psychology, and the remaining Chapters X and XI deal with the moral issue of inclusiveness and its opposite, xenophobia.

Chapter II is aimed at showing that "without satisfying at least two familiar and very weak consistency requirements of theoretical reason that are deeply embedded in the structure of a unified self, we could not be motivationally effective agents at all" (II:52). Both here and in Chapter III, which concerns the concept of a "genuine preference," Piper argues at length that the standard propositional calculus applies to "subsentential" elements as well, laying the groundwork for the claim that the most fundamental logical or formal requirements of rationality, such as the necessity of avoiding self-contradiction also apply to the domain of choice, modeled as a choice between representations of possible actions rather than propositions about possible alternatives. I will not go into the details of Piper's argument with Robert Brandom's interpretation of Kant here, although her claim that "judgment is not the fundamental unit of

awareness for Kant; intuitional representations are" (II:65) is consistent with my own qualms about Brandom's attempt to enlist Kant in the cause of his "inferentialism."[10] The point of Piper's painstaking logical work is to demonstrate that there can be such a thing as contradictory intentions, not just contradictory propositions, and that such intentions need to be avoided in order to have a unified self. In particular, Piper argues that unified selfhood requires the "two familiar and very weak consistency requirements." These are what she calls "horizontal" and "vertical consistency." Horizontal consistency requires "that I must conceive all the things and properties that are simultaneously rationally intelligible to me as logically consistent with one another" (II:85). This means that not only must the various beliefs that I hold simultaneously be consistent rather than contradictory with one another, but also that the various intentions I might be regarded as having at the same time must also be consistent with one another. Vertical consistency requires "that if I recognize some thing or property as a certain kind of thing, I must also be able to conceive it as of the same higher-order kind as is the kind of thing I originally recognized it to be. So, for example, if I recognize something as a three-dimensional thing, I must also be able to conceive it as a thing of a certain length; if I recognize going to the store as a tedious errand, I must also be able to recognize it as nothing extraordinary" (II:85-6). In other words, I must be able to move up from more concrete and determinate conceptions of things to less concrete and determinate ones, and, one would think, back down again as well, thus preserving coherence in my conceptions of things - and intentions -- by not, for example, conceiving of one and the same thing as both three-dimensional and yet as lacking one of the standard three dimensions, namely length. Piper's second illustration of vertical consistency seems less clear-cut: the idea seems to be that if I hold the maxim of not avoiding tedious errands when necessary and also conceive of going to the store as a tedious errand, then I would still accept rather than reject the intention of going to the store when necessary; this would seem to hold whether it being necessary for me to go to the store is a frequent or extraordinary occurrence in my life. But my more important concern with this phase of her argument is that while horizontal and vertical consistency certainly seem like necessary conditions of unified thought of

any kind, thus of unified selfhood, they do not seem nearly sufficient, for unified selfhood seems to be something that exists over time, even if with development and revision, and the kind of consistency that the unified self needs is surely as much diachronic as synchronic.

Any complaint here may be more about emphasis than substance, for Piper certainly does recognize the need for diachronic consistency in the unified self. At the beginning of the discussion of horizontal and vertical consistency, she alludes to the fact that an "agent's perspective changes over time, and with changes in her state, character, surroundings, and history," evolving "both progressively and regressively as the agent evolves over time," although she there focuses only on the conditions necessary to make "the sum total of things and properties" represented by the agent "simultaneously rationally intelligible to [the] agent at a particular moment" (II:84). But as the chapter continues, Piper adds to the conditions of rational agency -- for that is what she is talking about -- what she calls the requirement of self-consciousness or the "self-consciousness property," or the concept of "*my experience*" (II:97), and this requirement, inspired by Kant's conception of the unity of apperception, seems to imply a conception of oneself as a temporally-extended subject of particular representations, intentions, and ultimately choices and actions, with a requirement of diachronic consistency among those states and not just synchronic horizontal and vertical consistency. Piper does not come right out and say this, but when she refers to "concepts of properties that attach to [an agent] as a *subject* of experience, i.e. to the way she experiences . . . events, objects, and states of affairs, for example, being surprised by something, or open-minded to something, or desiring something," and then further argues that "In order for an agent to regard his experience of different things as objects of his experiences . . . he must be capable of viewing such experiences as not only *affecting* him, but also as being partly *determined* by him" (II:98), she is surely describing the temporally-extended experience of a temporally-extended self. Being surprised by something, for example, is a temporally extended experience, even if not a very extended one, involving a change of mental state from complacent ignorance to startled recognition, and being open-minded is being disposed to allow one's beliefs, desires, and intentions to change in reasonable

response to new information rather than to hold them fixed in spite of new information. Only a temporally extended self can be open-minded, or for that matter close-minded.

That it is actually the requirement of diachronic consistency (but also willingness to change in reasonable response to new information) that transforms the merely necessary conditions of synchronic horizontal and vertical consistency into something more like the sufficient conditions for unified selfhood becomes clear in Piper's Chapter IV, on "The Concept of a Genuine Preference." This chapter includes the statements that "Practical reasoning just is an application of theoretically rational rules of causal inference to the special-case event of intentionally conceptualized behavior of a goal-oriented kind," thus that "if decision theory is a formalization of practical reasoning, then it is a special case of the classical logic that formulates theoretical reason" (II:112). Both the reference to causal inference and that to goal-oriented behavior take us into the realm of the diachronic and not merely synchronic consistency, since a goal is something one aims to realize at some time later than the now when it is first formulated and causal inference concerns how one thinks one can get from now to then. The temporally extended character of the unified selfhood that Piper is discussing becomes even more explicit when she writes that:

> a conscious and intentional chooser [has] to satisfy two necessary conditions:
>
> (a) she must be able to form and apply consistently through time the concept of a thing's ranking superiority -- and therefore some other thing's ranking inferiority -- over a series of pairwise comparisons; and
>
> (b) she must remember the relation of the two alternatives she is presently ranking to the third she is not. (II:116)

(a) explicitly says that a unified agent has to apply its ranking of things (possible objects of action or actions) consistently "through time," and (b) adds that such an agent has to remember its previous pairwise rankings in order to consistently order multiple rankings (just as in "threefold synthesis" a Kantian cognitive subject has to "reproduce" its previous representations in order to recognize their unity under a concept)[11] -- and remembering obviously is possible

only for a being with a temporally extended experience. The concept of a genuine preference is then of a preference that can survive standard logical requirements such as transitivity in all pairwise rankings involving it as well as horizontal vertical consistency in its conceptualization. An agent can have multiple genuine preferences as long as they are each fully transitive with regard to relative rankings of relevant preferences and are consistent with each other. They have to fit together coherently in an agent's "perspective" or "experience," thus "genuine preferences as defined in [Piper's] variable term calculus are integrated into an agent's perspective as some among many other experiences that also include thoughts, beliefs, perceptions, and emotions . . . . this requirement, together with that of horizontal and vertical consistency, secures the rational intelligibility and logical consistency of a chooser's preference and the self-determining agency of that chooser" (II:259). The point of all this, worked out with a great deal of formal detail, is that "This approach subordinates the utility maximization model of rationality to the more general and universal requirements of classical logical consistency, and so divests it" -- the utility-maximization model, that is -- "of its pretensions to universality of application" (II:263). The rational agent is governed by an ideal of logical consistency, not utility maximization.

The implication of this position is then brought out in Chapter IV, entitled "McClennen on Resolute Choice." Edward McClennen's conception of a resolute choice is that of a choice to which a chooser remains committed through multiple choice-points at which it might *seem* reasonable to revise her choice on grounds of utility-maximization because it is actually sticking to the original choice that maximizes utility.[12] "The basic ideas is that under certain circumstances an agent might be disposed to guide his behavior according to certain sorts of rules in order to maximize freedom, flexibility, or scarce resources because he understands that violating the rule would be costly of these things" (II:166), which the agent values or which contribute to his conception of utility. A resolute choice is essentially one that promises to maximize utility all things --that is, all its foreseeable consequences over time -- considered, or, we might say, one that looks at the synchronic consequences of a choice from a rule-utilitarian point of view. A resolute choice "involves shaping future preferences in light of present rational

deliberation as to how future choices may maximize overall utility when coordinated with present ones" (II:172), thus "solving a coordination problem between the conflicting interests the self has at different times" (II:173). For McClennen, the value of a resolute choice is that it maximizes utility in comparison to an alternative pattern of irresolute choices; for example, keeping promises previously made, even when there might seem to be advantage in breaking them, turns out to be what really maximizes utility -- at least in most cases ("under certain circumstances"); presumably there are some circumstances in which resolute choice is not the best strategy for maximizing utility after all.

But this is not the point that Piper presses against McClennen. Her thesis is rather that "acting on her original resolve maintains the horizontal and vertical consistency of" an agent's "experience over time and at each moment is itself a reason. That is, preserving a unified and internally coherent self is a good that justifies" the agent's "resolve even though that unified self fails to maximize utility on" the particular occasion of some choice-point in this history of her resolute action" (II:176). She then glosses this claim by stating that it

> is not about a contingent psychological preference for consistency, but rather about the metaphysical consistency that genuine preference -- indeed, any kind of preference -- presupposes: I may or may not have a particular liking for consistency; but unless I am a unified and internally consistent self in the first place, the issue of my psychological likes and dislikes cannot arise. . . . although preserving an internally coherent self in this sense is a good, it is not an end, goal or intentional object that an agent can adopt or at which he can aim. Therefore while it can be a justifying reason for action, it cannot be the object of a preference. (II:177) --

and therefore presumably cannot be surrendered in the case in which maintaining does not after all promise maximal utility.

In this remark, Piper's position does sound like that of Korsgaard in *Self-Constitution*, which is, namely, that a unified agent, in Piper's terms, or an agent that is not a mere heap, in Korsgaard's, cannot but act rationally, although it might fail to correctly infer what is rational in

some particular case -- with this, to be sure important, difference, namely that Korsgaard interprets acting rationally by means of the relatively unanalyzed notion of acting on "a reason," whereas Piper interprets it by means of the model of synchronic and diachronic logical consistency which determines what can count as a reason in the first place. But the force of such a "metaphysical necessity" is obscure, at least to me. If it is meant to define an ideal of rational agency to which the agency of flesh-and-blood human beings ought to conform, then its normative hold on such actual human beings remains to be explained; if it is meant to describe a regularity to which flesh-and-blood human beings invariably conform, then it seems to be belied by the all too frequent divergence of human beings from it. Examples of human non-compliance with this logic of agency are all too frequent, just as the US is currently giving the world an example of a flesh-and-blood human being who regularly flaunts the law of non-contradiction in descriptive statements ("I made that statement to Billy Bush and I did not make that statement to Billy Bush"). Moreover, Piper's won rich and subtle description of the forms of human *pseudorationality* make it clear that she ultimately think that flesh-and-blood human beings have a psychological disposition to preserve the unity of their selfhood in many cases but at least the appearance of unified selfhood in yet others, and this recognition is not compatible with a metaphysical necessity that actual human beings preserve actual consistency. If they did, they would never have to strive so hard to preserve the mere appearance of it.

3. Piper's Moral Psychology

It seems to me that, talk of metaphysical necessity aside, Piper's commitment is actually to a normative position that unified selfhood is a good that human beings ought to strive to realize along with the claim in moral psychology that they often do but even when they do not they at least strive to preserve the appearance of it to others and even to themselves. I think that this is apparent even in her illustration of her purportedly metaphysical claim. She asks what sort of reason a would-be dieter has for sticking to his resolve to diet. The would-be dieter is called

Myron. Even if he is inclined at time t_3 to stuff himself instead of practicing the restraint on which he had resolved at t_1,

> Myron at t_3 still has a reason to abide by Myron-at-t_1's resolve to diet. The reason is that his resolve at t_1 makes coherent and intelligible his sticking to his diet at t_3, whereas it makes abandoning his diet at t_3 incoherent and disorienting: Sitting stuffed, queasy and stupefied at his dinner table after having gorged himself on food he had for good reason resolved to forego, Myron is naturally confounded by the empty plates and distended expanse of stomach before him. He asks himself, Did he really eat all that? And Why? To where did the sober and disciplined person he was at t_1 disappear? Myron chooses to avoid this condition of disconnected bewilderment by sticking to his diet at t_3, quite aside from the threat of precommitment or cost in resources of abandoning it. Once again, acting on a genuine preference is itself a good that, by ensuring the internal unity and coherence of his self at each moment and through time, justifies Myron's resolve. (II:177)

I find it difficult to read this passage without taking it to imply that Myron has a *genuine choice* whether or not to stick to his diet at dinner-time, thus that there is no *metaphysical* necessity that he maintain his resolve, and further that his reason for striving to maintain a unified self, should he choose to act upon it, is that an irresolute, "disconnected" self is bewildering and all things considered *unpleasant*, in spite of the momentary pleasure promised by the prospect of abandoning his diet and stuffing himself. And if it is in this way that preserving a unified self is a good, then I do not see why it cannot also be the object of a preference, although since a unified self is a structure or relation among preferences (and beliefs, emotions, etc.) that can only be manifest over time, it is in some sense not of the same order as the particular object of a particular preference ("A triple-scoop sundae now!") and choosing to act to realize or preserve it is not incompatible with having a first-order preference that is incompatible with that aim ("I wish I could have that sundae now, but I will not!").

   I think that the wonderful Chapters V and VI of Piper's Volume II can be understood only on the assumption that she does posit a psychological disposition but not a metaphysical

necessity to maintain unified selfhood, with the logical rules defining such unity serving as norms to be realized or, as she shows in Chapters VII and VIII, at least honored in the breach. Chapter V begins with a re-assertion of the metaphysical necessity of unified selfhood, and argues that it is because such selfhood is a precondition of any particular rational choice it cannot itself be seen as the object of any particular choice, thus in that sense as a good (II:192, see also 194, 198). But the psychological approach pops up its head in the statement that "if the promise of rational intelligibility is the carrot that disposes an agent to seek only those ends that satisfy the . . . consistency criteria, the threat of psychosis is the stick that discourages her deviation from them" (II:191): such a disposition is a psychological characteristic of a flesh-and-blood agent, the threat of psychosis is something disturbing and unpleasant to a normal human being, so conversely the prospect of avoiding such psychosis will be appealing, whether you call it a particular good or not. Further, Piper describes "the highest-order disposition to literal self-preservation as a kind of sentinel that repels all . . . threats to the theoretically rational unity of the self, filtering out inconsistent or conceptually anomalous beliefs, desires, and impulses, and admitting in only those that qualify as genuine preferences" (II:201); this describes a causally efficacious psychological disposition. And then she describes the commitment to coherence, including the coherence of genuine preferences, as something to which the agent must become "habituated" through a "long-term" "project of moral self-improvement" (II:209), success at which depends upon a variety of circumstances, including a "community of spontaneous agents" with adequate "material resources" (II:233), something which neither has been nor is guaranteed for all human beings, and which can hardly be considered a metaphysical necessity of human existence; it is rather a causal condition for flourishing human existence. Piper also observes that once the psychological disposition toward unified rather than psychotic selfhood "does take hold, it will bear a relation to individual act-tokens most of the time, even if the requirements of duty are violated occasionally, and even if one then need not be preoccupied with these requirements most of the time" (II:209); this is an eminently reasonable thing to say, but it could not be said if unified selfhood were a metaphysically necessary condition of genuine action. It is the kind of

thing that you can say about a psychological disposition that is usually but not always efficacious.
A "motivationally effective intellect," as Piper calls it (II:213), is not a metaphysical necessity
but a desirable psychological possibility for human beings in fortunate circumstances. Piper also
calls it an *ideal* of rational motivation (II:228) -- something at which we can and should aim, but
not something that is automatically given. She also characterizes this ideal as one of interiority
as opposed to mere spontaneity: "The second ideal is grounded in the value of interiority; of a
vivid and extended life of the mind that includes imagination, intellection, and reflection; these
are the foundations of transpersonal rationality" (II:233). This is an ideal that will appeal to most
people once they understand and experience it, and that will motivate them; but there is no
metaphysical necessity that human beings all recognize this value. This transpersonal rationality
can "flower" and undergo "development" and "growth": these are things that can happen to
psychological dispositions, not to metaphysical necessities.

Having left the talk of the metaphysical necessity rather than psychological value of
unified selfhood behind, Piper offers a moving account of its components (Chapter VI) and of the
ways in which we try to preserve at least the appearance of it, as much for our own benefit as for
anyone else's, even when we are undermining it (Chapters VII and VIII). The former comes in
her account of "moral interiority," the latter in her account of "pseudorationality." Interiority
depends upon "modal imagination," the ability to imagine oneself in different circumstances than
actually obtain, which in turn leads to the ability to imagine the interior state of others (one
thinks here of Adam Smith, though he is not one of the Smiths in Piper's otherwise
comprehensive bibliography). This ability is the foundation of impartiality and compassion,
which consists in empathy, sympathy, and recognition of the need for symmetry in response to
the suffering of others (II:241), and is the opposite of self-absorption (II:248). Piper offers
detailed analyses of all these concepts and accounts of how both ordinary and aesthetic
experience can foster compassion and ameliorate the tendency to self-absorption (II:254-5). Her
argument that genuine impartiality requires avoiding *both* self-absorption but also "vicarious
possession" by another -- identifying so completely with another's needs that one looses sight of

one's own -- is particularly illuminating (II:266). In a passage that might sound like updated Bishop Butler, she concludes that "strict impartiality requires the ability to balance the demands and interests of the self with those of others in accordance with a normative principle biased toward neither" (II:272).[13] This could also remind one of Kant's eventual conclusion, after the many twists and turns of his critique but then recognition of the moral significance of happiness, that "since all *others* with the exception of myself would not be *all* . . . the law making benevolence a duty will include myself, as object of benevolence."[14] Piper's analysis of "pseudorationality" then describes how we try to preserve the appearance of unified selfhood and its particular features even when we are not really aiming at it.

Kant is the inspiration for Piper's account of pseudorationality. One way one could put Kant's point, although not quite the way Piper does, is that while Kant characterizes rationality in practice, as in his comment about benevolence, as the recognition that one counts no less but also no more than all others, pseudorationality consists in insisting that the requirements of rationality apply to all others but not to oneself, at least when one thinks one can get away with that; one does not deny the demands of morality, but makes an exception for oneself.[15]f Piper instead invokes Kant's definition of evil from *Religion within the Boundaries of Mere Reason*, where, as she reports, Kant defines evil as the "subordinat[ion of] the requirements of principle to the demands of desire," where we then, Piper adds, "rationalize this by minimizing the authority of principle and magnifying the value of desire" (II:294). The point is that we still try to preserve the structure of rationality over all, but hope to get away with adjustments that will allow what is just the current structure of our desire, so to speak, to mimic the structure of rationality. In her detailed account, pseudorationality is then non-exhaustively analyzed into denial, dissociation, and rationalization. "In *denial* we suppress awareness of an anomalous particular, property, or state of affairs, by failing to recognize it as instantiating concepts supplied by our unified conceptual scheme," thereby releasing it from the requirements of horizontal and vertical consistency. "In *dissociation*, by contrast, the anomaly is not banished from awareness entirely, but rather identified solely in terms of the negation of some subset of the concepts that constitute

the agent's perception," thereby preserving the appearance of horizontal consistency, or, as Piper says, satisfying this requirement "degeneratively." And *"rationalization* consists in biased predication: in applying a higher-order concept too broadly or too narrowly to something, ignoring or minimizing properties of the thing that do not instantiate this concept, and magnifying properties of it that do," this time in order to preserve the appearance of, or degeneratively satisfy, the requirement of vertical consistency (II:292). Piper provides rich illustration of these forms of psuedorationality; for our purposes the general point is just that insofar as agents doing wrong strive to preserve at least the appearance of rationality, they must recognize the ideal of rationality and its appeal, but not as forcefully as they feel the pull of some desire. All of this makes sense only if we recognize rationality as a *norm* and an *ideal* with some, indeed considerable *motivational* and thus *psychological* pull on even wrong-doing human beings, but not as a *metaphysical necessity* or condition of the possibility of action as such. Unified selfhood might be constitutive for the *concept* of a genuinely rational agent, but for flesh-and-blood human beings it can only be a psychologically powerful ideal that is often strong enough to overcome particular desires but that it sometimes honored only in the breach.

This conclusion does not fit all of Piper's statements, but it seems to me the only way to make sense of her powerful descriptions of the motivational efficacy of the ideal of rationality on the one hand and the pseudorationality of wrong-doing evil on the other. So I will assume that she would ultimately agree with me in the resolution of the equivocation between a metaphysical and a psychological account of the necessity of rationality in favor of the latter. I now want to turn to a second issue, namely her transition from the ideal of unified selfhood and interiority to the requirement of moral impartiality. One may well see how the norm of unified *selfhood* entails consistent choices over time, or what we might think of as trans-momentary, *intrapersonal* impartiality, but how does this norm also entail *interpersonal* or what Piper calls "transpersonal" impartiality? Piper rejects Kant's own position that the application of a universalizability requirement to a proposed maxim will suffice to determine what is morally requisite, prohibited, or permissible on the basis of the classical objection that one can avoid

practical self-contradiction by carefully tailoring one's maxim for any proposed action so that its generalization would not undermine that proposed action (II:180-3). I have argued that Kant's has resources to reply to this objection, but will not rehearse my arguments on this score here.[16] The present question is how does Piper herself propose to make the transition from unified selfhood to full-blown morality? In what we have seen so far, Piper has shown how the logical requirements on genuine preference and on unified selfhood more generally provide the foundation for "the relationships of trust and responsibility that a stable interpersonal morality must presuppose" (II:188) -- no one could trust another person who does not maintain a unified self, because what that person says, promises, or does at one moment would imply nothing about what he would say, promise, or do at the next. But does acceptance of the requirements of unified selfhood itself entail commitment to the interpersonal principles of morality as normally understood? Is unified selfhood a sufficient as well as a necessary condition for commitment to morality?

The answer to this question again appears to lie not in logic or metaphysics but in moral psychology, confirming the psychological rather than metaphysical interpretation of Piper's conception of unified selfhood for which I have been arguing. The key move comes in her account of "interiority," although not so much in Chapter VI, explicitly devoted to this topic, which presupposes that the move has already been made, but in the final section of Chapter V, on "Transpersonal Rationality and the Ideal of Interiority" (II:233-9), which completes the account of "How Reason Causes Action." The idea, although once again I have to add some words to what Piper explicitly says, is that although it might seem logically possible to gain control over one's impulses by horizontal consistency, vertical consistency, and trans-temporal consistency within one's own personal set of concepts, the conceptual version of an idiolect, in fact, that is to say, in psychological reality, one does this by using interpersonally shared concepts with their relations of horizontal, vertical, and trans-temporal consistency. The key idea seems to be that self-conscious agents, who develop the capacity to reflect upon themselves and in so doing to imagine alternatives to their present feelings and desires that would better realize the norms of

consistency, thereby also learn to imagine the mental and emotional conditions of others and to respond to them in a non-impulsive way. "Thus interiorized agents of necessity develop the ethical capacities for impersonality, disinterest, selflessness, and impartiality that are engendered by the pleasures of abstract speculation and inquiry, in direct proportion to the vividness, clarity and power of the interior universes they are forced by circumstances to create." "Interiorized agents" do not automatically act upon such considerations, and indeed their ability to imagine the interior state of others also makes them able to carry out "calculated revenge, betrayal, deception, and self-aggrandizement," or to do so better than merely impulsive, unreflective agents could do; but memory -- presumably memory of indignities and injuries they themselves have suffered -- and imagination -- imagination of how others feel in such cases -- provide "the intrapersonal foundation for the negative moral emotions of guilt, shame, and resentment" and "the interpersonal foundation for the positive moral emotions of empathy, sympathy, pity and compassion" (II:237). The idea seems to be that the conditions for unified selfhood provide the conditions for the development of moral emotions, and these lead to moral principles, although they do not guarantee that even interiorized agents will always act in accordance with these principles. The "interiorized agent makes a comparative judgment about the superiority of her capacity for interiority itself. She is causally influenced" -- although presumably not necessarily determined -- "by the impartial directives it engenders because these define and make intelligible to her the kind of self she is. Interiority, then, is a necessary condition of transpersonal rationality; and this, in turn, is a necessary condition for the development of a recognizable morality" (II:238). But the latter necessity, at least, seems psychological rather than metaphysical. Reflectiveness or self-consciousness naturally tend toward sympathy and impartiality, although it is not in the end a logically or metaphysically sufficient condition for that, nor even when it does generate moral principles does it in any sense guarantee morally correct action.

Piper's use of the word "sympathy" is not the only thing here that makes one think of Adam Smith as much or more than Kant; in *The Theory of Moral Sentiments*, Smith transformed

the view of the relation between reason and emotion proposed by his teacher Hutcheson and his friend Hume argued into the view that general principles sum up the particular responses of individuals with well-developed impartial spectators within their breasts.[17]   This is not intended as a criticism of Piper, but as a confirmation of the psychological rather than metaphysical interpretation of her argument to which I find myself forced.

In fact, it is my view that Kant himself was tempted at least sometimes in his career by an ultimately psychological as well as a more purely rationalist foundation for his approach to morality.[18] In his lectures on ethics in the form in which they were given from the mid-1770s until the publication of his critical works in practical philosophy, he presented "self-consistency" in the use of our "powers," "choice" or freedom itself as the "essential end" of mankind and as the basis of our duties, in the first instance to ourselves but also to others.[19] About the basis of this claim itself he is less than clear, saying only that freedom "is the highest degree of life." This sounds more like an empirical or psychological claim -- it is in the self-consistent exercise of freedom that we *feel* most fully alive -- but Kant's comment is hardly decisive. In Kant's earliest recorded thoughts about moral philosophy however, namely some of the notes he wrote in his own copy of his 1764 book *Observations on the Feeling of the Beautiful and Sublime*, Kant seems to have tried out both a psychological and a rationalist argument for the foundational status of the self-consistent use of freedom. He suggests a psychological argument in remarks like these: "There may quite well be stimulations that the human being prefers to freedom for a moment, but he certainly must feel sorry right after that,"[20] and "if I was free before nothing can present s more dreadful prospect of sorrow and despair to me than that in the future my state shall not reside in my own will, but in the will of another."[21]   These sorts of remarks suggest that anyone who has ever experienced freedom will *like* it, and will therefore be disposed to preserve it from being undermined from intrapersonal inconsistency in its use as well as from abridgement by others. But Kant also uses different language when he says that "In subjection there is not only something externally dangerous but also a certain ugliness and a contradiction that at the same time indicates its unlawfulness. . . .But that the human being himself should, as it were,

need no soul and have no will of his own, and that another soul should move my limbs is absurd and perverse."[22] "Absurd" sounds like a logical rather than psychological term, and "contradiction" surely is; thus Kant seems to be saying that there is a logical absurdity, a violation of the principle of (non-)contradiction, that is, the principle that a contradiction is always false, in treating a being with a will of its own -- another being, but for that matter oneself as well -- as if it did not have one. Kant's thought seems to be that to *act* toward a being with a will of its own as if it did not have one is also to *assert* both that it has a will and that it does not, thus to assert a self-contradiction, the most fundamental violation of reason that is possible.

This is just suggested by a passing remark in some notes written for Kant's private use some time after but presumably not too much after 1764. But I think this way of thinking pervades Kant's thought even when it appears in different language, for example in the *Groundwork*'s remark that the "nature" of human beings "marks them out as ends in themselves."[23]   I will not try to provide further evidence for this interpretation here. And I will only mention that Kant's own conception of reason or rationality involves a second main principle in addition to the principle of non-contradiction, namely the principle of sufficient reason and its associated idea of pure reason, the idea of the unconditioned. These are also applied in a variety of ways in the foundation of his moral philosophy. The most obvious application of the idea of the unconditioned is in his argument for the necessity of the highest good as the object of morality, where Kant appeals to the idea that the unconditioned value of virtue (a good will) must condition all striving for happiness, but also at least tacitly appeals to this idea to ground the conception of the highest good as the greatest happiness possible *throughout the world* consistent with the greatest possible virtue, or the non-selfish conception of the highest good that he clearly expounds in the 1793 essay on "Theory and Practice."[24] Kant could also have used the principle of sufficient reason to justify the requirement of universalizability, through the idea that there should be no difference in how persons act or are treated -- what maxims they adopt or are applied to them -- for which there is no reason; of course, that would have to be supplemented with the premise that their common humanity or

personhood entails that they should all be treated the same with regard to the possibility of their exercise of freedom, their freedom to set their own ends, unless they have themselves done something to ground the rationality of differential treatment. Kant does not formulate such a principle in so many words, but he does assume such a canon of rationality in his exposition of the innate right to freedom in the Introduction to the Doctrine of Right in the *Metaphysics of Morals* when he asserts that this right includes "being a human being *beyond reproach*, since before he performs any act affecting rights he has done no wrong to anyone":[25] his assumption is that a right claimed for oneself cannot be denied to anyone else without a specific reason for so doing, one that should further be generated by the other and not by oneself. No doubt some further argument would be necessary to fully justify this inference. But the point here is that Kant does suggest a style of argument for the impartial treatment of others, or the equal treatment of oneself and others, that is the essence of morality in general and right in this particular case that is grounded in canons of rationality rather than just psychology, but canons that go beyond the mere requirement of non-contradiction to include the principle of sufficient reason as well. To the extent that Piper develops her Kant-inspired model of rationality on the basis of the principle of non-contradiction or consistency alone, she may be attempting something even more challenging than Kant himself attempted, and the difficulty of getting everything desired for a moral theory from the principle of non-contradiction might be part of what pushes her toward a more psychological and less purely logical and metaphysical approach than perhaps she originally intended.

My suggestion that Piper might have used aspects of Kant's complex conception of rationality beyond the principle of non-contradiction to her own advantage would need a great deal more development to be convincing. Be that as it may, my argument that there is much in Piper's work that should be regarded as moral psychology rather than metaphysics should hardly be taken as a criticism of her work. Those who are inclined to approach moral theory from a logical and metaphysical angle will find much to think about and profit from in her work, particularly in her detailed demonstration that the ordinary logic of theoretical reasoning applies

to intentions as well as to theoretical propositions. But those interested in moral psychology will also find a tremendous amount to think about and learn from in her detailed development of the conceptions of interiority on the one hand and pseudorationality on the other. Piper understandably wants logic and psychology to stay together, but without denying the desirability of that I will say that her work on moral psychology in *Rationality and the Structure of the Self* is as deep and important as any known to me. Piper's *magnum opus* has something for everyone interested in moral theory, and deserves the most careful attention from all philosophers.

Notes

[1] ISBN 978-3-9813763-2-6 and -3-3. Available at the website of the Adrian Piper Research Archive, Berlin, http://www.adrianpiper.com/rss/index.shtml.

[2] David Hume, *Enquiry concerning the Principles of Morals*, ed. Tom L. Beauchamp (Oxford: Clarendon Press, 1998), Section 7, p. 63.

[3] For the distinction between moral theory, or what we would now call meta-ethics, and practical ethics, a philosopher's catalogue of particular duties and/or virtues, see Colin Heydt, *Moral Philosophy in Eighteenth-Century Britain: God, Self, and Other* (Oxford: Oxford University Press, 2018).

[4] Christine M. Korsgaard, *The Sources of Normativity* (Cambridge: Cambridge University Press, 1996), chapter 3, and *Self-Constitution: Agency, Identity, and Integrity* (Oxford: Oxford University Press, 2009), especially chapter 4.4.

[5] It should also be noted that although Piper's view might be considered part of a family that also includes those of Onora O'Neill as well as Korsgaard, Piper does not use the term "constructivism" as a designation for her position. Rather, she designates her position as a species of "Rationalism." This would be because she does not consider the fundamental principles of rationality as such as something "constructed," but as the necessary conditions of coherent thought in general. Just what the relation of those conditions to flesh-and-blood human beings is will be an issue in what follows. For some more of my own reaction to "constructivism," see my "Constructivism and Self-Constitution," in Mark Timmons and Sorin Baiasu, eds., *Kant on Practical Justification: Interpretive Essays* (New York: Oxford University Press, 2013), pp. 176-200.

[6] Kant, *Groundwork for the Metaphysics of Morals*, 4:391; in Kant, *Practical Philosophy*, edited and translated by Mary J. Gregor (Cambridge: Cambridge University Press, 1996), p. 47.

[7] Kant, *Critique of Practical Reason*, 5:15-16; *Practical Philosophy*, p. 148.

[8] See Kant, *Critique of Practical Reason*, 5:29-30, 5:47-8; *Practical Philosophy*, pp. 163, 178.

[9] I am thinking particularly of Korsgaard, *Self-Constitution,* chapter 8.

[10] There is a large debate about "conceptualism" versus "non-conceptualism" in contemporary Kant scholarship, which concerns the question whether non-conceptualized intuitions can be objects of awareness. I think that this debate turns on phenomenological considerations alien to Kant, but cannot get into that here. I do think that both sides to this debate will accept that in some form intuitions are objects of awareness and reject the view that judgments are the most basic and immediate form of awareness for Kant, although agreeing that form him both intuitions and concepts can yield knowledge only when used in judgments.

[11] Kant, *Critique of Pure Reason*, A 100-102; in Immanuel Kant, *Critique of Pure Reason*, ed. Paul Guyer and Allen W. Wood (Cambridge: Cambridge University Press, 1998), pp. 229-30.

[12] Piper is referring to the view developed by Edward McClennen in *Rationality and Dynamic Choice: Foundational Explorations* (Cambridge: Cambridge University Press, 1990), and "Pragmatic Rationality and Rules," *Philosophy and Public Affairs* 23 (1997): 210-58.

[13] See Joseph Butler, *Fifteen Sermons Preached at the Rolls Chapel*, ed. David McNaughton (Oxford: Oxford University Press, 2017), Preface and Sermon 1, especially pp. 13-14 and 18-19.

[14] Kant, *Metaphysics of Morals*, §27, 6:451; *Practical Philosophy*, p. 570.

[15] Kant, *Groundwork*, 4:424; *Practical Philosophy*, p. 75.

[16] See Paul Guyer, *Kant's* Groundwork for the Metaphysics of Morals: *A Reader's Guide* (London: Continuum [now Bloomsbury], 2007), pp. 124-6.

[17] See Adam Smith, *The Theory of Moral Sentiments*, sixth edition, ed. D.D. Raphael and A.L. MacFie (Oxford: Clarendon Press, 1979, reprinted Indianapolis: Liberty Fund, 1982), Part III, Section 4, Chapter 7, especially pp. 159-61.

[18] See especially my 2011 American Philosophical Association Eastern Division Presidential Address, "A Passion for Reason: Hume, Kant, and the Motivation for Morality, reprinted in my *Virtues of Freedom: Selected Essays on Kant* (Oxford: Oxford University Press, 2016), pp. 201-15.

[19] Immanuel Kant, *Moral Philosophy Collins*, 27:343-6; in Kant, *Lectures on Ethics*, ed. Peter Heath and J.B. Schneewind, tr. Peter Heath (Cambridge: Cambridge University Press, 1997), pp. 124-7

[20] Immanuel Kant, *Observations on the Feeling of the Beautiful and Sublime and Other Writings*, ed. Patrick Frierson and Paul Guyer (Cambridge:Cambridge University Press, 2011), p. 130.   Or "There may well be attractions that a person prefers to freedom for a moment, but this must make him sorry in the end," from Kant, *Notes and Fragments*, ed. Paul Guyer, trans. Curtis Bowman, Paul Guyer, and Frederick Rauscher (Cambridge: Cambridge University Press, 2005), p. 13.

[21] Kant, *Observations*, p. 128; *Notes and Fragments*, p. 12..

[22] Kant, *Observations*, p. 129; *Notes and Fragments*, p. 12..

[23] Kant, *Groundwork*, 4:428; *Practical Philosophy*, p. 79.

[24] Kant, "On the Common Saying: That May be Correct in Theory, but It is of No Use in Practice," 8:279-80n; *Practical Philosophy*, pp. 282-3.

[25] Kant, *Metaphysics of Morals*, Doctrine of Right, 6:238; *Practical Philosophy*, p. 394.

Works Cited