23 October 2018

ADRIAN PIPER | Rationality and the Structure of the Self (2nd ed.) | APRA Foundation Berlin 2013

These two sets of comments on Volume II of my *Rationality and the Structure of the Self*[1] (henceforth RSS II), from the two leading philosophers in their respective areas of specialization – Kant scholarship and decision theory – are the very first to appear from any quarter within academic philosophy.[2] My gratitude to Paul Guyer and Richard Bradley for the seriousness, thoroughness and respect with which they treat RSS – and my admiration for their readiness to acknowledge the existence of books that in fact have been in wide circulation for a long time – know no bounds. Their comments and criticisms, though sharp, are always constructive. I take my role here to be to incorporate those comments and criticisms where they hit the mark, and, where they go astray, to further articulate my view to meet the standard of clarity they demand. While Guyer's and Bradley's comments both pertain to the substantive view elaborated in RSS II, my responses often refer back to the critical background it presupposes that I offer in RSS *Volume I: The Humean Conception* (henceforth RSS I). I address Guyer's more exegetically oriented remarks first, in order to provide a general philosophical framework within which to then discuss the decision-theoretic core of the project that is the focus of Bradley's comments.

I. RSS II Is Not A Species of Kantian Constructivism

I am deeply indebted to Paul Guyer for bringing to the reader's attention a range of issues raised by RSS II. My reply to his comments indicates that we are in agreement about many more of these issues than he supposes. I focus here on his valiant attempt to locate the view defended in RSS II within the prevailing taxonomy of contemporary metaethics. Guyer's comments afford me a welcome opportunity to distance my view in every significant respect from any of the existing templates on that familiar map. My view is neither Kantian Constructivist, nor Deductivist, nor Instrumentalist (much less deontological or consequentialist). In RSS II, I characterize my view as *Ur-Kantian* because it is grounded more closely in Kant's texts than any of these other contemporary views. But I distance mine even from those texts in significant respects.

Guyer begins by suggesting that the substantive view I develop in RSS II is a "versio[n] of … contemporary 'Kantian constructivism'." But this term is a misnomer for everyone but Rawls. Consider the definition of Kantian constructivism that Rawls offered in the Dewey Lectures, where he first coined that term:

> [1] What distinguishes the Kantian form of constructivism is essentially this: it specifies a particular conception of the person as an element in a reasonable procedure of construction, the outcome of which determines the content of the first principles of justice. [2] Expressed another way: this kind of view sets

[1] 2008; second edition 2013. Berlin, Germany: APRA Foundation Berlin.

[2] Both volumes are open access and available for download at http://www.adrianpiper.com/rss/index.shtml . As an additional free bonus, you get a video interview with jokes in it.

> up a certain procedure of construction which answers to certain reasonable requirements, and within this procedure persons characterized as rational agents of construction specify, through their agreements, the first principles of justice. … [3] The leading idea is to establish a suitable connection between a particular conception of the person and first principles of justice, by means of a procedure of construction. [4] In a Kantian view the conception of the person, the procedure, and the first principles must be related in a certain manner – which, of course admits of a number of variations.[3]

Originally, Rawls's main aim was to stipulate an intrinsic connection between a value-neutral conception of the person and a value-neutral decision procedure such that "within this procedure persons characterized as rational agents of construction specify, through their agreements, the first principles of justice [2]." The term 'procedure of construction' occurs, undefined, in three of the above four sentences, and the term 'procedure' in all four of them. The 'procedure' to which Rawls repeatedly alludes in this passage, and whose outcome "determines the content of the first principles of justice [1]," is the decision procedure he began to develop very early on in his 'Outline of a Decision Procedure for Ethics.' Although the details of this procedure undergo change over time, his aim for it until late in his life remained "to establish the objectivity of moral rules, and the decisions based upon them … the decision procedure, which can be shown to be both reasonable and reliable, at least in some cases, for deciding between moral rules and lines of conduct consequent to them." [4] And he reminded us in *A Theory of Justice* of the analogy he meant to draw with the decision procedure followed in social theory, in which "rational individuals with certain ends … are to choose among various courses of action … [and] …[w]hat these individuals will do is then derived by strictly deductive reasoning from these assumptions about their beliefs and interests, their situation and the options open to them (119)." Here, too, in constructing the metaethical foundations of a normative theory of justice, he maintains, "[t]he argument aims eventually to be strictly deductive (121)."[5] Thus despite the many differences among these texts, Rawls's position in all of them is that Kantian constructivism aims to justify basic normative principles as the necessary outcome of an underlying value-neutral decision procedure, enacted by rational agents, from which those first principles can be derived.

Rawls's account of Kantian constructivism is thus recognizable as a species of *Deductivism*, the method of moral justification that attempts to derive normative moral principles from a conceptual analysis of value-neutral premises that are presumed to generate it. I critique the general strategy of Deductivism at length in RSS I, Chapter IX, Section 2. Thomas Nagel, Alan Gewirth, Rawls, and Richard Brandt, in addition to Kant, are all Deductivists; I examine their views in RSS I, Chapters VII, IX.3, X and XI respectively. But Rawls's particular kind of Deductivism, which he calls 'Kantian constructivism,' is distinguished by the nature of the premises he stipulates, namely the idealized rational agents, together with the constraining conditions, quoted above, under which they choose. These premises

[3] John Rawls, 'Kantian Constructivism in Moral Theory,' The Dewey Lectures 1980. *The Journal of Philosophy LXXVII*, 9 (September 1980), 516-517.
[4] John Rawls, 'Outline of a Decision Procedure for Ethics,' *The Philosophical Review LX*, 2 (April 1951), 177.
[5] John Rawls, *A Theory of Justice* [First Edition] (Cambridge, Mass.: Harvard University, 1971).

identify Rawls's view as that subcategory of Deductivism in which the premises define a value-neutral decision procedure, utilized by rational agents, that is instrumental to the generation of foundational normative principles as the necessary result of that procedure, i.e. as *Instrumentalist*. I discuss Instrumentalism in general in RSS I, Chapter IX, Section 4; and Rawls's Instrumentalism in particular in RSS I, Chapter X. So a view that does not justify its normative first principles as derivable from a rational agent's decision procedure is not an Instrumentalist view, and therefore not a species of Kantian constructivism as Rawls intended that term to be used. Rawls's explicit criteria, and his emphasis on the centrality of a decision procedure to generating normative principles, exclude most self-identified Kantian constructivists, aka New Kantians, from their scope.[6]

However, there are other views that satisfy the Instrumentalist criteria Rawls enunciated for Kantian constructivism to which he also would have denied that label. Richard Brandt's 'rational desire' theory of moral justification deduces moral principles as the necessary outcome of choices made by fully informed rational agents who have undergone cognitive psychotherapy. I examine Brandt's view in RSS I, Chapter XI. Moreover, there are other varieties of Deductivism that are closer in spirit to Rawls's avowed Kantian loyalties, even though they violate his stated Instrumentalist criteria. Alan Gewirth's method of moral justification, which attempts to deduce moral first principles directly from the concept of action, would be one; I discuss Gewirth's arguments in RSS I, Chapter IX, Section 3. So the term, 'Kantian constructivism' does not delineate the "family of views" that Rawls claimed and to which Guyer refers.

II. RSS II Is Not A Species of Deductivism

What all of these views do have in common is their attempt to establish a deductive relationship between metaethical premises and normative conclusions – but not always specifically between an agent defined by the procedural guidelines of rational choice and the moral principles such an agent would choose. Rather, the deductive relationship is supposed to hold more generally, between some stipulated foundational concept related to agency and the resulting moral principles. It is the aspiration to that relationship of logical derivation of substantive moral principles from universal, value-neutral premises that unites those of Rawls's students known as Kantian constructivists. Rawls never explicitly disowned Kant's agenda of demonstrating a necessary relation of logical derivation between the universal requirements of reason and the particular deliverances of moral decision-making;[7] and those of his students who identify themselves as Kantian constructivists have not done so, either.

I am not one of those students. I do explicitly disown this agenda. Guyer ascribes to me the views that "commitment to morality is implicated as the condition for including intentions and choices in a self that is unified by the conditions of

[6] I offer a more thorough critique of the tenets of New Kantianism in 'Kant's Self-Legislation Procedure Reconsidered,' *Kant Studies Online* 2, 4 (October 2012): 203-277 and http://www.adrianpiper.com/docs/AMSPiperKantsSelfLegislationProcedureReconsideredKantStudiesOnlineVol2No14_20Oct2012_203-277.pdf .

[7] - although he wasted a great deal of ink deprecating his own attempt in *A Theory of Justice* to fulfill its demands. See his 'Justice as Fairness: Political not Metaphysical,' *Philosophy and Public Affairs 14*, 3 (1985), 223-251.

rationality as such;" and that "the conditions for unified selfhood provide the conditions for the development of moral emotions, and these lead to moral principles." But even if these views accurately summarized my claims, neither of them would describe deductive relationships. That P is implicated as a condition for Q does not mean that P implies Q. And that the conditions for P provide the conditions for Q, which leads to R, does not mean that P, or even P and Q, necessarily lead to or imply R. The necessary relation of logical derivation fails to hold in both cases.

In fact I would deny that a commitment to morality is a precondition for including intentions and choices in a rationally unified self, although it may well be an empirical outcome of that unity. On my view, rather, intentions and choices are part of a rationally unified self because and to the extent that – as on Kant's view – conscious experiences in general are part of a rationally unified self; and those intentions and choices, independently of their content, are conscious experiences. Similarly, I would deny that the development of moral emotions *necessarily* lead to any principles at all. An upbringing that develops only the moral emotions of shame and resentment may lead only to wholesale rejection of the very idea. For all of these reasons, I do not and could not identify my view as Kantian constructivist.

Moreover, my view diverges even further from the Instrumentalist variant of the Deductivist program: even were I to accept Rawls's insistence on the centrality of a deliberative procedure, which I do not, I in any case deny that there is any necessary connection between rational deliberation and moral choice. Rational deliberation consists in formal operations of the mind that can be applied to many different kinds of choice objects, contingently including but not limited to normative moral principles. In RSS I, Chapters IX.2, IX.4 and X, I point out that only by packing the premises of that deliberation with the desired moral assumptions can they then be derived as conclusions.[8] And in RSS II, Chapters IV and V, I often state that there is nothing about the logical function of universalization by itself that of necessity forces a moral conclusion – or, therefore, a moral choice. So, for example, in my discussion of Ned McClennen's concept of resolute choice, I conclude a critique of Kant's decision procedure by observing that

> Kant's problem is that simple universalization of a maxim is by itself too weak a criterion to rule out all cases of false promising, lying, suicide, self-neglect, or any other practices that are, in their most general descriptions, prima facie morally unacceptable. His rationalist disdain for "any law determined by particular actions" leads him to adopt as law a principle so weak and comprehensive in scope that it potentially justifies virtually all actions given sufficient specification of the circumstances. It is not fine-grained enough to distinguish between those specific actions that really are justifiable and those which are not (RSS II: 182).

Similarly, in the first paragraph of RSS II, Chapter V, 'How Reason Causes Action,' I explicitly disown any such Deductivist intentions (RSS II: 188). And throughout that chapter, I distance my own project repeatedly from that one (cf. RSS II: 189, 208, 218,

[8] In 'Kant's Self-Legislation Procedure Reconsidered,' (*op. cit.* Footnote 6) I join those who deplore Kant's own commission of what logicians now know as Mill's Paradox of Inference and what Jaakko Hintikka calls the 'scandal of deduction' (in his *Logic, Language-Games and Information: Kantian Themes in the Philosophy of Logic* (New York: Oxford University Press, 1973 and 2002), p. 222)), aka the 'garbage-in garbage-out' problem.

222, and 227). So to all three of Guyer's questions,

> [H]ow does Piper herself propose to make the transition from unified selfhood to full-blown morality? … [D]oes acceptance of the requirements of unified selfhood itself entail commitment to the interpersonal principles of morality as normatively understood? Is unified selfhood a sufficient as well as a necessary condition for commitment to morality?

my answer would be in the negative. I do not propose to make any such transition from unified selfhood to normative morality, and unified selfhood on my view does not entail commitment to normative morality. As Guyer himself then observes, "[T]he conditions for unified selfhood … do not guarantee that even interiorized agents will always act in accordance with these principles."

My view is rather that the canons of classical predicate logic structure all of the conscious experiences – including the beliefs, perceptions, desires and intentions – of a unified self; and that only such a rationally unified self is capable of formulating coherent moral intentions. But this does not imply that a rationally unified self must formulate them. Whereas the structure of an intention is supplied by the logical preconditions of thought, its content is always a function of empirical circumstance. So a particular, fully formed moral intention always *instantiates* the rational structure of intention, but can never be logically *derived* from it. My project is more modest than the Deductivist's, and therefore has a correspondingly greater chance of success.

In light of my wholesale rejection of the Deductivist project as misguided and futile, Guyer is right to distinguish my view from that of Christine Korsgaard, who "succumb[s] to the fantasy of assuming that … [we can] explain moral failure only as defective reasoning about what morality requires rather than as a direct choice not to be moral." But he then goes on to interpret my distinction between psychological consistency as a contingent preference and metaphysical consistency as a necessary precondition for having contingent preferences – or indeed, unified intentional states of any kind – as in some sense similar to Korsgaard's profoundly confused neo-Deductivist claim that "an agent that is not a mere heap … cannot but act rationally, although it might fail to correctly infer what is rational in some particular case." The suggestion that not being a mere heap implies rational action is an amusing one: picture a neatly sorted, packed, taped, bagged, labeled, and stacked heap of garbage, sitting at a computer in readiness for the 'garbage-in' stage of the Deductivist procedure, evaluating the choice options in a normal form decision tree. But even more amusing is the failure it expresses to grasp my view. In an anonymous reader's report on RSS II of 2003 for Oxford University Press that bears a remarkable resemblance to Korsgaard's distinctive manner of thinking and writing, the reader tasks me with the same project as does Guyer, of "explaining how the ideal of perfect practical rationality must motivate actual human beings in a *morally* normative way."[9] But we have just seen above that RSS II repeatedly disowns that project as a waste of time. In case the foregoing examples did not suffice, its futility

[9] This patently substandard OUP reader's report of 2003 on my manuscript (the hard copy of which the reader declined to return) is available with my original annotations upon request, for personal use only, from archive@adrianpiper.com . Another amusing read, it commits exactly the same elementary errors of reasoning that later surface in Korsgaard's *Self-Constitution: Agency, Identity, and Integrity* (Oxford: Oxford University Press, 2009), which appeared one year after the first edition of *Rationality and the Structure of the Self*, and which I warmly welcome as a recent addition to her ongoing *Kant for Dummies* pamphlet series.

is amply illustrated by Korsgaard's own failed attempts to realize it, in which this vague and bloated 'must' of metaphysical necessity is supposed to pack the same browbeating wallop as does the notion of a ground project for Bernard Williams or the notion of a second-order desire for Harry Frankfurt (I criticize both of these views in RSS I, Chapter VIII, Sections 2 and 3). Her approach applies the Deductivist methodology of Alan Gewirth, her former senior colleague at the University of Chicago, to whose *Reason and Morality* her arguments are clearly indebted. As I discuss in detail the failings of Gewirth's view in particular in 'Gewirth's Deductivism,' cited above, there is no need to dwell at greater length on those of his disciples.

The problem inherent to all such views is that if you duck the punch, you deflect the wallop. Guyer rightly complains that

> 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.

This brand of 'metaphysical necessity' is obscure to me, too. That is why I unconditionally reject it, and address the failings of this general approach, in the chapters and sections from RSS I listed above. As I join Guyer in rejecting this useless notion, his binary distinction between metaphysical necessity and moral psychology as a taxonomic device for identifying RSS II as a normative moral theory would require, at the very least, a respectably rehabilitated account of metaphysical necessity in order to carry any weight. As he does not offer such an account, I also reject that taxonomy.

III. RSS II Is Not A Species of Normative Moral Theory

My view has nothing whatsoever in common with the Deductivist tactics outlined above. Guyer's complaint does not apply to my argument, because mine does not invoke any such unexplicated notion of metaphysical necessity. In my usage, the term 'metaphysical' means what it standardly means according to the OED, i.e. it refers to explanatory first principles, preconditions or premises that a particular state of affairs presupposes. These may be material or nonmaterial, concrete or abstract, specific or general, as the case may be. My claim that a unified and internally consistent self is a metaphysical precondition for having psychological likes and dislikes is no more obscure than claiming that a magnetic field is a metaphysical precondition for a magnet's attracting some objects and repelling others. In both cases, the former is an explanatory premise of the latter. Thus there is no suggestion in my argument that any obscure notion of metaphysical necessity functions to define ideal agency, flesh-and-blood agency, or the relation between them.

Guyer acknowledges my thesis that "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 (RSS II, Chapter IV: 177; also see 198, passim)." Yet he then immediately goes on to ascribe to me the contradictory commitment, "actually to a normative position that unified selfhood is a good that

human beings ought to strive to realize etc." However, none of the passages Guyer cites in support of his claim that my philosophical commitment is the exact opposite of that which I explicitly defend at very great length in fact support his interpretation of RSS II as a normative moral theory. At best they show that metaphysical consistency is compatible with physical embodiment. But on my view, that is not in dispute.

(1) For example, he lifts out of context my example of Myron the puffy glutton, which occurs in the course of my argument that Ned McClennen's concept of resolute choice is materially equivalent to my concept of a genuine preference. This example directly follows and illustrates my thesis that the

> claim is not about a contingent psychological preference for consistency, but rather about the necessary 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 (RSS II, Chapter IV: 177).

About the subsequent discussion of Myron's gluttony, Guyer comments that he finds it difficult to read that 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 etc." My statement that "Myron chooses to avoid this condition of disconnected bewilderment etc." certainly does imply that Myron has a choice etc. But Guyer conflates Myron's choice to avoid a state of disconnected bewilderment with the highest-order disposition to metaphysical consistency it presupposes, and that transforms that choice option into a motivationally effective preference. And his observation that no metaphysical necessity is involved, with which I of course also agree, conflates this cryptic Deductivist notion of metaphysical necessity which we both reject with the more robust concept of the metaphysical consistency which I claim any such choice presupposes.

(2) Similarly, Guyer takes my carrot-stick metaphor for the disposition to metaphysical consistency to show that I am really talking about a "psychological characteristic of a flesh-and-blood agent etc." – as though metaphysical consistency could only be a transcendent characteristic of nonmaterial noumenal agents. Just as flesh-and-blood agents have a hard-wired predisposition to physical entropy, they can also have a hard-wired predisposition to metaphysical consistency. My thesis is that they must, to the extent that they are able to form coherent intentional states at all.

(3) Furthermore, he rightly finds in my sentinel metaphor for the highest-order disposition to literal self-preservation a "causally efficacious psychological disposition," but then wrongly infers from this that this disposition therefore must be a psychologically transient inclination rather than an enduring metaphysical precondition for forming inclinations of any kind. However, we have just seen that flesh-and-blood agents can have causally efficacious psychological dispositions that are enduring and hard-wired rather than transient. Other examples would include the psychological dispositions to react to stimuli, to explore the environment, or to ignore anomaly. All can and do serve as metaphysical preconditions for more specific states or behaviors – such as, respectively, to answer the telephone, to open a closed door, or to overlook E.T. crouching in the corner.

(4) Finally, Guyer claims that I

> describ[e] 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" (RSS II:209), ...

But this claim is false. The quoted terms are lifted out of a passage analyzing Marcia Baron's distinction between primary and secondary motives, both of which would count as occurrent psychological states on my view. That passage has nothing to do with a commitment to coherence. It has even less to do with the normative moral theory Guyer groundlessly ascribes to me. He then runs this passage together with one in a section occurring much later in the text, on a completely different topic. There I describe a utility-maximizing ideal of rational motivation that I explicitly criticize as inadequate. Thus he continues:

> … success at which depends upon a variety of circumstances, including a "community of spontaneous agents" with adequate "material resources" (RSS 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,

now lifting these quoted terms out of a different context that bears no relation to my explication either of Baron or of my contrasting view. Finally, Guyer then splices these latter terms back into my earlier analysis of Baron's view, and wrongly ascribes it to me:

> 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" (RSS II:209) (ibid.); 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.

However, I say this reasonable thing about Baron's view, not my own. Guyer's assemblage of unrelated passages critiquing other people's views would not be relevant to a defense of his claim that mine is normative. Actually he offers no such defense.

I infer from exegeses (1)-(4) that Guyer is using the term 'metaphysical' to refer to states of affairs that are necessarily nonphysical and nonmaterial – i.e. beyond the scope of the empirical world as Kant defined it in the *Critique of Pure Reason*. This use of the term conforms well to Kant's in his critique of speculative metaphysics in the Antinomies. But it does not conform to mine. It is a straw man, cobbled together out of disparate and irrelevant passages, that Guyer and I both agree is implausible.

As the view I defend in RSS II is not in fact a normative theory, nothing could show that it is. I offer a taxonomy that accurately locates that view as a descriptive and explanatory metaethical theory – for example, at RSS II pages 18, 27, 33, and 48; and at page 221, where I repeat a point from RSS I, Chapter III that really cannot be repeated often enough:

> [T]he question whether a theory is explanatory or not can be answered independently of the question whether it has a normative or a descriptive metaethical status. The metaethical status of any principle is fully exhausted by specifying the relation between two descriptive versions of it: that which

> describes actual behavior and that which describes ideal behavior. A theory can be both explanatory and normative if it explains the behavior of an ideal agent who sets a standard we are exhorted to emulate (RSS II: 221).

The ideal agent I describe in the first half of RSS II does not set a standard we are exhorted to emulate, because it is not the kind of ideal an agent can aspire to emulate. So getting bogged down in searching for someone to normatively exhort would have delivered even less bang for the buck than getting bogged down in the flatulent cloud of metaphysical necessity. A careful reading of RSS II will confirm that it makes neither of these mistakes.

IV. Genuine Preferences Are Logically Consistent

Richard Bradley's comments are focused on RSS II Chapter III, 'The Concept of a Genuine Preference,' in which I defend the thesis that, like all other conscious intentional states of a rationally unified self, preferences – understood in the broad, decision-theoretic sense as encompassing intentions and resolves in addition to desires – are similarly structured by and subject to the requirements of classical predicate logic; and therefore that rational preferences must meet exactly the same constraints of logical consistency. I am relieved and grateful to have his endorsement of my claim that preferences express categorical judgments. I now explore some of the implications of that claim. I argue here that Bradley cannot consistently accept that claim yet deny that intransitive preferences violate the requirement of logical consistency even when the agent's preferences are incomplete.

1. The Symbolic Representation of a Pairwise Comparison

A categorical judgment ascribes a predicate to a subject. A notation adequate for symbolizing a categorical judgment must expose the subsentential relationship between subject and predicate. Classical first-order predicate logic represents and exposes this subsentential relationship in the formula Fx. However, understanding a binary categorical judgment of preference in particular requires more syntax than this; for in its traditional form, classical predicate logic does not expose enough of that substructure to distinguish between a judgment of preferences involved in pairwise comparisons and other categorical judgments. A *pairwise comparison* is a binary categorical judgment of the form

(1) Agent S prefers option x to option y.

Traditional predicate logic would express this judgment as

(2) Sxy

But (2) does not suffice to distinguish an intensional judgment of preference involving a pairwise comparison from other relational categorical judgments of an extensional nature, such as

(3) x is y's sister.

The traditional semantics of classical predicate logic permits (1) and (3) as two conflicting interpretations of (2), but its syntactical cache is not rich enough to

represent the structural difference between them. I defend this claim at length in RSS II, Chapter III, Section 1, 'A Problem about Cyclical Inconsistency.'

Unfortunately, the notation of standard Ramsey-Savage decision theory is even less adequate to this challenge, because it not only disregards the intensional nature of preference judgments. It fails to represent the subject-predicate substructure of preference judgments at all. Instead, it treats pairwise comparisons as though they were extensional mathematical formulas. In these formulas, preference alternatives are isolated from the contextualizing intentional attitude in which they are nested, represented as though they were extensional objects like real numbers, and ordered by the extensional mathematical connective '>,' thus:

(4) $x > y$

In mathematics, (4) would symbolize the natural language sentence,

(5) x is greater than y.

'Greater than' is an extensional phrase that relates two numerical values. In traditional Ramsey-Savage decision theory, (4) symbolizes the sentence,

(6) x is preferred to y.

'Preferred to' is an intensional phrase ranging over two intentional objects. But the passive grammatical construction of (6) obscures its intensionality by eliminating mention of the agent S whose intentional attitude takes x and y as objects. Armed with the appearance of extensionality, Ramsey-Savage decision theory constructs from a sequence of such pairwise comparisons an ordinal ranking

(7) $x > y > z \ldots$

on the basis of which mathematical calculations of subjective utility are then to be performed.

But this won't do. (7) is a rank-ordered series of intensional objects of someone's intentional attitude, whether their intensional status is visible or not. In (7), it is not visible. The objects x, y and z are abstracted from their necessary dependence on that attitude. But without a modifying phrase that expresses it, the intentional objects nested in it cannot be related using an extensional connective. I can relate

(8) I prefer that x

to

(9) I prefer that not-y

using the Boolean connective 'and' because nesting x and y in their respective intentional attitudes extends to them the propositional status of the intentional attitude in which they are nested. Eliminate those nesting intensional contexts and

there is nothing for an extensional connective to connect. Isolated intentional objects are not the kind of entity that can have extensional propositional status.

Similarly, you cannot connect a hope to a dream using telephone cables. You cannot line up a memory of summer camp with a memory of college with a memory of internship along an interstate highway. You cannot link up an imagined garden with an imagined lawn sprinkler using pliers and copper wire. *A fortiori*, you cannot connect isolated intensional objects of any kind, including preference alternatives, using extensional connectives of any kind, including mathematical ones. That the traditional notation of Ramsey-Savage decision theory renders the intensionality of such objects invisible does not render it nonexistent or unimportant. Rather, such a notation reveals its inadequacy to the requirement of veridical representation. It needs to be revised. I defend this claim at length in RSS I, Chapter IV, Section 2, 'The Ramsey-Savage Concept of a Simple Ordering;' and RSS II, Chapter III, Section 2, 'Savage's Concept of a Simple Ordering Reconsidered.'

Bradley offers a revision of that traditional notation that aims to expose the intensionality of categorical judgments involving pairwise comparisons by replacing the extensional mathematical connective '<' meaning 'less than' with the symbol '$\preceq$,' such that

(10) $X \preceq Y$

is interpreted as the weak preference relation

(11) Y is at least as preferable as X.

As with (6), 'at least as preferable as' is an intensional phrase that ranges over two intentional objects. But (11) obscures its intensionality even more than (6). In addition to its passive grammatical construction, (11) substitutes the modal locution 'as preferable as' for the indicative locution "is preferred to." But preferable to whom? The modal locution approaches extensional status only to the extent that something can be preferable independently of the existence of any possible subject to whom it is in fact preferred. As the concept of being preferable without being preferred to some possible subject is incoherent, the implicit intentional attitude cannot be eliminated. Hence there are no grounds for ascribing extensional status to Bradley's modal locution. (10) remains an intensional statement consisting in an intentional operator '$\preceq$' that governs a pair of intentional objects whose intensional status is obscured by the notation.

This means that whereas (4) purported to connect two intensional objects using an extensional operator, (10) in fact does not aspire to connect two objects at all, because '$\preceq$' is not a connective but rather an intentional operator that expresses a subject's intentional attitude. It does not contain the syntactical resources for rank-ordering its subsentential constituents. (10) denotes a relation between X and Y only in the semantic sense of (2). By itself – and like (2), (10) does not expose any syntactical substructure at all. How the two objects X and Y are related in (10) depends on the semantic interpretation of the expression, rather than – as in (4) and (7) – the syntactic ordinal substructure that the symbolization itself identifies. Then this interpretation provides no warrant for ascribing the ordinal ranking required for calculations of subjective utility to a series of such statements expressing this

intentional attitude toward pairs of intentional objects. And we have just seen that the isolated intentional objects themselves also fail on this count.

Bradley attempts to meet this objection by proposing a link between extension and intension that translates the extensional mathematical connective '<' into the intentional operator '$\preceq$', such that

(12) $X < Y$,

i.e.

(13) X is less than Y,

is defined as

(14) $(X \preceq Y) \,.\, \sim(Y \preceq X)$,[10]

i.e.

(15) Y is at least as preferable as X and X is not at least as preferable as Y.

Thus Bradley intends (14) to represent the same ordinal ranking of X and Y as does (12). However, they cannot represent the same ordinal ranking, because (14) lacks the equivalent syntactical resources to fix a relation between X and Y in both conjuncts. If (10) exposes no more subsentential structure than (2), then neither conjunct of (14) can connect two preference alternatives in any particular order at all, and therefore cannot connect them in the manner of (12). Neither conjunct of (14) exposes the syntactical substructure necessary for interpreting the parenthecized expression in the second conjunct as reversing a relation between X and Y symbolized in the first. There is no such extensional relation that can be presumed to hold for both. Thus (14) can be replaced by

(16) $P \,.\, \sim Q$

without loss of meaning. Categorical judgments of this description are *internally intransparent*. (14) does not contain the syntactical resources necessary for illuminating an ordinal relationship among those constituents because it is internally intransparent. The only structure Bradley's notation can expose is that supplied by the Boolean connectives that relate these internally intransparent judgments.[11] So (14) works neither as an intensional definition of (12), nor as a symbolic representation of (15).

However, (15) presents further intransparencies. Something cannot be at least as preferable as something else without there being some implied third thing to which the first two are at least as preferable. An apple can only be at least as preferable as an orange if there is also, for example, a prune, to which they are at least as

[10] As in RSS II Chapter III, I substitute Quine's aesthetically pleasing traditional notation for the more muscular, plumber's tool look of Bradley's '∧' and '¬.'

[11] Here I apply the same argument to Bradley's proposal as that directed at traditional Ramsey-Savage decision theory at RSS II: 112-115.

preferable. Thus (15) is grammatically incomplete without mention of the third term of comparison that reveals it to be a triadic rather than a pairwise comparison – and therefore unsuitable for purposes of this discussion.

None of these internal intransparencies would seem problematic for Bradley's assumption that

(17) $(X \precsim Y)$ and $\sim(X \precsim Y)$

are contradictory. That Sxy and $\sim Sxy$, or P and $\sim P$, or $(X \precsim Y)$ and $\sim(X \precsim Y)$ are contradictory certainly can be established with reference to the law of non-contradiction

(18) $(\forall x)\sim(Fx \,.\, \sim Fx)$

However, (18) is a premise of classical first-order predicate logic; and that premise is not among the axioms of Ramsey-Savage decision theory. In fact, (18) cannot be stated within the notational constraints of traditional Ramsey-Savage decision theory at all, and we have just seen that Bradley's replacement of '<' with '$\precsim$' does not address this problem. Without piecemeal importations of syntax from natural language, mathematics, and predicate logic,[12] Bradley would be unable even to formulate (17).

That absence in turn further undermines Bradley's claim, about an intransitive statement, that "someone who judges that

(19) $X \precsim Y$ and that $Y \precsim Z$ but not $X \precsim Z$

does not make any strictly contradictory judgments, and so cannot in my opinion be said to violate the law of non-contradiction." Like (14), (19) is similarly internally intransparent. It, too, could be replaced by a logically contingent expression

(20) P and Q but not R

without loss of meaning. As it stands, (19) does not expose enough sentential substructure to ascertain whether it conceals a contradiction or not. There is no way solely within the syntactical constraints of traditional Ramsey-Savage decision theory of ascertaining whether $(X \precsim Y)$ and $\sim(X \precsim Y)$ are contradictory or not; nor, therefore, whether someone who judges that $X \precsim Y$ and that $Y \precsim Z$ but not that $X \precsim Z$ makes any strictly contradictory judgments or not. On this count, (17) and (19) must stand or fall together. Like Davidson, McKinsey and Suppes and the many others

[12] So what is the matter with piecemeal importations from other symbol systems such as natural language, predicate logic, set theory, and mathematics into decision theory, you might ask? Unless this is done with the same rigor and elegance as Sen's analysis of preference in *Collective Choice and Social Welfare* (San Francisco: Holden-Day, Inc., 1970), they make it extremely difficult to construct a non-idiosyncratic and uncontroversial axiomatized system that, like first-order predicate logic, can be shown to be both consistent and complete. That is why contemporary decision theory has generated such a wide variety of kinds of consistency proofs, and why it is still so far from settling on an authoritative one. That is not going to happen until its intensional character is acknowledged and adequately formalized in its syntax.

who have subsequently addressed this problem,[13] Bradley must resort to the extrinsic resources of natural language even in order to state it. His notation, too, is not fine-grained enough to detect the differences, and the extensional mathematical conventions on which it relies obscure what the differences might be.

2. A Variable Term Calculus

For the reasons just described, neither classical predicate logic in its traditional form, nor traditional Ramsey-Savage decision theory, nor Bradley's proposed reform of it, suffice to expose the intensional and subsentential structure of binary categorical judgments expressing pairwise comparisons between preference alternatives. Without making this substructure visible, I would agree with Bradley that it is difficult to see in what sense preferences are subject to the same logical consistency constraints as other intentional attitudes; and therefore in what sense (19) might violate the law of non-contradiction. As it stands, (19) does not comprise the kind of judgments that can be contradictory.

The concept of a genuine preference that I defend at much greater length in RSS II Chapter III is designed to redress this problem. Rather than rehearse the details of that proposal here, I offer merely a brief summary and refer the reader to that and the preceding chapter for a fuller exposition. In RSS II Chapter II, I argue that all cognitively operative intentional states of a unified self must satisfy two rationality criteria, of horizontal and vertical consistency. *Horizontal consistency* applies the quantified law of non-contradiction to the subsentential constituents of categorical judgments. These *subsentential constituents* consist in the variable terms and *n*-adic predicates ascribed to them of sentences in predicate logic. *Vertical consistency* applies the quantified law of modus ponens to the inferential relationships among such constituents, such that the lower- and higher-order concepts that figure in such judgments are logically consistent with one another.

The *variable term calculus* introduced in RSS II Chapter III offers a syntax for structuring subsentential constituents as they appear in categorical judgments expressing intentional attitudes. I focus on preference in the broad decision-theoretic sense, and in particular on preference over multiple alternatives, because that is the most troublesome intentional attitude for a Kantian theory of action capable of anchoring a plausible and robust metaethical analysis. The basic idea is to extend the Boolean connectives and quantificational conventions of classical predicate logic to these constituents. I impose on such sentences the same familiar opacity constraints as in any intentional logic. The resulting apparatus preserves the traditional conventions of classical predicate logic, but also extends them to the subsentential structure of sentences expressing intentional attitudes in that logic. My proposal differs from other intentional logics in that it treats the symbol for preference, *P*, as a first-order predicate rather than an intentional operator. This enables me to more closely conform the variable term calculus to those well-established conventions, and to then express the consistency constraints on *n*-adic judgments of preference in terms of them.

RSS II, Chapter III, Section 5, walks the reader through the basics of this apparatus; and demonstrates how to navigate with it. Section 6 contrasts my approach with those developed by Mark Kaplan on the one hand and Richard

[13] Donald Davidson, J. C. C McKinsey, and Patrick Suppes, 'Outlines of a Formal Theory of Value, I,' *Philosophy of Science* 22, 2 (April 1955), 140-160. See the Bibliography of RSS II for a fuller list.

Jeffrey and Ethan Bolker on the other; and argues for the superiority of mine over both. Preference is inherently bound up with action, and therefore must be represented dynamically.[14] Hence Sections 7 and 8 formulate five further familiar criteria that such a genuine *n*-adic judgment of preference must meet: Asymmetry, Connectivity (basically equivalent to Bradley's Completeness), Irreflexivity, Transitivity, and Ordinality. I argue that these seven criteria constitute a conceptual truth about in what a genuine preference consists.

With that background in place, I now put this apparatus to work by showing how a modified version of Bradley's formulation (19) that is responsive to the criticisms made in Section IV.1 above can, in fact, be shown to violate the law of non-contradiction and therefore is subject to the requirements of logical consistency after all. We have just seen that (14) is not a satisfactory definition of (12), and that (15) is not a satisfactory expression in natural language of (14). Then (19) is not a satisfactory rendering of the intransitive preference ranking of *X*, *Y*, and *Z* that Bradley wants to claim does not violate the law of non-contradiction. For a more transparent and intuitive representation of Bradley's claim, I revert to a more familiar natural language rendering of an intransitivity among weak preferences that nevertheless both preserves its role as a relation among pairwise comparisons, and also exposes its intensionality using an active grammatical construction that explicitly expresses its concealed intentional attitude:

(21) Subject *W* prefers *X* to *Y* or is indifferent between them; and prefers *Y* to *Z* or is indifferent between them; but it is not the case that *W* prefers *X* to *Z* or is indifferent between them.

Then we can invoke the resources of the proposal developed in RSS II, Chapter III, Section 5 to symbolize (21) in a way that veridically represents both its intentional character and its subsentential structure. Given variables *w*, *x*, *y* and *z*, let *w* be a chooser and *x*, *y* and *z* be any alternatives – actions, states, events, gambles, compound lotteries, plans, prospects or discrete objects – among which that chooser decides. Define *w*'s strict preference *P* for *x* over *y* as

(22) $Pw(x \,.\, \sim y)$

and *w*'s indifference between *x* and *y* as

(23) $Pw(x \vee y)$

Then define *w*'s weak preference for *x* to *y* as

(24) $Pw[(x \,.\, \sim y) \vee (x \vee y)]$

(24) says literally that *w* is indifferent between either strictly preferring *x* to *y*, or being indifferent between them. I discuss the weak preference relation at greater length at RSS II: 131-133. Then (21) can be symbolized as

[14] I argue that judgments involving pairwise comparisons are inherently dynamic in 'The Money Pump is Necessarily Diachronic' (2014), at https://philpapers.org/rec/PIPTMP .

(25) $Pw[(x \,.\, {\sim}y) \vee (x \vee y)] \,.\, Pw[(y \,.\, {\sim}z) \vee (y \vee z)] \,.\, {\sim} Pw[(x \,.\, {\sim}z) \vee (x \vee z)]$ — Intransitive ranking of weak preferences

Section 6.2.1 of RSS II, Chapter III introduces a Quinean truth-table method for subsentential constituents nested in intentional attitudes that shows (24) to be a tautology:

(26)

Pw	*[(x*	.	~y)	∨	*(x*	∨	*y)]*
			F				**T**
		F				**T**	
				T			
T							

If (24) is a tautology, then (25) is a contradiction, and satisfaction of a completeness (or connectivity) condition is not required in order to show this. For convenience I represent the table for each conjunct separately:

(27)

Pw	*[(x*	.	~y)	∨	*(x*	∨	*y)]*
			F				**T**
		F				**T**	
				T			
T							

(28)

Pw	*[(y*	.	~z)	∨	*(y*	∨	*z)]*
			F				**T**
		F				**T**	
				T			
T							

(29)

~Pw	*[(x*	.	~z)	∨	*(x*	∨	*z)]*
			F				**T**
		F				**T**	
				T			
F							

The conjunction of (27), (28) and (29) constitutes a contradiction. So Bradley's claim that "someone who judges [in the manner of (25)] does not make any strictly contradictory judgments" is false.

Furthermore, that (25) violates the law of non-contradiction can also be brought out in a different way. The rule of transitivity T^3 for strict preference developed in RSS II, Chapter III, Section 7 can be reformulated for weak preference thus:

(30) $\{Pw[(x\ .\ \sim y)\ v\ (x\ v\ y)]\ .\ Pw[(y\ .\ \sim z)\ v\ (y\ v\ z)]\}$ → $Pw[(x\ .\ \sim z)\ v\ (x\ v\ z)]$ — Transitivity for weak preference

It is easy to generate a contradiction from the foregoing assumptions, again without requiring satisfaction of a completeness (or connectivity) condition:

(31)	$\sim Pw[(x\ .\ \sim z)\ v\ (x\ v\ z)]$	(25)
(32)	$\sim\{Pw[(x\ .\ \sim y)\ v\ (x\ v\ y)]\ .\ Pw[(y\ .\ \sim z)\ v\ (y\ v\ z)]\}$	(30), (31)
(33)	$\{Pw[(x\ .\ \sim y)\ v\ (x\ v\ y)]\ .\ Pw[(y\ .\ \sim z)\ v\ (y\ v\ z)]\}$	(25)
(34)	$\sim\{Pw[(x\ .\ \sim y)\ v\ (x\ v\ y)]\ .\ Pw[(y\ .\ \sim z)\ v\ (y\ v\ z)]\}\ .$ $\{Pw[(x\ .\ \sim y)\ v\ (x\ v\ y)]\ .\ Pw[(y\ .\ \sim z)\ v\ (y\ v\ z)]\}$	(32), (33)

If transitivity is, as I have argued in RSS II Chapter III, a requirement of logically consistent preferences, then (25) violates that requirement.

Both of these conclusions hold whether *w*'s preferences are complete or not. Nevertheless I think there is truth to be gleaned from Bradley's claim. For the fact that (25) is intransitive does not imply that it is cyclical. Unlike a cyclical ordering of strict preferences, (31) – the third conjunct of (25) – in fact does not express a determinate ranking of *x* and *z* at all. Rather, it is ambiguous between the cases in which agent *w* does not weakly prefer *x* to *z* because (a) she is strictly indifferent between them; (b) she weakly prefers *z* to *x*; (c) she strictly prefers *x* to *z*; (d) she is not offered a comparison between *x* and *z* in the first place. That intransitivity is not logically equivalent to cyclicity can be shown using the truth table method above. But I leave that as an exercise for the reader.

Adrian Piper did her dissertation on rationality with John Rawls at Harvard in 1981 and studied Kant and Hegel with Dieter Henrich at the University of Heidelberg in 1977-78. She taught at Harvard, the University of Michigan, Stanford, Georgetown, UCSD, and Wellesley. Her primary philosophical publications are in Kant, metaethics, and the history of ethics. Recently, she published her *Escape to Berlin: A Travel Memoir* http://www.adrianpiper.com/books/Escape_To_Berlin/index.shtml (APRA Foundation Berlin, 2018). Piper is also an artist.