

Intrinsic Value and the Supervenience Principle[†]—forthcoming in *Philosophical Studies*.

Dale Dorsey

Department of Philosophy
University of Kansas
1445 Jayhawk Boulevard
Wescoe Hall, rm. 3090
Lawrence, KS 66045
ddorsey@ku.edu

An important, and popular, constraint on the nature of intrinsic value—the “Supervenience Principle” (SP)—holds that some object, event, or state of affairs ϕ is intrinsically valuable only if the value of ϕ supervenes entirely on ϕ ’s intrinsic properties.¹ To deny SP is to accept that an object, state, or event ϕ can be intrinsically valuable in a way that is explained at least in part by its extrinsic properties. SP does not tell us which objects, events, or states of affairs actually *are* intrinsically valuable, on the order of, say, hedonism. Rather, SP purports to identify the *concept* of intrinsic value, a concept to which any particular first-order *conception* of intrinsic value must live up.

In this paper, I argue that SP should be rejected. SP is inordinately restrictive. In particular, I argue that no SP-respecting conception of intrinsic value can accept the importance of *psychological resonance*, or the positive endorsement of persons, in explaining value. Given that psychological resonance is often, and plausibly, thought to be an important explanatory dimension of the *prudential good*, or what is intrinsically valuable for a person, SP’s inability to accommodate it is significant, and is strong reason to seek an alternative characterization of the nature of intrinsic value.

1. Terminology

Value-theoretic terminology is diverse. Philosophers often speak of “intrinsic

[†]I would like to thank Ben Bradley, Chris Heathwood, Doug Portmore, and participants in a discussion of the issue of this paper at PEASoup: <http://peasoup.typepad.com>, for their extremely helpful comments. I would also like to thank an anonymous reviewer.

¹Cf. Ben Bradley, “Is Intrinsic Value Conditional?” in *Philosophical Studies* 107 (2002), 23. See also G. E. Moore, “The Conception of Intrinsic Value” in *Philosophical Studies* (Oxford: Oxford University Press, 1951), 206.

value”, “extrinsic value”, “instrumental value”, “non-instrumental value”, “value as an end”, “final value”, etc. Thus some terminological explication is in order. Traditionally, “intrinsic value” is understood as synonymous with the idea of being “valuable as an end”. I will employ the traditional usage here. This differs from the practice adopted by, for instance, Christine Korsgaard. According to Korsgaard, intrinsic value just *is* value that supervenes on intrinsic properties. According to Korsgaard, SP is analytic. But that SP is analytic says nothing about whether SP is a proper account of the nature of value *as an end*—or, in Korsgaard’s terminology, “final value”.

Much of the traditional identification of “intrinsic value” with “final value” or “value as an end” relies on the implicit assumption of SP. As Korsgaard notes, those who would prefer to use the term “intrinsic value” to mean “value as an end” and contrast it with instrumental value, are advocating a substantive thesis about the relationship between final value and intrinsic value: that to be valuable as an end just is to be valuable in a way that supervenes on intrinsic properties.² This position is stated by Michael Zimmerman as follows:

[M]y keeping to the traditional practice of using the term “intrinsic value” to refer to [final] value is vindicated, for two reasons. First, since a state’s having the constituents that it has is part of its intrinsic nature (in both the nonrelational and qualitative senses of “intrinsic”) and since its final value supervenes on and only on its having the constituents that it has, *all final value is intrinsic value*. Second, there is... no nonfinal value that something can have that supervenes entirely on its intrinsic properties. Thus, *only final value is intrinsic value*. Hence Korsgaard’s distinctions coincide.³

In this paper, I keep the traditional terminological link between intrinsic value and “value as an end”. However, this paper will oppose Zimmerman’s view; I argue that intrinsic value (in the sense of “value as an end”) cannot be adequately characterized by means of SP. If one prefers Korsgaard’s terminology, my substantive claim can be stated as follows: SP cannot be understood to characterize the nature of final value. My terminology admittedly divorces the terminological identification of intrinsic value and “value as an end” from its philosophical basis (i.e., the implicit acceptance of SP).

²Korsgaard, “Two Distinctions in Goodness” in *Creating the Kingdom of Ends* (Cambridge: Cambridge University Press, 1996), 253.

³Michael J. Zimmerman, *The Nature of Intrinsic Value* (Lanham, MD: Rowman and Littlefield, 2001), 62.

Nonetheless, it seems to me sensible to oppose SP in the language in which it is advocated.⁴

2. *Well-Being and Intrinsic Value*

With this terminology in mind, the first point to make is that intrinsic value can take at least two forms. Intrinsic value can be *relational* as well as *non-relational*. An object ϕ is relationally intrinsically good if it is intrinsically good *for* something or someone. Claims about the nature of well-being or prudential value are claims about relational intrinsic value. My experience of pleasure at t_1 is intrinsically good *for me*. An object ϕ is non-relationally intrinsically good if it is intrinsically good “period”, “*tout court*” or “full stop”. Indeed, particular objects can be relationally intrinsically valuable but not non-relationally intrinsically valuable, and vice versa. That something is good *for me*, or “relationally”, does not guarantee that that thing is intrinsically good *tout court*, or “non-relationally”. My experience of pleasure might be intrinsically good for me, but might nevertheless lack non-relational intrinsic value. For instance, if I am a cold-blooded murderer, my experience of pleasure might be relationally intrinsically valuable (valuable for me), but it might, nevertheless, fail to be non-relationally intrinsically valuable, viz., intrinsically valuable *tout court*, or on the whole.⁵ It might be intrinsically better, some have claimed, that cold-blooded murderers feel pain than pleasure, though such pain is certainly intrinsically worse *for them* than pleasure.

Some deny that intrinsic value can be relational. For instance, Noah Lemos writes: “according to the traditional view, intrinsic value is a *non-relational* concept. When one says that something is intrinsically good, in the sense with which we are concerned, he means just that, that it is intrinsically good *period*.”⁶ However, Lemos does attempt to capture something like an account of relational intrinsic value:

we may also say, for example, that a state of affairs is intrinsically better for one person than for another. We might say that the state of affairs consisting in Smith’s being very happy and Brown’s being very unhappy is intrinsically better for Smith

⁴Not all people who identify “intrinsic value” with “final value” accept SP. See, for instance, Shelly Kagan, “Rethinking Intrinsic Value” in *Journal of Ethics* 4 (1998).

⁵See, for instance, Larry Temkin, *Inequality* (Oxford: Oxford University Press, 1993), 260.

⁶Noah Lemos, *Intrinsic Value: Concept and Warrant* (Cambridge: Cambridge University Press, 1994), 3-4.

than for Brown. But this may be taken to mean roughly that Smith and Brown have certain characteristics, F and G , such that Smith's having F is intrinsically better than Brown's having G . The fact that one person is better off than another when certain states of affairs obtain does not imply that intrinsic value is a relational concept.⁷

Lemos' accommodation is inadequate as an analysis of many common-sense claims about relational intrinsic value. For instance, the case of the cold-blooded murderer appears not to be adequately captured on Lemos' account. We might think that it is intrinsically good for a cold-blooded murderer to suffer. But this doesn't mean that suffering is intrinsically good *for the cold-blooded murderer*. The cold-blooded murderer might have a characteristic, suffering, such that the cold-blooded murderer's having that characteristic is intrinsically good. But nevertheless the cold-blooded murderer's suffering is surely intrinsically *bad* for the cold-blooded murderer. Lemos is correct to note that intrinsic value is not relational in all cases. It permits of non-relational forms. However, it also permits of relational forms, such as when something is intrinsically good for a person.

At first glance, however, SP has serious difficulties accommodating many popular theories of *relational* intrinsic value, at least in the form of prudential intrinsic value, or well-being. Though some theories of well-being—simple hedonism, for instance—can be accommodated straightforwardly (a pleasure is just as intrinsically valuable as its intrinsic pleasurable-ness), many views cannot. Take, for instance, some popular refinements to hedonism. Though two pleasures might have identical intrinsic properties, i.e., identical *pleasurableness*, some might hold that the value of these pleasures is affected by things that are extrinsic to the pleasure itself. My pleasure in solving Fermat's last theorem might be identically pleasurable to my pleasure in playing tiddlywinks. But, it might be said, my pleasure in solving Fermat's last theorem is better for me than my pleasure taken in playing tiddlywinks because the former was caused by the use of my "higher faculties".⁸ To explain this fact, we must appeal to facts that are extrinsic to the pleasure itself, including the extent to which one used one's higher faculties in the generation of said pleasure. Furthermore, consider the suggestion that that which is good for me is the satisfaction of my preferences or desires. For SP, what is beneficial for me? Is it the *desire*? Surely not—the desire itself is not good for me. Is it the object of desire? Though this might be

⁷Lemos, 4.

⁸The obvious inspiration for this view is J. S. Mill, *Utilitarianism*, ch. II.

plausible, it is not compatible with SP. The value of that which is desired cannot rest on its intrinsic properties; in order for a particular desired ϕ to be good for me, I must *desire* it.

These are serious problems for SP, but they can be overcome. Most who accept SP accept a monistic account of intrinsic value bearers: the only things that can bear intrinsic value are *states of affairs*.⁹ If this is correct, SP appears able to capture the views noted in the previous paragraph. According to SP, a state of affairs is intrinsically good for me if and only if its intrinsic value for me is explained by its intrinsic properties. Thus I might take pleasure in solving Fermat's last theorem. That this pleasure is *in solving Fermat's last theorem* is a relational property of the pleasure. However, the intrinsic value for me of the *state of affairs* in which I take pleasure in solving Fermat's last theorem can be perfectly well explained in terms of *its* intrinsic properties. Furthermore, one can perfectly well say that the state of affairs in which I take pleasure in playing tiddlywinks is less intrinsically valuable than the state of affairs in which I solve Fermat's last theorem. If states of affairs are proper value bearers, rather than pleasures *per se*, SP appears able to accommodate a relatively easy restatement of hedonism in its various dimensions.

Furthermore, this move allows SP access to a preference or desire-satisfaction theory. For SP, the desire satisfaction view insists that that which is intrinsically good for me is not the object of desire, but rather a particular state of affairs: a *conjunctive* state of affairs in which I desire that ϕ and ϕ . For the sake of brevity, call the conjunctive state of affairs in which x desires that ϕ and ϕ , a "D-state". An SP-respecting translation of a desire-satisfaction view would simply insist that D-states, rather than objects of desire *per se*, are the proper bearers of intrinsic value.

Selecting states of affairs as the proper value bearers is important for SP not only in accommodating important theories of prudential intrinsic value but also in accommodating additional claims about non-relational intrinsic value. In critiquing SP, Korsgaard writes:

Consider also the example of an extraordinarily beautiful painting unsuspectedly locked up, perhaps permanently, in a closet. Now a beautiful painting, I am supposing, is valued for its own sake. If [intrinsic value and final value] are equated, we must say it has intrinsic value. Yet it is locked in a closet, utterly unseen, and no one is the better for its existence. [...] Now on the Kantian type of account we can say that the painting is valuable for

⁹Bradley, 26. See also Zimmerman, ch. 3.

its own sake, yet so long as it remains locked up and unseen, it is no good at all. The condition of its goodness—the condition of the goodness of its beauty—is not met.¹⁰

In this case what seems to be missing from the locked-up painting is its being enjoyed by someone. If so, SP cannot accommodate this first-order fact about value: the value of the painting is explained in part by its relational properties, viz., its being enjoyed. Korsgaard insists that certain objects can only be good *given* the interest we take in them, given the fact that we value them as ends—as Korsgaard writes, “Value is, as I have put it, ‘conferred’ by choice.”¹¹ But if states of affairs, and not objects, are proper bearers of value, Korsgaard’s rejection of SP is too fast: SP can accommodate Korsgaard’s intuition: by insisting that intrinsic value is not borne by the *objects* we value, but is rather borne by the state of affairs in which the painting is valued or enjoyed. If this translation is plausible, Korsgaard’s critique fails.

SP’s acceptance of states of affairs as bearers of intrinsic value appears to vindicate SP’s ability to accommodate many common-sense first-order claims about both relational and non-relational intrinsic value. In what follows, I argue that this move is not enough. Though I allow that SP might plausibly escape Korsgaard’s critique, and though I accept the claim that intrinsic value is borne only by states of affairs, I hold that SP has special problems when it comes to relational intrinsic value or well-being. SP cannot capture the plausible thought that one important fact that can contribute to the prudential value of ϕ (whether a state, event, or object) for x is ϕ ’s psychological resonance with x .

Furthermore, SP’s failure to adequately capture an account of relational intrinsic value is seriously problematic. Given that intrinsic value permits of relational and non-relational forms, we should expect that that which characterizes the concept of intrinsic value in one form should do so in the other. If SP cannot adequately characterize relational intrinsic goodness, there is strong reason to believe that it cannot characterize intrinsic goodness *tout court*.¹² (I will argue for this claim in §4.3) However, even if we reject the claim that relational intrinsic value and non-relational intrinsic value should be treated as different forms of an identical concept, whether SP characterizes the concept of relational intrinsic value is itself worthy of investigation, as many have held that SP must hold of any acceptable theory

¹⁰Korsgaard, 264.

¹¹Korsgaard, 261.

¹²See, for instance, Chris Heathwood, “Fitting Attitudes and Welfare” in *Oxford Studies in Metaethics* v. 3, ed. Shafer-Landau (Oxford: Oxford University Press, 2008), 51.

of well-being.¹³

3. *Desire-Satisfaction and Internalism*

Take, for a moment, the desire-satisfaction theory. There are many versions of a desire-satisfaction view, including so-called “simple” or “actual” desire fulfillment views,¹⁴ “full-information” views,¹⁵ “ideal advisor” views,¹⁶ or “dispositional” views.¹⁷ Which version one accepts is not important for the purposes of this paper. But it is worth noting that important statements of a desire-satisfaction view do not seem to fit the model of the desire-satisfaction view accommodated by SP, i.e., that treats D-states as the bearers of intrinsic value.

For instance, consider Peter Railton’s view. Railton writes: “an individual’s good consists in *what he would want himself to want*, or to pursue, were he to contemplate his present situation from a standpoint fully and vividly informed about himself and his circumstances, and entirely free of cognitive error or lapses of instrumental rationality.”¹⁸ Going back a bit further in history, Hobbes:

But whatsoever is the *object of any man’s appetite or desire*; that is it which he for his part calleth *good*: and the object of his hate, and aversion, *evil*; and of his contempt, *vile* and *inconsiderable*. For these words of good, evil, and contemptible, are ever used with relation to the person that useth them: there being nothing simply and absolutely so.¹⁹

Consider also the view offered by Sidgwick:

[I]t would have to be said that a man’s future good on the whole is *what he would now desire and seek* on the whole if all the

¹³According to Ben Bradley, SP is a “non-negotiable” constraint on acceptable theories of well-being or relational intrinsic value. See Ben Bradley, *Well-Being and Death* (Oxford: Oxford University Press, 2009), 19, 27.

¹⁴James Griffin, *Well-Being: Its Meaning, Measurement, and Moral Importance* (Oxford: Oxford University Press, 1989), 10-11; Mark C. Murphy, “The Simple Desire-Fulfillment Theory” in *Noûs* 33 (1999).

¹⁵See Griffin, ch. 2; for criticism, see Sobel, “Full Information Theories of the Good” in *Ethics* 104 (1994).

¹⁶Most importantly, see Peter Railton, “Facts and Values” in *Facts, Values, and Norms* (Cambridge: Cambridge University Press, 2002).

¹⁷See David Lewis, “Dispositional Theories of Value” in *Papers in Ethics and Social Philosophy* (Cambridge: Cambridge University Press, 2000).

¹⁸Railton, 54. My emphasis.

¹⁹Hobbes, *Leviathan*, I 6. My emphasis. See also *Human Nature*, VIII 3.

consequences of all the different lines of conduct open to him were accurately foreseen and adequately realised in imagination at the present point of time.”²⁰

These views share a common feature.²¹ According to them, that which is intrinsically good for someone is *that which is desired by that person* (perhaps under certain idealized or counterfactual circumstances). For SP’s interpretation of the desire-satisfaction view, that which is intrinsically good for x is not the object of desire, but rather a D-state, i.e., the conjunctive state of affairs that includes x ’s desire of ϕ , and ϕ —whatever ϕ might be. The views above claim, on the other hand, that what is valuable for a person is the object of desire, i.e., ϕ —whatever ϕ might be—and the explanation of ϕ ’s goodness is that ϕ is desired. The latter view clearly violates SP. ϕ ’s intrinsic value for x supervenes, at least in part, on a strictly relational property: being desired by x .

But so what? Why can’t supporters of SP simply insist that—just as Korsgaard was incorrect to insist that the bearer of value was a beautiful painting itself rather than the state of affairs in which the painting is enjoyed—desire-satisfaction views are incorrect to hold that bearers of value are objects of desire? The answer runs as follows. One important motivation for a desire-satisfaction theory of well-being is a plausible, though controversial, constraint on well-being: existence internalism.²² Existence internalism about the personal good—very roughly—is the view that for ϕ to be good for x , x must be motivated to promote ϕ (or motivated to promote ϕ under relevant idealized or counterfactual conditions). Connie Rosati puts existence internalism in the following way: “The label ‘existence internalism’ . . . refers to the general thesis that there is a necessary connection between motivation and normative status. The thesis tells us that something X can have a certain normative status N only if someone A would be motivated by it in sense M.”²³ Existence internalism is plausible, at least at first glance, and explains, at least in part, the appeal of a desiderative view of the personal

²⁰Henry Sidgwick, *The Methods of Ethics* (Indianapolis, IN: Hackett Publishing Company, 1981 [1907]), 111-12. My emphasis.

²¹This feature is shared by a number of other desiderative views. See, for instance, Mulgan’s account of the desire-satisfaction view. See Tim Mulgan, *Understanding Utilitarianism* (Acumen, 2007), 73.

²²Many have held that existence internalism is a crucial plank in the argument for desiderative views of practical reason and the personal good. See, for instance, Bernard Williams, “Internal and External Reasons” in *Moral Luck* (Cambridge: Cambridge University Press, 1981); David Brink, “The Significance of Desire” in *Oxford Studies in Metaethics* v. 3, ed. Shafer-Landau (Oxford: Oxford University Press, 2008); Railton, 47.

²³Connie Rosati, “Internalism and the Good for a Person” in *Ethics* 106 (1996), 297-8.

good. Because—so the argument goes—only desires motivate, it must be the case that for ϕ to be good for a person, that person must desire (or desire under some set of counterfactual conditions) ϕ .²⁴ Only in this way will there be a necessary connection between ϕ 's “normative status” and the motivation to ϕ .

However, in insisting that the bearer of intrinsic value is a D-state, rather than an *object* of desire, SP fails to accommodate existence internalism, and hence fails to accommodate a major argument in favor of a desire-satisfaction theory. For SP to accommodate existence internalism, we must have some guarantee that, for every D-state, there is some further, positive motivational state taken toward that D-state. There can be no such guarantee. According to SP's translation of a desire-satisfaction view, for any given person, *any* D-state of desiring to ϕ and ϕ is good for them. *This holds no matter what ϕ actually is, and no matter whether we actually desire to ϕ , desire to desire that ϕ , or desire the relevant D-state.* Take an example. I cannot stand the song “Ball and Chain” by the rock band Social Distortion. I have no desire to hear it ever again (and hence I have no desire for the state of affairs in which I hear that song). Furthermore, I have no desire to desire to hear “Ball and Chain”. Actually, I am strongly averse to ever developing such a desire: I regard this song as objectively awful, and would regard a desire to hear it as a symptom of a general disintegration of my aesthetic sensibilities. For the same reason, I have no desire for the conjunctive state of my desiring hear the song and hearing the song. But according to SP's understanding of the desire-satisfaction view, the state of my desiring to hear the song and hearing the song is intrinsically good for me despite that state's utter impotence when it comes to my motivational states. This is a straightforward violation of existence internalism. According to SP, this state can be good for me even if I am strongly averse to it.

One might respond by restricting the range of intrinsically valuable D-states just to those that would motivate a given individual. One might do so by suggesting that a D-state is only intrinsically good for x if x desires the D-state in question. But this move is no help; it would obviously violate SP to be told that that a particular D-state is only good if I *desire* it; this would make the value of this D-state supervene in part on its relational properties.

²⁴Of course, the view that only desires motivate is certainly controversial. For defense, see Michael Smith, “The Humean Theory of Motivation” in *Mind* 96 (1987) and Neil Sinhababu, “The Humean Theory of Motivation Reformulated and Defended” in *Philosophical Review* 118 (2009). Critics include Stephen Darwall (*Impartial Reason* (Ithaca, NY: Cornell University Press, 1983)) and T. M. Scanlon (*What We Owe To Each Other* (Cambridge, MA: Harvard University Press, 1999)).

However, one might argue that it is perfectly compatible with SP to restrict the range of intrinsically valuable D-states not by x 's attitude toward the D-state itself, but rather by x 's attitude toward the relevant ϕ , out of which the D-state is constructed. The most obvious strategy would be to claim that a particular D-state of x 's desire to ϕ and ϕ is only intrinsically valuable for x if x desires ϕ . This would have the effect of ruling out as intrinsically valuable the D-state in which I desire to hear “Ball and Chain” and hear “Ball and Chain” because I have no desire for the relevant ϕ , i.e., to hear “Ball and Chain”. But there are two problems with this proposal. First, it would not accommodate existence internalism as stated. It would not guarantee that the bearer of intrinsic value for x motivates x ; it would only guarantee that one *element* of the D-state (i.e., ϕ) motivates. I might desire ϕ , but I might not desire to desire ϕ , and hence I do not desire the relevant D-state, which, according to SP's version of the desire-satisfaction view, is intrinsically good for me. But leave this aside. More importantly, this proposal also violates SP. Say, for instance, the relevant D-state is composed of my desire to hear “Ball and Chain” and my hearing “Ball and Chain”. In order to accommodate existence internalism, this D-state must not be good for me, because I fail to be motivated by it. The proposal now on the table is that this D-state is not intrinsically good for me because I fail to desire to hear “Ball and Chain”. But if so, facts about the intrinsic value of this D-state are settled by facts that are extrinsic to the D-state itself: the fact that I fail to desire to hear “Ball and Chain”—which is *not* a constituent of the relevant D-state. Hence it would appear that facts about the intrinsic value, or lack thereof, of a particular state of affairs supervenes on facts that are extrinsic to the state of affairs. It thus seems implausible that any potential restriction of relevant D-states could be found that (a) accommodates existence internalism *and* (b) accommodates SP.

4. *Resonance*

The failure of SP to plausibly accommodate existence internalism—and with it an important motivation for a desiderative theory of the personal good—leaves us with two choices. First, we can reject SP. Second, we can reject existence internalism, and with it a central motivation for the desire-satisfaction view. And there may be very good reasons to adopt the second course of action rather than the first. SP appears to be a reasonably well-entrenched and reasonably explanatory theory of intrinsic value; on the other hand, the overall philosophical status of existence internalism, and indeed a desire-satisfaction view of the nature of well-being, is far more

controversial. Hence if we feel more strongly about SP than we do about the desire-satisfaction view (if in fact we do), we should not feel uncomfortable about jettisoning the latter to maintain the former.

However, *even if* we reject existence internalism, there are further reasons to reject SP. First, SP is supposed to be a claim about the *conceptual* nature of intrinsic value. Even if existence internalism is false, it seems wrong to say it is *analytically* false, or ruled out merely at the level of concept. Rather, we are much more likely to point out some of the more implausible features of existence internalism or a desire-satisfaction view, and claim that this view fails at a first-order, rather than second-order level. Second, and more importantly, existence internalism is merely one way of specifying a more general feature of the nature of well-being, on which *many* views rely. In particular, it is one version of a *psychological resonance* constraint.

One version of a psychological resonance constraint is stated as follows by Railton:

Is it true that all normative judgments must find an internal resonance in those to whom they are applied? While I do not find this thesis convincing as a claim about all species of normative assessment, it does seem to me to capture an important feature of the concept of intrinsic value to say that what is intrinsically valuable for a person must have a connection with what he would find in some degree compelling or attractive, at least if he were rational and aware. It would be an intolerably alienated conception of someone's good to imagine that it might fail in any such way to engage him.²⁵

As noted by Railton, a resonance constraint claims, simply enough, that that which is intrinsically good for a person must *resonate* with that person. Put more precisely, for some state ϕ to be good for x at time t_1 , ϕ must be positively endorsed by x at t_1 . Positive endorsement can take a number of forms. For Railton, endorsement is counterfactual: ϕ *would* resonate with x , were x rational (and by this he means “instrumentally rational”²⁶) and aware (and by this he means in possession of full information). Some will construe positive endorsement more strongly, i.e., that ϕ must resonate with

²⁵Railton, 47.

²⁶Railton writes: “One might join the Kantians in challenging the premise of instrumentalism, and attempt to argue that some substantive ends or actions are indeed mandated by rationality. Such challenges have a noble history, and even a noble representation in the present, but I find the ignoble instrumentalist view the clearest idea we have of what it is, at a minimum, to have a reason for acting.” (Railton, 47).

x under x 's actual conditions. Furthermore, one can identify a number of potential pro-attitudes as a form of positive endorsement. Existence internalism is one form of resonance constraint: here the relevant positive attitude is the attitude of being motivated (or being disposed to be motivated under certain idealized or counterfactual conditions). But one could easily imagine a resonance constraint that insists that certain purported goods must be *believed* to be good, or must be the subject of autonomous *choice*, or must be *enjoyed*, or must be the subject of any number of additional attitudes (or dispositional, or counterfactual, attitudes) on the part of individuals for whom those goods are good.

Though a resonance constraint is controversial, it is responsive to a roughly compiled set of considered judgments about the proper connection between a person and her own good. In general, it is plausible to say that a person (perhaps under idealized conditions) ought to have at least some limited *control* over that which is intrinsically good for her, that that which is good for her ought not to be identified in a way that is *alien* to her own attitudes, ought to be explained, at least in part, by her assent.²⁷ Though these various intuitions are slippery and ill-defined, a natural and plausible way to capture them is to insist that a necessary condition of ϕ 's goodness for x is ϕ 's psychological resonance with x . As Sumner notes, subjectivism about welfare—which he defines simply in terms of a resonance constraint—is plausible in part because subjectivism makes “your well-being depend on your own concerns: the things you care about, attach importance to, regard as mattering, and so on,”²⁸ (whether these “concerns” are actual or counterfactual).

To properly satisfy the resonance constraint and its motivation, one must insist that ϕ 's psychological resonance with x is at least part of the explanation of ϕ 's intrinsic value for x —it is one fact upon which the value of ϕ supervenes. First—as in the case of existence internalism—unless the prudential value of ϕ for x depends at least in part on its resonance with x , we cannot guarantee that all prudential goods for x will resonate with x (this will be a matter of some discussion in §5.2). Second, for any view to plausibly satisfy the demand that x ought to have the right sort of control over, or connection to, x 's good, at least one part of the explanation of ϕ 's

²⁷Larry Temkin notes this feature of a desiderative view: “on this view, the agent is . . . the ultimate arbiter of her own good. What she desires is good for her, and, importantly, it is her desiring it which makes it so.” Larry Temkin, *Inequality* (Oxford: Oxford University Press, 1993), 264.

²⁸L. W. Sumner, *Welfare, Happiness, and Ethics* (Oxford: Oxford University Press, 1996), 42.

goodness for x must be that ϕ is endorsed (however this is characterized) by x . Otherwise x has no control at all: ϕ 's value for x is in no way controlled by x or x 's attitudes. But SP cannot allow that resonance will play any explanatory role in the goodness of ϕ for x . This is true even if we accept, as Railton does, that ϕ need only resonate with x 's attitudes under counterfactual conditions: such attitudes are extrinsic properties of ϕ , no matter how they are construed.

The problem for SP is magnified once we see that resonance can play an explanatory role in ϕ 's intrinsic goodness for x , even if ϕ 's resonance is not a *necessary condition* for ϕ 's intrinsic goodness for x . In other words, some views might suggest that not all goods must resonate, but that ϕ resonates with x is sufficient for ϕ to be intrinsically good for x . Alternatively, one might suggest that resonance is neither a necessary nor sufficient condition, but that ϕ 's resonance along with some other condition (such as ϕ 's being at least neutral when it comes to some account of "objective" goodness) might be sufficient for ϕ 's prudential value. Here resonance is neither a necessary nor sufficient condition for ϕ 's value for x , but *is* a contributing factor: ϕ 's resonance with x is one thing on which the intrinsic goodness of ϕ for x might supervene. Take, for instance, the perfectionist theory of well-being defended by David Brink. This view insists not on the authority of *desire* over an agent's good, but rather on the authority of *autonomous choice*. Brink holds that choices have normative significance insofar as they help to explain why some chosen object, ϕ , is good for an agent x who autonomously chooses it.²⁹ On Brink's view, the fact that an agent autonomously chooses an object ϕ can function as part of the explanation of ϕ 's value, hence ϕ 's value can supervene on ϕ 's relations as well as ϕ 's intrinsic properties. For Brink, the authority of choice is limited. Chosen states must be, at least to some degree, inherently choiceworthy.³⁰ However, choice still enters into an explanation of ϕ 's intrinsic value. Hence because, for Brink, choice *can* enter into an explanation of the value of particular states for persons, Brink's view is incompatible with SP. Furthermore, no SP-compatible translation of Brink's view is available. Such a translation must hold that a state in which I *choose* ϕ , and ϕ occurs or obtains, is good for me, no matter the value of ϕ (assuming other relevant conditions, including the independent choiceworthiness of ϕ , are met). (Call a state composed of my choosing ϕ and ϕ a "C-state".) But there is no reason to believe I have *chosen* the relevant C-states that, on such an SP-compatible translation, are beneficial

²⁹Brink, 40-45.

³⁰Brink, 39-40.

for me.

Hence SP cannot respect any view for which resonance plays an explanatory role. Call any view that makes welfare depend, at least in part, and at least occasionally, on psychological resonance a “resonance-respecting” theory of well-being. Of course, the mere fact that some philosophers have held resonance-respecting theories is no reason *by itself* to reject SP. But the reach of resonance is important insofar as the resonance constraint is an important and attractive explanatory principle of the personal good. It is certainly plausible to say that at least *some* limited range of welfare goods (even if not all) depend on the pro-attitudes of the person for whom they are good. And though the explanatory power of resonance may well turn out to be chimerical, it would be substantially disappointing to be told that any theory that employs resonance as an explanatory principle is ruled out as a matter of concept.

5. Responses

That the relational intrinsic value of ϕ can supervene on the pro-attitudes of persons for whom ϕ is good is a reason to reject SP. However, there are three responses worth consideration here. The first response argues that though SP cannot accommodate the importance of psychological resonance to well-being, nothing is lost by rejecting resonance-respecting formulations of these views, because an SP-compatible translation yields identical verdicts. In other words, the difference between a genuinely resonance-respecting view and an SP-compatible translation of a resonance-respecting view is a distinction without a (philosophically significant) difference. The second response holds that SP could accept one version of the resonance constraint if we accept that there are substantive requirements of rationality. The third response accepts my critique as it stands, but holds that SP is only a proper theory of *non*-relational intrinsic value. Third, and finally, I consider an argument, offered by Ben Bradley, that no matter how important resonance *seems* in discussions of well-being, no resonance-respecting theory can be a proper conception of relational intrinsic value, because no resonance-respecting theory can accomplish that which a conception of well-being must.

5.1. Does Resonance Matter?

To begin this response, it is worth noting that SP can offer a translation of any resonance-respecting view in terms of states that are similar in structure

to D-states: SP will simply insist that, in the case of a resonance constraint (such as existence internalism), all intrinsically good states of affairs for x must be—at least—composed of x bearing a resonating pro-attitude toward ϕ and ϕ (call this an “R-state”). In the case of resonance as a merely explanatory principle (such as in the case of Brink’s perfectionism), SP will hold that which is good for x will include *some* particular, independently specified R-states (though it may include some states of affairs that do not qualify as R-states).

Of course, I have argued that SP-respecting translations of resonance-respecting views cannot themselves be resonance-respecting. But one obvious response might be that, when all is said and done, the selection of value bearers matters little to an account of relational intrinsic value. Assume, for the purposes of illustration, a resonance-respecting desire-satisfaction theory of well-being. Such a view must hold that were a particular D-state to obtain, this would improve x ’s well-being, for whatever value of ϕ . After all, in any D-state, x desires some particular object or state, is motivated by that state, and this state occurs. Hence, were this particular D-state to obtain, indeed *any* D-state, x ’s well-being would improve, even for a resonance-respecting view. Thus the difference between a resonance-respecting desiderative view and an SP-respecting desiderative view might be substantially less than advertised, and perhaps not enough to reject SP on the grounds of inability to accommodate genuinely resonance-respecting views.

My response is as follows. A resonance-respecting view is surely committed to saying that, whenever these adequately translated R-states obtain, that they *do* so obtain is good for x . Nevertheless, there is a large and philosophically important difference between a counterfactual claim, e.g., “were this conjunctive state to occur, my well-being would improve,” and a non-counterfactual, e.g., “this R-state is itself intrinsically good for me”. One might put this in temporal terms. Any theory of well-being should, at least in principle, be able to come up with a rank-ordering (within the bounds of commensurability) of goods that can deliver a verdict about what is good for a person at t_1 . What distinguishes resonance-respecting theories is that x ’s pro-attitudes at t_1 are relevant to the list: what is good for x at t_1 is determined, at least in part, by x ’s pro-attitudes at t_1 . If we accept a resonance constraint, it will not be the case that the conjunctive state of my taking a pro-attitude toward hearing “Ball and Chain” and hearing it is good for me at t_1 . The list of things that are good for me at t_1 are composed of things that actually resonate with me at t_1 , and the state in which I take a pro-attitude toward hearing “Ball and Chain” and hear it does not.

Any resonance-respecting view will hold that *were* it the case that an adequately translated R-state obtains, I would be better off *at the time at which this R-state obtains*. This is substantially different, however, than SP’s treatment. According to SP at t_1 , this R-state is good for me, no matter the value of ϕ (assuming, of course, that other conditions of value as dictated by the relevant theory are met by the relevant R-state). This is because SP understands that which is good for a person at a time in an entirely different way than resonance-respecting theories. This is no superficial difference between first-order theories of well-being, and any theory of the concept of intrinsic value should be able to represent both such approaches.

5.2. Resonance and Substantive Rationality³¹

In discussion of SP’s inability to accommodate a resonance-respecting theory of welfare, I lumped together two distinguishable versions of the resonance constraint: first, a resonance constraint that insists on resonance with x ’s actual attitudes; second, a resonance constraint that insists on resonance with x ’s attitudes under counterfactual conditions, including, as specified by Railton, conditions of rationality and awareness. But if we allow that resonance might be construed in the latter way, this might allow an entry window for SP. I have so far not considered the possibility that there may be *substantive* requirements of rationality such that, were x rational, x would positively endorse D-states (or, perhaps, other SP-compatible welfare goods). In other words, it might be that *any rational person* will endorse the state in which that person desires to ϕ and ϕ (or other SP-compatible welfare goods), because rationality itself requires their endorsement, and hence if the resonance is construed as counterfactual on x ’s positive endorsement under conditions of rationality, SP can accommodate a resonance constraint.

Of course, Railton rejects this interpretation of “rational”, insofar as he rejects such substantive requirements of rationality.³² I myself, however, will refrain from entering this vexed dispute. Even if we accept that there can be such substantive requirements of rationality, however, SP remains unacceptable, for four reasons. First, there is at least some reason to believe that the resonance constraint should be construed in Railton’s way, i.e., one’s resonating attitudes are idealized only to the extent that one is instrumentally rational. It is plausible to believe that for ϕ to be good for x , it must bear the right sort of connection to x , a connection it might not bear to y , who

³¹Thanks to an anonymous reviewer for suggesting this response.

³²See note 26.

happens to possess very different attitudes, interests, desires, etc. Welfare goods, in other words, are not one size fits all—that ϕ is good for a person must be, in part, explained by something about *her*, and her concerns. Suggesting that resonance is construed counterfactually given the assumption of substantive rationality would leave x 's and y 's differing concerns irrelevant to the comparative determination of that which is intrinsically good for x and y .

Leaving this aside, however, even if we accept the claim that there is a substantive requirement of rationality to positively endorse whatever welfare goods, SP cannot be resonance-respecting. In particular, SP cannot accept that ϕ 's resonance with x (perhaps even under conditions of substantive rationality) can *explain* ϕ 's welfare value for x . That x would, were x substantively rational, value ϕ is certainly an extrinsic property of ϕ . Of course, if one rejects Railton's limitation of "rationality" to "instrumental rationality", there is less pressure to declare that to satisfy the resonance constraint, the value of ϕ must depend on, or supervene on, ϕ 's resonance: it remains the case, if we accept substantive principles of rationality of the right kind, that ϕ would not alienate x , were x substantively rational. But rejecting the explanatory importance of resonance is not without cost. Doing so yields a view that cannot accommodate a central motivation for insisting on the evaluative importance of resonance in the first place. Recall that it is at least occasionally plausible to say that x should have *control* over x 's good, and hence that the explanation of ϕ 's goodness for x should appeal to something *about* x , in particular, x 's own positive endorsement of ϕ (under whatever relevant conditions). If ϕ 's value does not supervene on anything about x , x does not have such control— ϕ 's goodness is determined in ways that are totally independent of x 's endorsement, however this endorsement is understood.

Again, I do not wish to enter the vexed dispute concerning whether there are any substantive requirements of rationality. There may very well be. The point here concerns whether SP could accommodate the evaluative importance of resonance even in light of such substantive requirements. The answer is no, given that SP cannot allow that the prudential value of ϕ for x will supervene on, or at least be partially explained by, ϕ 's psychological resonance. And only admitting that value can, at least to some extent, supervene on instances of psychological resonance is responsive to considered judgments that motivate the evaluative significance of psychological resonance.

Third, even if we allow that SP could accommodate the resonance *constraint*, it cannot accommodate all the ways in which resonance might be

important to prudential value. Even if we agree that not all welfare goods must bear a connection to x 's pro-attitudes, it is at least occasionally plausible to declare that x 's attitudes can *render* ϕ prudentially valuable for x . SP cannot allow such a claim, insofar as this would require an explanatory principle: ϕ 's value for x is explained in part by x 's attitudes, or x 's attitudes were x rational (however this is understood) and aware. If so, we should reject SP.

Finally, even if we are steadfast in rejecting the claim that resonance can have any role in explaining the nature of prudential value, it remains the case that SP is inordinately restrictive and should be rejected. SP cannot accommodate any theory of well-being for which (a) resonance is not construed counterfactually or (b) resonance is construed counterfactually, but not dependent on x 's rationality or (c) resonance is construed counterfactually and dependent on x 's rationality, but only on x 's instrumental rationality or (d) resonance is construed counterfactually, on x 's substantive rationality, but there are no substantive requirements of rationality to value particular welfare goods or (e) any view for which resonance plays an explanatory role or (f) any theory that allows x control over that which is intrinsically valuable for x . And though such theories may well be false, they are certainly not ruled out conceptually; they are surely well-formed theories of what it means for a particular ϕ to be relationally intrinsically valuable for a person. And this, it seems to me, is enough to declare that SP cannot accommodate a wide and important range of theories of welfare. To sum up: though the denial of (a)-(f) is on the table, their acceptance is on the table, as well. If so, SP cannot—as it aspires to—be an adequate mark of the concept of intrinsic value.

5.3. *The Supervenience Principle and Non-Relational Value*

We have seen reason to reject SP as a proper mark of the concept of relational intrinsic value. Of course, one option might be to divorce the second-order analyses of relational and non-relational intrinsic value. Given the *prima facie* importance of resonance to well-being, we might simply be content to allow SP to supply the best account of *non-prudential* or *non-relational* intrinsic value. If so, the argument I offer against SP is dulled. Though it may be compelling to hold that persons should have some say over that which is good *for them*, which in part explains the appeal of the importance of resonance, it is less appealing to hold that that which is good *tout court* must similarly resonate.

However, I think this strategy cannot be plausibly sustained. Recall

what SP is. SP is not a theory about what makes something intrinsically valuable—it is not, in other words, a *conception* of intrinsic value. Rather, it is an *identifier*, or *mark*, of intrinsic value as a concept. Hence if SP is not to identify relational intrinsic value, it would appear that the term “intrinsic value” simply refers to two different concepts—one that is identified by SP, and one that is identified by some other principle. But this seems implausible. All the evidence seems to point to the fact that relational and non-relational intrinsic value are simply two forms of the same concept. For instance, one important mark of the nature of intrinsic value is its relation to practical reasons. It is plausible to say that if x is intrinsically valuable, I have at least *pro tanto* reason to promote it.³³ However, this is surely true of relational intrinsic value: if ϕ is intrinsically valuable for x , x certainly has *pro tanto* prudential reason to promote ϕ .

Furthermore, consider three platitudes noted by Bradley that are generally associated with the concept of intrinsic value:

DEG: *Intrinsic value is the sort of thing that comes in degrees.*

MAX: *Intrinsic value can be added up and maximized.*

BET: *Adding something with intrinsic value to the world makes the world, or a life, or an outcome, better (other things being equal).*³⁴

Each of these platitudes applies to relational intrinsic value. First, relational intrinsic value comes in degrees. Marrying my beloved is better for me than being a professional philosopher (or vice versa), for instance, though I might desire both for their own sakes. Second, well-being can be added up and maximized. The quality of my life at a given time is the aggregate of those things that count in favor of and against my well-being at that time. Furthermore, relational intrinsic value can certainly be maximized; indeed, maximizing my overall well-being appears to be the paradigmatic demand of prudential rationality. Finally, as BET notes, it is certainly true that adding something with intrinsic value *for me* to the world makes my life better (other things being equal). If a state of affairs is intrinsically good for me, and that state of affairs is brought about, or “added” to the world, my life goes better. Though this is not dispositive evidence that these concepts should be treated identically, it does shift the burden to those who would

³³This could be for many different reasons, and I remain neutral concerning which of the relata, “value” or “reason to promote”, is explanatorily prior.

³⁴Ben Bradley, “Two Concepts of Intrinsic Value” in *Ethical Theory and Moral Practice* 9 (2006), 119. The “other things being equal” qualifier is meant to rule out bad instrumental effects of intrinsic goods.

choose to accept SP for non-relational intrinsic value, reject it for relational intrinsic value. Given their similarity, relational and non-relational intrinsic value appear to be simply different forms of the same concept: intrinsic value.

As I noted before, however, even if we reject the claim that relational intrinsic value is a different form of the same concept as non-relational intrinsic value and claim instead that they are different concepts altogether, the failure of SP to adequately characterize the concept of relational intrinsic value is independently interesting. Many hold that SP is a requirement for *any* theory of well-being. One such thinker is Bradley, to whose defense of this claim I now turn.

5.4. *The Concept of Well-Being*

In a recent book, Ben Bradley claims that no theory of well-being that rejects SP can be a genuine theory of well-being at all. Hence it would appear that, if this is correct, a resonance-respecting theory of well-being must not be a theory of welfare proper.³⁵ Insofar as an extremely wide range of theories are resonance-respecting, Bradley's claim, if true, would be substantially surprising. What is the argument for this claim?

According to Bradley, any acceptable theory of well-being must do three things:

- (1) First, we must say what sorts of states of affairs are intrinsically good or bad for us in the *most fundamental* way, or are *basically, non-derivatively*, good for us. Call these the *value atoms*. The value atoms are what fundamentally and completely determine how well things go for us. (2) We must then say what determines *how* good or bad those atoms are. (3) Finally, we must say how to determine the values of non-atoms we might be particularly interested in.³⁶

At first glance, this list seems extraordinarily reasonable. One caveat, however, is worth mentioning. (1) is stated in a way that is difficult to understand. Bradley insists that we must identify the states of affairs that are "intrinsically good or bad for us in the most fundamental way," or that are "basically, non-derivatively" good for us. It certainly seems plausible to say that any theory of well-being should strive to identify those states of affairs

³⁵Ben Bradley, *Well-Being and Death* (Oxford: Oxford University Press, 2009), 19.

³⁶Bradley, *Well-Being and Death*, 5.

that are fundamentally good, if that simply means those states of affairs that are themselves *intrinsically good* for us. But what does it mean to be intrinsically good “in the most fundamental way”?

The most obvious possibility is that “intrinsically valuable in the most fundamental way” means *not intrinsically valuable simply as a conjunction of other, more basic, states that are themselves intrinsically valuable*. Indeed, this appears to be precisely Fred Feldman’s understanding of “basic intrinsic value”. Feldman writes:

When I say that these states of affairs have their intrinsic values in “the most fundamental way”, part of what I mean is that they do not have their intrinsic values because they are conjunctions or other combinations of smaller items that are intrinsically good. Thus, if some state of affairs, P, is intrinsically good, and some other state of affairs, Q, is intrinsically good, then their conjunction, P&Q, might also be intrinsically good. The value of P&Q might just be the sum of the value of P and the value of Q. In that case, P&Q would not have its value “in the most fundamental way”. It would have its value in virtue of the fact that it is composed of these parts, and they have their specified values. So P&Q could not be a basic intrinsic value state.³⁷

Of course, a resonance-respecting view could agree: value atoms are those states of affairs that are *themselves* resonating (e.g., the objects of desire), not states of affairs that are merely conjunctions of other, more basic, resonating states.

However, it is unclear whether this is the account of “basic intrinsic value” that Bradley means to offer. Consider now Bradley’s direct argument for SP, which refers directly to his use of the term “fundamental” in the desiderata offered above.

[SP] is a requirement of any acceptable theory of well-being. This is because, as noted above, the value atoms should be *instantiations of the fundamental good- or bad-making properties*—the properties that are fundamentally and completely responsible for how well a world (or a life, or . . .) goes. Suppose [SP] were false. Then there could be two properties, F and G, such that the only intrinsically good states of affairs are those involving the instantiation of F alone, but whose values are determined by whether

³⁷See Fred Feldman, *Pleasure and the Good Life* (Oxford: Oxford University Press, 2004), 173-4.

there are any instantiations of G. But if that were true, F would fail to be a *fundamental* good- or bad-making property, for instantiations of F would fail to completely determine what value there is.³⁸

It is true, of course, that any resonance-respecting view will accept that though F is that which is intrinsically good, F's intrinsic goodness will—at least occasionally—require the presence not just of F (and F's intrinsic properties), but also of some additional G. One might, for instance, replace F with “the state of affairs in which I listen to ‘Ball and Chain’”, and G with “the state of affairs in which F resonates with me”. For at least some true claims about well-being, assuming a resonance-respecting view, the value of F will depend on the presence of G, and hence F, according to Bradley, would fail to be a “fundamental” good- or bad-making property. Hence it would appear that Bradley means that a property can only be “fundamentally” intrinsically good if that property is itself sufficient to “completely determine what value there is.”

We are now in a position to see the error in Bradley's argument for SP. For the argument to yield the truth of SP, it must be the case that, in (1), “fundamental” is understood to exclude the possibility of the intrinsic value of a value atom supervening on relational properties of these value atoms. But this presents two serious problems. First, using “fundamental” in this way this simply begs the question in favor of SP. Second, if (1) is to be understood as claiming that a theory of well-being must offer a theory of value atoms that excludes the value of value atoms from supervening on relational properties, (1) has no plausibility. It would exclude, for instance, any resonance-respecting theory as a genuine theory of well-being. But as we have seen, many important verdicts about the nature of welfare and well-being depend upon respecting a resonance condition. If so, under this understanding of “fundamental,” it (a) begs the question and (b) renders (1) implausible.

If we are to interpret (1) in a way that doesn't straightforwardly beg the question, we must understand “fundamental” differently than is required by Bradley's argument for SP. We must understand it in Feldman's way. (1) is independently plausible only if (1) holds that any theory of well-being must identify those states of affairs that are intrinsically good in a basic way: value atoms that do not simply depend on their being conjunctions of further intrinsically valuable states. Surely any theory of well-being must provide a plausible account of such value-atoms. Hedonism, will, for in-

³⁸Bradley, 19.

stance, suggest that value atoms are discrete moments of pleasure.³⁹ The desire satisfaction view will claim that value atoms just are those most basic states of affairs that are desired for their own sakes. But (1) can be satisfied by any resonance-respecting theory of well-being if at least part of the explanation of the value of at least one basic intrinsic good ϕ for x is that ϕ resonates with x . Hence we should not accept Bradley’s argument for SP. Either Bradley equivocates on “fundamental” (between Feldman’s understanding, and the understanding that implies SP), or interprets “fundamental” in a way that begs the question in favor of SP.

6. Conclusion

SP cannot plausibly accommodate the importance of resonance to judgments of well-being, or what is intrinsically good for a person. If so, it seems to me, we have strong reason to reject SP as an account of the concept of intrinsic value. The concept of intrinsic value certainly permits of conceptions for which resonance plays an important explanatory role. This SP cannot do.

It should be noted that the denial of SP does not commit me to any *particular* first-order facts about intrinsic value, nor does it commit me to any particular account of the distinction between intrinsic value and various other sorts of non-intrinsic value, nor does it commit me to any particular account of that which can properly bear intrinsic value. The rejection of SP, however, is essential to capturing a wide-range of judgments about relational intrinsic value, including judgments about the nature of intrinsic value for a person. The true nature of intrinsic value—or “final” value, in Korsgaard’s terminology—must await further examination.

³⁹Cf. Feldman, 174.