

Justice for Hedgehogs, Conceptual Authenticity for Foxes: Ronald Dworkin on Value Conflicts

Jack Winter¹

Published online: 26 October 2015

© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract In his 2011 book *Justice for Hedgehogs*, Ronald Dworkin makes a case for the view that genuine values cannot conflict and, moreover, that they are necessarily mutually supportive. I argue that by prioritizing coherence over the conceptual authenticity of values, Dworkin’s ‘interpretivist’ view risks neglecting what we care about in these values. I first determine Dworkin’s position on the monism/pluralism debate and identify the scope of his argument, arguing that despite his self-declared monism, he is in fact a pluralist, but unusual in denying conflict between plural values. I then set out the structure of his interpretive theory of value relations and present a case of value conflict which I think interpretivism cannot deal with. Following this I argue that there are structural reasons why cases like this are liable to occur and suggest that interpretivism will frequently fail to properly reflect people’s moral commitments because reinterpretation of values has the side effect of excluding important moral commitments from our conceptions of values. While, as Dworkin argues, there are no brute moral facts concerning values, moral psychology constrains the range of acceptable conceptions of values. Given the shortcomings of interpretivism I conclude that we should acknowledge that values may conflict.

Keywords Dworkin · Pluralism · Monism · Interpretation · Values

In his book *Justice for Hedgehogs*, Ronald Dworkin (2011) defends the view that genuine values cannot conflict and, moreover, that they are necessarily mutually supportive. In this paper I argue that by prioritizing consistency over the conceptual authenticity of values, Dworkin’s ‘interpretivist’ view risks neglecting what we care

✉ Jack Winter
jack.winter@manchester.ac.uk

¹ University of Manchester, Manchester, UK

about in these values. In the first section I determine Dworkin's position on the monism/pluralism debate and identify the scope of his argument. In the second section I examine the structure of his interpretive theory of value relations. The third section presents a case of value conflict which I argue interpretivism cannot deal with. Finally, in the last section I argue that there are structural reasons why cases like that in the previous section are liable to occur and suggest that interpretivism will frequently fail to properly reflect people's moral commitments because reinterpretation of values has the side effect of excluding important moral commitments from our conceptions of values. Given the shortcomings of interpretivism I conclude that we should acknowledge that values may conflict.

Before beginning I should clarify some terms. For the purpose of this paper a 'value' is a moral principle directing that particular actions, attitudes or states of affairs be promoted or avoided on pain of wronging some person. Dworkin distinguishes between values and desiderata: while both are things a society might want, values 'have judgmental force' (Dworkin 2011, p. 118) and a society or person which fails to respect a value wrongs someone (cf. Dworkin 2001a, pp. 79–80). While desiderata may conflict with one another or with values, Dworkin believes that values, with their attendant claims, cannot conflict with one another. A 'value-conception' is a particular conceptual articulation of a value, specifying what sorts of actions, attitudes or states of affairs the value in question is correctly understood to impel us to promote or avoid. A value-conception is 'authentic' when it captures the sort of moral concerns moral agents have regarding the relevant states of affairs, attitudes or actions, and 'inauthentic' when it does not. The 'realization' of a value is the performance of the actions, holding of the attitudes, or establishment of the states of affairs that the value directs be pursued. 'Conflict' is the impossibility of realizing two values together in a given situation, and 'consistency' is an antonym for conflict.

An Idiosyncratic Pluralism

Dworkin's target is pluralism, which holds that there are many distinct values which may conflict (Dworkin 2011, p. 2); but it is less apparent what alternative position he wants to stake out (Fallon 2010, pp. 545–550). Pluralism is typically counterposed to monism, which denies that our moral landscape is populated by distinct and irreducible values, arguing that despite appearances there is in fact only one fundamental value. At some points, Dworkin appears to intend a fully-fledged monism of this type. On the first page of *Justice for Hedgehogs* he declaims that the book 'defends a large and old philosophical thesis: the unity of value. [...] Value is one big thing' (Dworkin 2011, p. 1). However, unlike true monists, he does not suppose some higher value like utility from which the value of other moral goals is derived. Indeed, contrary to the quoted passage, the substance of his project seems to be to show that the things we value are distinct and non-derivative, but all exist in relations of mutual support:

I hope to develop *integrated* conceptions [of values] that all seem *right in themselves*, at least after reflection.

(Dworkin 2011, p. 5, my italics)

[T]he hedgehog's faith [is] that all true values form an interlocking network,

(Dworkin 2011, p. 120)

Both these statements seem to imply an acknowledgement of multiple distinct and fundamental values; entities cannot be integrated or located in a network unless they are themselves distinct. In fact, Dworkin goes as far as to refer to 'the obvious fact that there are many values' (Dworkin 2011, p. 93). Despite the advertised aim of *Justice for Hedgehogs* I think the best understanding of Dworkin's position is that he recognizes the independence of different values but wants to show that they can never truly conflict with one another (cf. Baker 2010, p. 762; Kalderon 2013, p. 130; Smith 2012, p. 385; Williams 2001, p. 92). Dworkin is not a monist but a pluralist because he recognizes the separateness of different fundamental values. However, his is an idiosyncratic pluralism which differs from orthodox pluralism in denying that those separate values can ever conflict with one another. On Dworkin's view genuine values cannot conflict either conceptually or even in application. This means that, firstly there are no values which for logical reasons contradict one another, and secondly, and more stringently, that even for two such logically compatible values no situation can ever arise in which they conflict for contingent reasons (Dworkin 2011, p. 119).

A view recognizing distinct fundamental values but asserting their universal consistency seems quite surprising and is open to an obvious phenomenological objection. As Stuart Hampshire (1983, p. 151) puts it, '[o]ur everyday and raw experience is of conflict between contrary moral requirements at every stage of anyone's life'. This experience provides a *prima facie* case for the possibility of conflict. To take a familiar example, politicians are eager to justify policies in terms of prioritizing the claims of one value over those of another: responsibility over welfare, freedom of choice over equality, security over freedom, and so on. Granted, these characterizations of policy choices are often mistaken or disingenuous, but even so, at first glance it seems fantastic to claim that no value conflicts ever occur, if only because of the sheer number of such claims. The question, then, is how Dworkin seeks to support his position.

Interpretivism

There are two possible strategies that might show that values cannot conflict: the first is to show that consistency is logically necessary. This would mean identifying some logical implication of our understanding of values which precludes their conflicting. This is not the strategy Dworkin opts for, and, given the widespread adherence to pluralism among philosophers it would be surprising if this strategy were fruitful, since it would imply that pluralists have committed some basic conceptual error in their understanding of values. The second option would be to investigate value conflict in a way analogous to a scientific enquiry, treating

consistency as a hypothesis and instances of apparent conflict as something like empirical phenomena whose resolution cannot be assumed a priori (cf. Fallon 2010, p. 545). And in fact, Dworkin appears to endorse this second approach: '[w]e confront it [apparent conflict] at retail, case by case, but we confront it through a conceptual rearrangement that works toward eliminating it' (Dworkin 2011, p. 120). Dworkin regards consistency, then, not as a logical necessity, but as a hypothesis to be tested against cases of apparent value conflict. The more cases that turn out not to involve irresolvable conflict between values, the stronger the support for the hypothesis. Conversely, if a case of genuine conflict is discovered the hypothesis will be falsified.

An alternative interpretation of Dworkin's position is that consistency is not intended as a defeasible hypothesis, but as a commitment to attempt to resolve conflict wherever it appears.¹ In fact, I believe these two readings need not contradict one another; and the understanding of consistency as a hypothesis explains how Dworkin's argument might be persuasive to those not already devoted to consistency as a prior position. I can see two ways to understand a commitment to consistency: either the idea would be that we incorporate consistency into our concept of a value so that should two values appear to conflict we can infer by definition that (at least) one is not a genuine value, or the commitment could be a procedural protocol requiring that as a first step we attempt to find conceptions that do not conflict. The first understanding would be a version of the unpromising argument for consistency from logical necessity mentioned above. The second interpretation of commitment—that we approach value enquiry with the aim of producing consistency—is more persuasive; but since 'commitment' here describes an aspiration, rather than giving a reason to think that aim is realizable, it cannot support an assertion of consistency in advance of an enquiry into whether values actually are consistent. Even aiming for consistency we may find that it cannot be attained (cf. Dworkin 2006, p. 162; 2011, p. 263). If the commitment is a procedural rule then I do not think it contradicts my reading of Dworkin's view of consistency as a hypothesis, since consistency might well be a hypothesis which Dworkin believes we should hope is true and strive to make correct if possible (as we might hope that the hypothesis that people can coexist peacefully is correct and attempt to make it correct, while recognizing that it might be false). Consistency would then be an ambition, but one which may turn out to be unrealizable. If consistency as a hypothesis is falsified, then consistency as an aim is futile.

Given the *prima facie* scope for conflict between the conceptions of values that we initially bring to the table, an argument against conflict will have to appeal to a theory purporting to explain how those initial conceptions can be *revised*. This is the strategy endorsed by Dworkin in the form of his interpretive theory (henceforth 'interpretivism'). *Justice for Hedgehogs* appears in some respects to be an extension of Dworkin's interpretive account of the law to the sphere of morality (see Dworkin 1986a, pp. 146–177; 1986b, especially chs. 2, 3 and 7).² I lack space here to consider the merits or otherwise of interpretivism in law or literature; regardless, its

¹ Thanks to an anonymous reviewer for suggesting this interpretation.

² The relevance of Dworkin's legal theory here was brought to my attention by David Archard.

extension to morality stands in need of justification: Dworkin argues for interpretation in the law by analogy with interpretation of literature on the basis that both domains have an external purpose in light of which their practice can be interpreted (Dworkin 1986a, pp. 141–145, 149–152, 158–162; 1986b, pp. 47, 58–59, 62, 87–88, 92–93, 228–238). Given this, it may be that interpretivism is peculiarly applicable to law and literature since these practices are human constructs, whose purposes are given by the rationales for our engagement in them: regulation of people's conduct and aesthetic appreciation respectively (Dworkin 1986a, p. 160). It is not obvious that there is an appropriate external purpose to morality in light of which to interpret moral concepts. This does not prove that an external purpose cannot be found, but it indicates a potential difference between morality on the one hand and law and literature on the other, which should serve to caution those who are persuaded by interpretivism in the latter spheres not to assent too quickly to its applicability to morality, since there may be disanalogies between these fields (cf. Williams 2001, pp. 98–99). In the remainder of this section I outline the structure of interpretivism as applied to moral values before critiquing it in the following sections.

Dworkin begins from the idea that values do not conflict (Dworkin 2001a, pp. 87–88). Working on this assumption he then proposes an explanation eschewing value conflict to be deployed in cases where values apparently enjoin arrangements which would compromise other values. Dworkin proposes that apparent conflicts actually indicate a misinterpretation of one or more of the values involved (Dworkin 2011, pp. 118–120; see also Dworkin 2001a, pp. 88–89; cf. Plaw 2004, p. 109). Therefore he recommends that when values appear to conflict, we should look to reinterpret our conceptions of those values. The process by which this new understanding is reached forms the core of his theory:

What is the right thing to do? Is the apparent conflict real? These questions cannot be so independent as my distinction suggested. The first question requires us to think further, and the way we think further is to further refine our conceptions of the two values. [...] We reinterpret our concepts to resolve our dilemma: the direction of our thought is toward unity, not fragmentation. (Dworkin 2011, p. 119)

According to Dworkin, the thought process through which we determine what we ought to do also illuminates the proper conceptual character of our values. Faced with an apparent value conflict and committed to avoid it, something has to give. Dworkin proposes to relieve the pressure by conceding that we previously misunderstood (at least) one of our values. He insists, however, that interpretivism does not grant a licence to simply reinterpret our values in whatever way suits our overall value system: the interpretation must also be independently plausible (Dworkin 2011, pp. 120–121). By asking whether a reinterpreted conception of a value coincides with existing ideas about the implications of the value, Dworkin hopes to establish the independent plausibility of the reinterpreted value while also avoiding the apparent conflict.

Summarizing Dworkin's approach, it takes something like the following form³:

- (i) We hypothesize that values do not conflict (Dworkin 2011, p. 331).
- (ii) In light of (i), if apparent value conflicts arise then we enquire whether there is a possible reinterpretation of one or more values which would relieve the conflict (Dworkin 2001a, pp. 83–88; 2001b, pp. 253–256; 2011, pp. 118–119).
- (iii) If yes to (ii), then we consider whether this interpretation is also consistent with our prior commitments regarding that value (Dworkin 2001a, pp. 87–88; 2001b, p. 256; 2011, pp. 5, 120–121; cf. Dworkin 1986b, pp. 66, 75–76).
- (iv) If yes to (iii), then we adopt the reinterpreted conception of the value identified in (ii) and thus dissolve the apparent value conflict (Dworkin 2011, pp. 119–120; cf. Dworkin 1986b, p. 66).

Now, as a first port of call when an apparent value conflict looms this is probably a sensible approach. We don't want conflict, and if we can avoid it by reinterpretation while retaining the appeal of the relevant values then we lose nothing by adopting the reinterpreted conceptions and we gain consistency. Reinterpretation may succeed in many cases, and to this extent interpretivism may constitute a useful initial approach to apparent value conflict, particularly for political values, where the consequences of a confirmed conflict may be grave. But the strategic priority of interpretivism does not guarantee the success of reinterpretation for any given instance of value conflict, and it remains the case, as I will argue, that some instances are not susceptible to reinterpretation, however much we might wish they were.⁴ It looks like the most problematic step will be (iii). It is trivially true that for any values that appear to conflict, *some possible* reinterpretation of one or more of them will yield consistent conceptions, since without a restriction on the scope of reinterpretation one could simply vary the values in whatever *ad hoc* way would alleviate the conflict (e.g. reinterpreting security to mean safety from all violence but those sorts of terrorist actions whose prevention would require violation of civil liberties).⁵ But, as Henry Richardson (1994, p. 112) cautions, 'one must start with the actual values of human beings, rather than allowing one's sense of the possibilities to be influenced by merely conceivable constellations of values'. This is not to say that values can never admit any shifts from their pre-theoretical conceptions; considering and revising our understandings of values is part of almost any moral theory. But if value-conceptions are to retain the moral significance enjoyed by pre-theoretical conceptions, any reinterpretation must bear at least a reasonable degree of resemblance to them. The precise degree of resemblance required will vary from one value to another, but I hope it is uncontroversial to

³ This structure corresponds roughly to the three stages of interpretation set out by Dworkin (1986b, pp. 65–66).

⁴ Thanks to an anonymous reviewer for suggesting that I clarify the practical utility of interpretivism. John Holme and Miriam Ronzoni also pressed me on this point.

⁵ Dworkin (2011, p. 118) characterizes security as a desideratum rather than a value, but it does not seem implausible that citizens have a moral right to protection from violence.

suggest that there is a point beyond which a reinterpretation will no longer be recognizable as the same value, as in the case of the reinterpretation of security above. In order to anchor the reinterpreted conceptions of values to people's actual commitments, step (iii) is required as a failsafe against 'stretching and lopping conceptions of the great political virtues so that they neatly fit one another' (Dworkin 2011, p. 5; cf. Dworkin 1986b, p. 52). Having outlined the structure of interpretivism, I return here to its resemblance to a scientific theory. Even accepting a commitment to consistency in the sense that it is regarded as an aspiration, if interpretivism is also to have persuasive power rather than being merely an assertion, then it is not obvious that presenting a way to revise value-conceptions will overcome the problem of value conflict. In line with the reading of Dworkin's theory as treating consistency as a hypothesis, interpretivism establishes nothing until put to the test; that is the purpose of the next section.

A Test Case

In this section I present a case of value conflict for which, so I claim, no plausible reinterpretation is possible. Specifically, no reinterpretation of either value can be found which will ensure consistency with the other while also remaining independently plausible as required by step (iii). If I am correct that reinterpretation fails in this case, this will show that Dworkin's view is unpersuasive, since in at least some cases conflict persists. Though there are certainly many cases where, as Dworkin contends, the appearance of conflict is due to a faulty understanding of one of the values, for reasons I discuss in the final section, we should expect reinterpretation to fail in many other cases of value conflict. This section presents one such case.

Dworkin's commitment to both freedom and equality is laudable, and if he limited himself to the claim that those two values can be made consistent under the requisite political conditions I would have a lot of sympathy for his position. But unfortunately he goes further and claims not just that freedom and equality are not inherently (logically) conflicting, but that they cannot conflict, even under unfavourable circumstances (Dworkin 2000, p. 123; 2011, p. 119). Contrary to Dworkin's position, I think it is possible for freedom and equality to come into conflict even on his own conceptions of these values. Dworkin (2000, pp. 73–83) famously endorses a conception of equality as equal distribution of resources, according to which compensation for brute bad luck should be provided by society at a level determined by the hypothetical insurance premium a rational consumer would choose to pay. An egalitarian society must provide universal healthcare at the level provided by a hypothetical insurance premium which 'most prudent people' (Dworkin 2000, p. 315) would opt to pay. For reasons of prudence and efficiency that healthcare should be provided as treatment rather than as a monetary payment (Dworkin 2004, p. 361).

Dworkin recognizes that many readers will surmise that this commitment to universal healthcare risks conflict between equality and freedom on the presumption that taxation reduces freedom (Dworkin 2001a, p. 80). In line with step (ii) above, Dworkin's response to this danger is to reinterpret liberty to mean 'freedom to do

whatever you like so long as you respect the moral rights, properly understood, of others' (Dworkin 2001a, p. 84). On the premise that everyone has a right to basic healthcare he then excludes taxation from the set of liberty violating practices (Dworkin 2001a, pp. 84–85). Accept, for the moment, that this reinterpreted conception of liberty passes step (iii), so that Dworkin's strategy is successful in averting the looming conflict (I discuss this move further in the next section). Nevertheless, even on the revised conception, there remain cases which are not due to taxation where the commitment to universal basic healthcare causes conflicts between liberty and equality. Suppose the following situation: an egalitarian society whose citizens each hold the resource bundles they would obtain under Dworkin's hypothetical auction is faced with an epidemic of a debilitating disease. In this society there is a definite group of persons who possess the necessary natural medical talent to treat the disease, and whose number is sufficient that if they all practice the epidemic could be curbed. However, despite the good faith attempts of the state, it is unable to secure their medical labour on a regular basis. While they are happy to treat infected persons in summer, the medically talented persons are unwilling to provide treatment in winter because they all suffer from severe seasonal affective disorder. However much money is offered, the potential doctors will not sign contracts which oblige them to work in winter. Consequently a sizable proportion of the population must go without treatment for the disease, the members of the untreated group being determined randomly by the season in which they are infected. Now, on Dworkin's view, if the state fails to provide universal (equal) basic healthcare then it violates its citizens' claims, rooted in the value of equality, to that provision. But given the intransigence of the talented, the only way it could provide that healthcare would be to force them to practice in winter, and this would violate their freedom 'to do whatever you like so long as you respect the moral rights, properly understood, of others' (Dworkin, 2001a, p. 84); whatever freedom means, it seems uncontroversial that it precludes forced labour. Given the circumstances, the state must, therefore, sacrifice either its citizens' equality-based claim to basic healthcare, or the liberty-based claim of the talented group not to be forced to labour.

How might Dworkin respond to this scenario? To forestall one possible objection, it cannot be contended that the conflict is not genuine because it is only the wrongdoing of the potential doctors that lands society in this mess. Even if the talented are behaving unreasonably, they have a freedom-based right to choose whether or not to perform socially beneficial labour (cf. Dworkin 2000, p. 158). Though their behaviour is socially suboptimal it does not constitute an injustice (and given the stipulated psychological condition of the potential doctors it hardly seems clear that their behaviour is even morally objectionable). The ill citizens' claim is not against the talented intransigents, but against the state; hence the potential doctors' behaviour can be treated as an exogenous circumstance. On Dworkin's 'statocentric' (Cohen 2000, p. 165) theory of equality, in an unjust situation 'it is not the [individual's] duty to do what the state should be doing' (Cohen 2000, p. 164). The state must commit an injustice on either course of action, since both the sick persons and the unwilling doctors have value-based claims against it which it cannot jointly satisfy. If it allows the sick to suffer it fails to meet their claim from equality,

but if it secures their treatment by force it violates the claim of the talented not to be forced to work in ways they choose not to. One way to secure equal healthcare provision in this situation would be to refuse treatment to patients in summer in order to put them in the same boat as those who succumb in winter. However, this inhumane proposal might implicate the state in a failure to treat citizens equally, since it would then fail to provide healthcare for those whom it could treat, even though it would treat everyone if possible. This would indicate an inequality between its intentions towards those infected in winter, whom it *would* treat if possible, and those infected in summer, whom it could treat, but *will not* do so. Another option might be to offer monetary compensation for those unable to access healthcare. However, it can be stipulated that they would still envy the resource bundles of those who obtained treatment, even if extravagant compensation packages were provided, in which case equality could not be satisfied by compensation and would demand treatment in kind. Nor is this a straightforward case of bad brute luck attracting compensation, since it is the conjunction of brute luck concerning the time of infection with inconsistent state provision which causes the inequality. Alternatively, Dworkin might attempt to reinterpret equality along the lines of his previous reinterpretation of liberty so that equality means 'equal treatment, so long as this can be secured without compromising the moral rights of others'. This, however, would occasion a problematic circularity in the reinterpreted conceptions of freedom and equality: as things stand, with liberty defined as 'freedom to do whatever you like so long as you respect the moral rights, properly understood, of others' (Dworkin 2001a, p. 84), the content of freedom is defined by the list of moral rights deriving from equality. Taxation is no invasion of liberty because equality demands a moral right to welfare which requires taxation for its satisfaction (Dworkin 2001a, pp. 84–85). But if equality were made rights-conditional, then equality could no longer play this defining role in the conception of liberty, since we could not know the content of equality until we knew what liberty-derived rights it must respect. But we don't know what rights liberty demands until we have the list of equality-based rights. Equality would be defined in terms of liberty, but liberty is indeterminate until equality is specified, giving rise to a definitional circle. And if Dworkin proposed to avoid this circle by balancing or prioritizing the moral interests in equality and freedom against one another, then that would amount to a concession that consistency can only be obtained at the cost of failure to satisfy (iii), since, by hypothesis, one of the values would no longer accord with the full interest it previously reflected.

The case outlined above is a useful test of interpretivism because it seems to me that no reinterpreted conception of freedom or equality that makes them consistent in this scenario will also pass the requirement of step (iii) that it respect our prior commitments regarding that value. In order to make equality and freedom consistent in this case, either equality would have to allow unequal provision of basic healthcare, or freedom would have to allow forced labour, neither of which reinterpretations seem likely to pass step (iii). A conception of equality which allows that people are accorded equal concern by a state which allows some to endure severe physical suffering while arbitrarily sparing others the same fate hardly coincides with what most egalitarians believe equality requires. Likewise,

even in these extreme circumstances, no conception of freedom which rules that forced labour causes no unfreedom will satisfy our expectations regarding that value. Moreover, those conceptions have already undergone substantial reinterpretation to reconcile them in the taxation case (see Dworkin 2000, pp. 120–183). Though I do not think they are the most persuasive, I have interrogated Dworkin's own conceptions of freedom and equality here to show that, even on charitable conceptions, interpretivism cannot produce conceptions of freedom and equality which avoid conflict in all cases.

This case shows that in at least some situations Dworkin's procedure cannot achieve successful reinterpretation in a way that accords with both (ii) and (iii). Our pre-theoretical conceptions of values give us *prima facie* reason to think that some values may conflict with one another and these initial conceptions cannot always be plausibly reinterpreted to yield consistency. So, on my reading of Dworkin's argument as analogous to a scientific investigation, this case falsifies his hypothesis of consistency. Where does this leave interpretivism? Defending Dworkin, Dale Smith has objected to the presentation of seemingly falsifying cases as undermining his theory:

[E]ven if the best available understanding of two values indicates that they do conflict, Dworkin may respond that the conflict is only apparent and would disappear if we developed better interpretations of those values.

(Smith 2012, p. 388)

There is some evidence that Dworkin endorses this move (e.g. Dworkin 2013, p. 140), but he shouldn't. Deferring to some possible future interpretation of the values would mean abandoning the analogy with science. Disregard of countervailing evidence is, as Karl Popper (1965, sections 6, 11, 19–20, pp. 40–42, 53–54, 78–84) argued, a cardinal sin in scientific practice, since it renders a theory unfalsifiable. Since it looks like values sometimes conflict and it is not a logical truth that they cannot, Dworkin must, perforce, adopt the scientific analogy to support his conclusion; but his theory then falls foul of falsifying cases like that described in this section. Notwithstanding the hope that many conflicts may turn out to be avoidable, and the propriety of *attempting* reinterpretation where they appear, reinterpretation does not offer a magic bullet solution to conflict. I will next attempt to consolidate this conclusion by analysing the conceptual implications of interpretivism. I hope to show why cases of value conflict intractable by reinterpretation are likely to occur.

Reinterpretation of Values and Moral Thought

Thinking further about the values involved in an apparent conflict does not present itself as a likely candidate for the most fruitful means of reconciliation. More likely, when faced with an apparent conflict one might try to think of a *practical* solution, i.e. a way of acting differently or altering the circumstances in order to avoid the values coming into conflict. Merely thinking about the values themselves does not look like a promising panacea, yet this is exactly what Dworkin (2011, p. 119) recommends. So how can thinking about values produce a change in their relations without altering the circumstances? The answer is that reinterpretation eliminates

those aspects of value-conceptions which conflict. As we have seen, there are in fact cases in which values cannot be reinterpreted to avoid incompatibility without sacrificing their plausibility. In what follows I explicate the source of this problem by elaborating an effect I call 'value erosion', which results from reinterpretation. Value erosion occurs when consistency is achieved at the cost of ignoring moral commitments which rightly belong in our value-conceptions; hence cases like that presented in the previous section are bound to arise where we must choose between consistency and the conceptual authenticity of value-conceptions.

The properties of value-conceptions generally respond to our moral commitments regarding some single feature of interpersonal relations (or a small set of closely aligned features).⁶ If freedom responds to the view that people should not force one another to do things, then the conception of freedom mandated by this commitment will be one which specifies types of situations which involve coercion and condemns them. A conception of freedom is authentic when it plausibly represents the conceptual implications of the moral view that coercion is to be avoided. The commitments grounding different values are independent of one another. For freedom the commitment might be that people should not use force against each other, for equality perhaps the view that people should hold equivalent bundles of resources. Since these immediate commitments are independent of one another, the conceptions of different values will also be formed, pre-theoretically at least, without reference to one another, meaning there is no reason to expect that they cannot conflict with one another in application (Williams 2001, p. 95). Each value-conception will describe which social arrangements respect or compromise the moral commitments affirmed regarding a different aspect of interpersonal relations. A conception of freedom will respond to moral commitments regarding force in interpersonal relations, a conception of equality will respond to commitments regarding the acceptability of comparative differences in the fortune of different people's lives, and so on. Though the specific sets of commitments attributed to a value by different persons may differ, since each of those sets is independent of that person's commitments regarding other values, it is unlikely that anyone's commitments with regards to any two values will perfectly overlap. Indeed, if the commitments behind different values were identical it would not be clear why we have different names for those values (cf. Williams 2001, p. 92). But reinterpretation requires that when forming a value-conception we do so in light of our commitments regarding at least *two* disparate features of interpersonal relations and derive a conception that indicates what states of affairs will respect those elements of the two sets of commitments that are consistent. Therefore, those commitments regarding one feature of interpersonal relations which conflict with commitments regarding another are eliminated from the relevant value-conception in order to

⁶ For brevity's sake I refer to 'our' conceptions of values, but in any moral community there are in fact various conceptions of any value at large. I do not mean to suggest that everyone shares a single conception of each value, but that any person's conception will be negotiable only to a certain extent without losing something of what she regards as important in that value. Consequently, for many value conflicts there will be a group of people whose conceptions of the values in question will not admit reinterpretation in a way that would alleviate the conflict without loss of moral content. Thanks to an anonymous reviewer for pressing me to clarify this point.

obtain consistency between values; hence, the content of the value-conception is eroded in line with the content of the conflicting value.

Value erosion can be seen at work in Dworkin's approach to resolving the putative conflict between freedom and equality. Pre-theoretical conceptions of freedom and equality are formed independently of one another and at risk of conflict. Dworkin recognizes, for instance, that on Berlin's negative conception of freedom as 'freedom from the interference of others in doing whatever it is that you might wish to do [...] it is very plausible that this commitment will often conflict with other commitments, including even minimal egalitarian ones' (Dworkin 2001a, p. 84). Therefore, he seeks to provide a conception which avoids this risk. In order to immunize freedom from conflict with equality, he reinterprets freedom as 'liberty', meaning 'freedom to do whatever you like so long as you respect the moral rights, properly understood, of others' (Dworkin 2001a, p. 84).⁷ Since Dworkin believes that people ought not to be free to amass extreme wealth at the expense of social failure to meet others' basic needs (see Dworkin 2011, p. 375), this capacity does not enter into his reinterpreted conception of liberty (Carter 1999, p. 72). Freedom is made compatible with equality by purging from the conception those arrangements which protect people from coercion but also permit inequality.

The effect of eliminating moral commitments on the resulting value-conceptions can be seen by considering scenarios designated as cases of unfreedom by a pre-theoretical conception of freedom but denied this status by the reinterpretation. Take Robert Nozick's (1974, pp. 160–164) Wilt Chamberlain example of an apparent conflict between freedom and equality. The idea here is that freedom conflicts with equality because it is only by interfering coercively in individuals' lives that society could prevent Chamberlain from accumulating wealth by charging people to watch him play basketball. On Dworkin's reinterpreted conception of liberty there is no value conflict here because liberty does not include the right to accumulate excessive wealth. But what will happen if Chamberlain tries to charge people to watch him play? The coercive power of the state will be brought to bear to prevent him and his fans from completing their transactions. Imagine you see the police carting the hapless athlete off to jail for performing in black market basketball games. To maintain that preventing equality reducing transactions cannot conflict with freedom you must be willing to say, with a straight face, that the police are not infringing his liberty—despite the fact that they are *forcibly detaining him*. Coercion just is the infringement of freedom, and even if we think the infringement is justified it is an infringement nonetheless. Dworkin's reinterpreted conception of liberty will fail (iii) because denying that a jailed prisoner is unfree stretches the concept of freedom to breaking point.

Despite the apparent intuitive cost of denying that justly imprisoned persons are unfree, Dworkin (2000, p. 125; 2011, p. 367) argues that this is merely a function of linguistic intuitions concerning the 'flat' non-normative sense of liberty and that no morally significant content is lost by reinterpreting freedom in this way

⁷ On the other side of the coin, Dworkin (2001b, p. 254) characterizes his 'dynamic' conception of equality as set out in *Sovereign Virtue* as a reinterpretation of the 'traditional [...] flat' socialist conception of equality of outcome.

(Dworkin 2001a, pp. 88–89).⁸ This is because Dworkin denies that freedom is valuable in a general sense (non-specifically valuable) rather than just in virtue of the specific good things it allows us to do (Carter 1999, pp. 33–34; cf. Dworkin 2000, pp. 129–130; 2011, p. 4). Since everyone agrees that some things people might otherwise do must be prevented (such as murder), Dworkin claims that there can be no moral loss in such preventions and concludes that only freedom to act justly has value and not freedom *simpliciter* (Dworkin 2001a, pp. 88–89). However, this claim that freedom lacks non-specific value is certainly not uncontroversial. I do not have space here to provide a detailed defence of the non-specific value of freedom, but I will briefly suggest some reasons that might support it. Hopefully the reader will accept at least one of these, but even if not, I think they give enough plausibility to the claim that freedom is non-specifically valuable to suggest that many people will endorse it, and that they will not do so unreasonably. If this is the case then at least those who do accept freedom's non-specific value will feel that some important moral content is lost by reinterpreting freedom to avoid conflict.

So, here are some reasons one might think that freedom is non-specifically valuable: (1) people often value choice, probably for the opportunity to exercise agency. Even rightful prevention reduces people's options, so situations of greater freedom are (*pro tanto*) preferable (Carter 1999, p. 42). (2) Unfreedom in the 'flat' sense implies violence; in standard cases someone prevents another person doing something by deploying or threatening violence against her. If one objects to violence as such (believes it is intrinsically disvaluable), even while believing it is sometimes necessary all things considered, then one should subscribe to the corollary that freedom is non-specifically valuable. (3) Freedom may be one component of an intrinsically valuable complex phenomenon. Even if freedom is not intrinsically valuable itself, it might be constitutively valuable as a necessary condition of autonomy, which might be considered intrinsically valuable (Carter 1999, pp. 58–62). It is plausible that to be autonomous an agent must have the opportunity to do what she does not choose as well as what she does, in which case more freedom is *pro tanto* better, whatever it is freedom to do. If any of these reasons (or another I haven't listed) is plausible then Dworkin is wrong to think that only the freedom to do what we should be permitted to do matters; rather, there is some disvalue in any prevention, even one that is right on balance. For anyone who accepts this view Dworkin's reinterpretation will involve not just linguistic infelicity but more importantly the loss of some moral significance for the value of freedom.

Rather than allowing the independent concern that motivates our interest in freedom to shape our conception of it uninhibited, which might result in conflict with equality, Dworkin jettisons this concern in favour of consistency. Consequently, where before we possessed two values reflecting two independent concerns, we are now left with one value—equality—and a hollow conception of liberty which reflects the concern that people should not coerce one another only to the extent that equality can be secured without coercion. In order to make room for the concerns behind equality within his conception of liberty, Dworkin has eroded the

⁸ Thanks to an anonymous reviewer for raising this objection.

role of the concerns motivating freedom where the two groups of concerns clashed. Dworkin's reinterpreted conceptions will often fail to chime with the moral concerns behind values because they are not derived from those concerns alone, but from those concerns *and* the concerns behind any values that might conceivably conflict with them. His conception of freedom is unpersuasive because it is derived from the concerns behind freedom *and* the concerns behind equality. Anyone who is interested in freedom for the reasons that motivated the original conception is going to be dissatisfied with the reinterpreted conception, which reflects those concerns only half-heartedly. It approaches the concerns of freedom with one eye on those of equality, with the unattractive consequence that it fails to treat cases like the jailing of Wilt Chamberlain as instances of unfreedom (Carter 1999, p. 72; see also Flathman 1987, pp. 106–107); in this sense it is conceptually inauthentic. In order to produce consistency, Dworkin's value-conception must disregard certain cases flagged up by the initial moral commitment, but then the reinterpreted conception only reflects some fraction of the full complement of concerns it should reflect.

Conceptions of values are, of course, disputed, and I do not mean to suggest that there is some precise and immutable conceptual specification of each value which cannot be altered in any way without a complete loss of meaning. However, recall the ludicrous reinterpretation of security to mean 'safety from all violence but those sorts of terrorist actions whose prevention would require violation of civil liberties' described above. As in that case, my claim is that if the content of a value-conception differs too greatly from what it is intuitively understood to include, then we no longer have the same reason to care about it morally and are unlikely to identify it as representing the same value. Part of what we care about when we talk about security is precisely safety from terrorist attacks, and a conception of security which ignored this would fail to capture our moral interest in security. Similarly a conception of freedom should allow us to identify even rightly jailed prisoners as unfree. That said, it might be argued that at least when it comes to politics, consistency has certain benefits, such as promoting stability or averting social conflict, which could outweigh conceptual authenticity so that it would be preferable to opt for consistent conceptions of values even if these are necessarily inauthentic.⁹ There are two reasons why I think this policy would not rescue Dworkin's view. First, simply as regards Dworkin's own project, he wants to explain how values relate, not just to offer a politically expedient policy (Dworkin 2001a, p. 76; 2011, pp. 5–6). Second, it is not clear that sacrificing authenticity to consistency would in fact reap the mooted benefits. As Bernard Williams (2001, pp. 94–95, 99–103) observes, responding to those who lose out in a value conflict by asserting that their grievance is mistaken is unlikely to encourage them to drop their claims. More likely they will feel ignored or patronized and this will provoke them to greater vehemence in affirming their claims. Might it not be better for stability and social harmony to admit that one value has been compromised and explain why, regrettably, this was necessary, than to sweep this under the carpet while everyone concerned still feels the loss? As I mentioned above, searching for reinterpretations which satisfy step (iii) may be a sensible first response to apparent value conflict

⁹ Thanks to an anonymous reviewer for this suggestion.

since it would avoid conflict where possible, but such reinterpretations cannot always be found, and in cases of irresolvable conflict it is better to acknowledge the impossibility of satisfying all claims than to deny this.

Dworkin is well aware that his theory will come under fire from objections like that I have just developed and replies as follows:

[W]hat in the world could that supposed plain fact [of conflict] consist in? Kindness and honesty cannot *just* have one content or another, because moral claims cannot be barely true.
(Dworkin 2011, p. 120, Dworkin's italics)

The claim here is that since there are no brute moral facts to fix the proper content of value-conceptions there is no independent source by reference to which the appropriate content of a value-conception can be determined. Counterarguments like mine which rely on asserting the inappropriate content of some reinterpreted conceptions thus have no privileged value-conception with which to compare the reinterpretation. Dworkin concludes from the absence of foundational brute moral facts that there can be no independent source of moral input to fix the content of values other than coherent interpretation. He comes back to this claim again and again in replying to objections that interpretivism whitewashes conflict (see, e.g. Dworkin 2010, p. 1079; 2011, pp. 120, 263; 2012a, b). The theme is that if there is nothing external to morality to fix values' contents, then that content cannot be undermined.

I am happy to agree with Dworkin that there are no brute facts about morality which fix how values can legitimately be interpreted (see, e.g. Dworkin 2001b, p. 255; 2010, p. 1079; 2011, pp. 114, 120, 263; 2012a, b). It is true that '[v]alues are not like rocks that we might stumble across in the dark. They are not just stubbornly there' (Dworkin 2011, p. 114). But it is too strong to infer from the absence of foundational moral facts that there is *nothing* to fix any prior content for how we appropriately understand values. The natures of values are not fixed by external objective reality in the way that the natures of rocks are fixed. But would we understand someone who claimed that prisoners are free and that slaves enjoy equality with their masters? The fact that the intelligibility of moral claims becomes stretched to breaking point in extreme cases like these indicates that though value conceptions are not fixed by brute facts, there is some independently fixed content which is necessary to the appropriate understanding of these concepts. I don't think this will be surprising, nor should it be, because there is an obvious external source for values' contents in moral psychology.

Like many psychological phenomena we should expect that though moral claims do not map on to any objectively real entities, our deployment of them remains broadly determined by our pre-existing moral intuitions and vocabulary. The intuitions which form the ultimate basis of morality do not directly or precisely determine the content of political values, and our specific moral beliefs may be as much the product of society and culture as moral psychology. But the reason that we care about how people treat one another at all is that we have moral intuitions regarding these relations which identify certain ways of acting as right or wrong. A person's intuitions regarding the moral propriety of particular actions set limits

within which conceptions of values must remain if they are to accord with what she endorses or objects to in people's treatment of one another. For instance, some people may well be happy to accept that no *pro tanto* wrong is done by coercive taxation, in which case Dworkin's reinterpreted conception of freedom would remain within the intuitive limits prescribed for their values by their moral psychology. But it seems likely that far fewer people could accept that forced labour is not wrong, so a conception of freedom which embraced forced labour would fall outside their intuitive limits and be rendered implausible. I would no longer care about freedom (or at least not in the same degree) if that value was understood to allow forced labour, since I object to this practice on intuitive grounds. Even if society adopted a new conception of freedom which allowed forced labour I would still maintain a moral objection to it, though I might be linguistically disabled from objecting in terms of freedom. My point is that Dworkin's reinterpreted value-conceptions cannot be objectively false in the way that it is false to say that the moon is made of cheese; but if you use reinterpretation to stretch the conception of a value too far you will lose sight of the reasons you cared about it in the first place. Your use of the term has changed, but neither the empirical phenomenon, nor your moral-psychological reaction to it will have done. Dworkin's step (iii) of reinterpretation tacitly acknowledges this, but if this is to be understood as anything more than lip service to independent plausibility it must also be acknowledged that step (iii) excludes reinterpretation in at least some cases, as with that described in the previous section. So the absence of brute moral facts does not grant a blanket licence for reinterpretation of moral values with its accompanying value erosion and Dworkin's theory is liable to miss the point of the values we care about.

The problem with value erosion is that those concerns which go unreflected do not disappear because we say they do not count as freedom (or equality etc.). *If* we were worried that Wilt Chamberlain's freedom was being infringed in the first place, agreeing not to call his jailing unfreedom is unlikely to assuage our worry. The problem with reinterpretation is that, while it may succeed in producing consistent and convincing value-conceptions in some cases, in many others consistency can only be accomplished by value erosion. But value erosion does not address the real issue, which is the infeasibility of pursuing two values at once. Rather, it neglects the issue by ignoring those concerns which produce the conflict and consequently curtailing the meaning of one or both values. The implication of this is that the case I presented above is not isolated; we can expect that reinterpretation will frequently be unable to provide consistent reinterpretations of values without incurring a serious cost in terms of the conceptual authenticity and intuitive plausibility of those values. Dworkin has identified an approach to apparent value conflict which may help to bypass conflict, so long as step (iii) is properly adhered to as a barrier against the risk of value erosion. But he is wrong to represent this approach as a comprehensive solution, and if we want to show that our values are consistent we must go about this project through conceptual analysis of the implications of our moral commitments and empirical investigation of the feasibility of their joint realization. We will be rewarded with consistency on some occasions and not on others, and sometimes we will find that we can obtain coherence by implementing political solutions. But this will be an ongoing project,

and we should be wary of quick fix solutions that offer big value consistency at a discount rate of enquiry.

Acknowledgments For helpful written comments and suggestions I would like to thank Richard Child, John Holme, John O'Neill, Liam Shields, Tom Sinclair, Hillel Steiner and two anonymous reviewers. A version of this paper was presented at the Brave New World graduate conference at the University of Manchester on 13th June 2014, and I thank the audience for an informative discussion. This work was supported by the Economic and Social Research Council [ES/J500094/1].

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Baker, C.Edwin. 2010. In hedgehog solidarity. *Boston University Law Review* 90: 759–818.
- Carter, Ian. 1999. *A measure of freedom*. Oxford: Oxford University Press.
- Cohen, G.A. 2000. *If you're an egalitarian, how come you're so rich?*. Cambridge, MA: Harvard University Press.
- Dworkin, Ronald. 1986a. *A matter of principle*. Oxford: Oxford University Press.
- Dworkin, Ronald. 1986b. *Law's empire*. Oxford: Hart.
- Dworkin, Ronald. 2000. *Sovereign virtue: the theory and practice of equality*. Cambridge, MA: Harvard University Press.
- Dworkin, Ronald. 2001a. Do liberal values conflict? In *The legacy of Isaiah Berlin*, ed. Mark Lilla, Ronald Dworkin, and Robert Silvers, 73–90. New York: New York Review of Books.
- Dworkin, Ronald. 2001b. Do values conflict? A hedgehog's approach. *Arizona Law Review* 43: 251–260.
- Dworkin, Ronald. 2004. Ronald Dworkin replies. In *Dworkin and his critics*, ed. Justine Burley, 339–395. Malden, MA: Blackwell Publishing.
- Dworkin, Ronald. 2006. *Justice in robes*. Cambridge, MA: Harvard University Press.
- Dworkin, Ronald. 2010. Response. *Boston University Law Review* 90: 1059–1088.
- Dworkin, Ronald. 2011. *Justice for hedgehogs*. Cambridge, MA: Harvard University Press.
- Dworkin, Ronald. 2012a. Response to Colin Bird. <http://www.justiceforhedgehogs.com/response-to-colin-bird/>. Accessed 27 Jan 2014.
- Dworkin, Ronald. 2012b. Response to William A. Galston. <http://www.justiceforhedgehogs.com/response-to-william-a-galston/>. Accessed 26 Mar 2014.
- Dworkin, Ronald. 2013. Replies. *Analysis* 73: 139–146.
- Fallon Jr, Richard H. 2010. Is moral reasoning conceptual interpretation? *Boston University Law Review* 90: 535–550.
- Flathman, Richard E. 1987. *The philosophy and politics of freedom*. Chicago: The University of Chicago Press.
- Hampshire, Stuart. (ed) 1983. Morality and conflict. In *Morality and conflict*, 140–169. Cambridge, MA: Harvard University Press.
- Kalderon, Mark Eli. 2013. Does metaethics rest on a mistake? *Analysis* 73: 129–138.
- Nozick, Robert. 1974. *Anarchy, state, and utopia*. Malden, MA: Blackwell Publishing.
- Plaw, Avery. 2004. Why monist critiques feed value pluralism: Ronald Dworkin's critique of Isaiah Berlin. *Social Theory and Practice* 30: 105–126.
- Popper, Karl. 1965. *The logic of scientific discovery*. New York: Harper & Row.
- Richardson, Henry S. 1994. *Practical reasoning about final ends*. Cambridge: Cambridge University Press.
- Smith, Dale. 2012. Law, justice and the unity of value. *Oxford Journal of Legal Studies* 32: 383–400.
- Williams, Bernard. 2001. Liberalism and loss. In *The legacy of Isaiah Berlin*, ed. Mark Lilla, Ronald Dworkin, and Robert Silvers, 91–103. New York: New York Review of Books.