

Velleman on Love and the Ideal of Rational Humanity
Edward Harcourt

Faculty of Philosophy, University of Oxford, and Keble College, Oxford OX1 3PG, UK
edward.harcourt@philosophy.ox.ac.uk

David Velleman, *Self to Self* (Cambridge: Cambridge University Press, 2006), xiii+385 pp., ISBN 0521670241 (pbk)/0521854296 (hbk), £40.00 (£16.99 pbk).

This collection of David Velleman's papers in ethics and moral psychology from 1996 to 2005 (plus two new papers and an introduction) is so rich in food for thought that any review of manageable length must leave a great deal out. Rather than leaving out a bit of everything, I shall depart from usual reviewing conventions and leave out almost all of some things for the sake of leaving in more of others. So of the collection's three main themes, one – Velleman's revisionist interpretation of Kant – I shall ignore altogether, and another – personal identity and its relations to narrative and motivation – I shall discuss only briefly, insofar as it bears on the third. This third theme - set out mainly in 'Love as a Moral Emotion', 'The Voice of Conscience', and 'A Rational Superego' - is that it's possible to marry some Kantian thoughts about the role played in motivation by an ideal of oneself as a rational agent, some Freudian thoughts about internal agencies of self-regulation, and some still psychoanalytic but perhaps less Freudian thoughts about the part played by love and loving nurture in shaping the way these agencies take root and work within us. To make this attempted marriage succeed would be, in Velleman's words, 'to humanize Kant's ideal of ourselves as rationally autonomous'.¹

1. On Velleman's Kantian view of moral motivation, I conceive of myself as a rational agent, and this ideal motivates me to live up to it, and to reproach myself if I don't. But *how* do ideal self-conceptions motivate? According to Velleman, my ideal self-conception – which Velleman identifies with a (quasi-) Freudian ego-ideal – is an 'internalized object of love',² derived from loved and loving parental figures. So one might imagine, if one weren't attending carefully, that Velleman's 'humanizing' of Kant consisted in proposing that our love for those figures explains how the ideals derived from them motivate us. But this isn't what Velleman says: 'moral requirements motivate us via an ideal image of our obeying them'.³ And though Velleman does have more to say about the way ideals motivate, love does not play a lead role in the story. First of all the attitude most to the fore in Velleman's account of our relation to the ego-ideal is not love but reverence: 'the motive that induces us to obey the [Categorical Imperative] is our reverence for the ideal it conveys'.⁴ Secondly, we

tend to act in accordance with motives and traits of character that we conceive of ourselves as having. ... [T]his tendency is due to a cognitive motive, to find ourselves explicable and predictable.⁵

The motivation to live up to our self-conceptions generally is the intellectual one of wanting to make sense to ourselves, and moral motivation is to be seen as a special

¹ Velleman, 'The Voice of Conscience', *Self to Self*, p. 128.

² *Ibid.*

³ *Ibid.*, p. 127.

⁴ 'A Rational Superego', p. 131.

⁵ 'From Self-Psychology to Moral Philosophy', in *Self to Self*, pp. 243-4.

case of this, its specialness residing in the specialness of the self-conception involved – the ideal of ourselves as rational autonomous agents – not in the way the self-conception motivates.

Velleman's humanizing project therefore turns out to be a different and perhaps more modest one, that of explaining the psychological *origin* of 'what Kant called "reverence for ... the mere dignity of humanity"', and it's only in this genetic story, not in the story of how the ideal motivates, that love is involved: this reverence is 'our response to something that we have internalized from real people in the course of our moral development. ... The ideal of ourselves as rationally autonomous [is acquired] in the course of loving our parents.'⁶ However, in combining his intellectualist account of how ideals motivate with his psychoanalytically inspired account of where they come from, Velleman sets himself a challenge. For of course love has a reputation as a *oiseau rebelle*. In order to ensure that an ideal that results from internalizing loved others motivates us to act not just any old how but in line with the requirements of practical rationality, Kantianly conceived, Velleman presents a strongly moralized conception of love: love (more precisely, love of persons) is, Velleman argues, a response to the very same thing which elicits Kantian respect, the value a person possesses as an instance of rational nature.⁷

As Velleman admits, this cannot be correct as a fully general account of love, at least if there is such a thing as love for infants or cities or good causes, since these are not instances of rational nature. This limitation aside, is Velleman right about the love of one rational human being for another?

Since 'the capacity for a rational ... will is that better side of a person which constitutes his true self', the claim that not only respect but also love is a response to that value fits the intuition that 'we love people for their true and better selves'.⁸ There is surely something exactly right about the idea that love finds its way to people's true and better selves, although this will strike some as already burdening the notion with too much moral baggage. Nonetheless there are objections to the way Velleman develops the thought. First, since the property that elicits love is the same as the property that elicits respect, Velleman needs to explain why we respect more people than we love. Secondly, if love is elicited by a property that all rational beings share, it seems to follow that it cannot matter *which* rational being one loves, and this seems contrary to another expectation we have about love, that it involves focus on the particularity of the loved individual.⁹

To the second problem Velleman responds that the claim that we love on the basis 'of our personhood, in which we are no different from other persons' and the claim that loving someone is 'valuing them as special'¹⁰ only appear inconsistent if we fail to distinguish 'appreciating the value' of an object from 'judging it to have that value'.¹¹ Though we may *judge* a person to have the same value as that possessed by others, we

⁶ 'The Voice of Conscience', p. 128.

⁷ 'Love as a Moral Emotion', pp. 99-100.

⁸ *Ibid.*, for both quotations.

⁹ The objection is raised by Robin Dillon, 'Respect and Care', *Can. J. Phil.* 22 (1992), cited by Velleman, 'Love as a Moral Emotion', p. 102.

¹⁰ *Ibid.*, p. 101.

¹¹ *Ibid.*.

respond to that value by ‘an unwillingness to replace her or to size her up against potential replacements’, and this is just as we should expect things to be because the value we correctly judge her and others to have is a ‘dignity not a price’.¹²

Velleman may be right that the value of rational agency is ‘a dignity’ – that it is incomparable - and also that the value of each individual rational agent is incomparable with any other’s. However, to say that the value of being a rational agent is incomparable is surely to say that the value of being a rational agent cannot be compared with the value of any other *property* – intelligence, say, or being part of God’s creation. It is *not* to say that the value of different *instances* of this property cannot be compared with one another. Indeed insofar as two individuals instantiate the same (comparable or incomparable) value-property to the same degree, their value must be the same, and *a fortiori* comparable. So what explains the fact the incomparability in value of any two rational beings cannot, *pace* Velleman, be that they have a shared incomparable value-property – either their common rational agency or anything else. By the same token, to claim that love is a response to a value-property – albeit an incomparable one – that we all share does not explain why love involves treating the loved individual as incomparable.

As to the first objection – ‘why we love only some people’¹³ – Velleman’s reply takes the form of an account of the relation between the rational self, in respect of which everyone is the same, and the ‘manifest person’, i.e. the person embodied in flesh and blood: ‘We can see the [rational] person only by seeing him in or through his empirical persona’. But we can see each and every rational person only through their respective empirical personae, while what the reply needs is something in respect of which persons differ from one another. Perhaps this is why Velleman adds a second line of reply: ‘the value that makes someone eligible to be loved [i.e. rational autonomy] does not necessarily make him lovable in our eyes’ because ‘whether someone is lovable depends on how well his value as a person is expressed or symbolized for us by his empirical persona’.¹⁴

This looks to be a false move, thanks to the implied distinction between (real) eligibility to be loved and lovability ‘in our eyes’: if we love only a subset of the people we respect because we miss their eligibility to be loved, the fact that we don’t love more people is our mistake. But if it is a mistake, we must be able to conceive of correcting it, and yet (assuming all rational beings are entitled to respect) a world in which the set of loved persons and the set of respected persons coincided would not, it seems, be a world that contained personal love at all.

There are also other difficulties with the reply. Certainly if some ‘personae’ express each rational being’s (equal) value better than others, that explains why it’s not true that each rational being loves every other one. But it explains this by explaining, more fundamentally, why only a subset of those really entitled to respect are manifestly entitled to it, and thus leaves unexplained why only a subset of those manifestly entitled to respect are also loved (by any one person); alternatively, why one and the same subset of rational beings isn’t the object of all the love that’s going. Perhaps, then, there’s

¹² *Ibid.*

¹³ *Ibid.*, p. 106.

¹⁴ *Ibid.*

an implied relativity to individual lovers in how well empirical personae express or symbolize the value of persons loved (cf. the 'for us' in the last passage quoted). This allows the theory to acknowledge that we don't all love the same people. But it still doesn't seem to leave room for divergence between the objects of a given individual's love and the objects of his respect. Respect and love are supposedly elicited by the same property, so for any individual who is unloved (by someone) because their empirical persona expresses that property poorly (to that same someone), it's hard to see why that individual won't also be unrespected (again, by that someone). Velleman adds that 'we are constitutionally limited in the number of people we can love'.¹⁵ True: but the question is whether that fact can be made sense of by a theory on which love is elicited by the same property that elicits respect, given that *that* attitude is *not* subject to any constitutional limitation. I conclude that love is not to be analyzed as a response to the very same value that also elicits respect, namely the value we have in virtue of exemplifying rational nature.

The very close relationship between love and respect urged by Velleman is required by his idea that reverence is our response to our ideal of ourselves as rational autonomous beings, which ideal – the ego-ideal – is in turn 'an internalized object of love'. That is, it's required because Velleman combines an intellectualist account of motivation by ideal with a psychoanalytic story about how ideals get there: they are internal representatives of loved others. To cast doubt on the closeness of that relationship is to cast doubt on whether something that motivates us in the way in which, according to Velleman, an ideal of ourselves as rational and autonomous does *could* be an internalized object of *that* emotion, properly understood. I now want to explore another aspect of Velleman's attempted 'marriage of Freud and Kant',¹⁶ by focusing on the developmental history of the ego-ideal as he sees it.

According to Freud, morality governs us via the superego. Velleman rejects Freud's scepticism about morality, but nonetheless thinks Freud has something to contribute to Kantian theory, 'namely a story of moral development'. For there's a puzzle as to how we can be, as Kant says, both bound by and authors of moral norms. By invoking 'the familiar psychological process of internalizing other people', Freud makes this possibility 'concretely imaginable': 'the external authority of morality is represented as the authority of another person, the parent; [while] the autonomous exercise of that authority is represented as the assumption of the parent's role by a part of the self'.¹⁷

Developing this thought, Velleman assumes that if we picture internalized others in such a way that they can explain guilt, we won't be far off explaining the authority of morality, since Freud was right that 'a sense of having disobeyed that authority is a prerequisite to feeling guilty'.¹⁸ But Velleman is also vividly aware of the inadequacy of Freud's own account of guilt, which explains guilt via the internalization of an object of fear: 'fear differs from guilt and cannot come to resemble it just by being internalized'.¹⁹ So, as Velleman rightly asks, why do

¹⁵ *Ibid.*, p. 107.

¹⁶ 'A Rational Superego', p. 131.

¹⁷ All three quotations from 'A Rational Superego', p. 131.

¹⁸ *Ibid.*, p. 132.

¹⁹ *Ibid.*, p. 135.

the parents, and their internal surrogate, come to be conceived as authorities administering punishment rather than as arm-twisting bullies[?]²⁰

An alternative is that 'the superego can inspire guilt only because it is formed out of figures already conceived as authorities'.²¹ But if the process of internalization doesn't explain this authority, what does? Apparently, love. Parents are feared as disciplinarians, Velleman says, but they are also loved and admired, so the object internalized will be an object *both* of fear *and* of love and admiration. This object - the superego - punishes the child for falling short of the ideal. But because the superego (in its aspect of ego-ideal) is also loved and admired, it must meet the ego's standards of admirability. Hence its authority: this ideal is the ego's *own* ideal, so 'the ego thinks it is being criticized and punished for a failure to meet its own standards'.²²

But does Velleman's identification of the superego with the ego-ideal really explains the authority of the superego's demands, and thus render 'concretely imaginable' Kant's thought that we are authors of the moral law that binds us? Let's remind ourselves why this thought is puzzling. Consider the UK parliament which, notoriously, cannot bind itself. Why not? Take any law made by Parliament which presumes to bind it in future, but suppose Parliament doesn't subsequently comply with it. Is this non-compliance a case of *breaking* anything? Only if it is that law authoritative for parliament. But it isn't a case of breaking anything: precisely because Parliament is the supreme legislative body, it's merely a case of Parliament's changing its mind. It is not only not bound by its own laws, but not bound by them *because* they are its own. One can see why Anscombe stated so bluntly that 'there is no such thing as legislating for oneself'.²³ To make the norms invoked by the superego the ego's own seems therefore not so much to explain the superego's authority as to state the second horn of a dilemma. The superego was meant to show how to combine 'the external authority of morality' with 'the autonomous exercise of that authority'.²⁴ But it now looks as if either the norms invoked by the superego aren't the ego's, in which case morality is external and compelling, but a mere tyranny (the first horn); or (the second horn) they are, in which case it's clear why we accept the norms, but not at all how they compel us as from outside – that is, how our conforming to them is doing something we *have* to do.

There's also a problem lurking in the idea that the norms or ideals the superego invokes are the ego's own, since it implies that the ego *has* values prior to the internalization of the relevant figures – that is, prior to the formation of the superego. So even if the community of values between ego and superego succeeded in explaining the latter's authority, it would do so at the cost of denying us something else the Kant-Freud marriage was said to promise - 'a story of moral *development*'.²⁵ The concept of the superego doesn't therefore do much for the project of humanizing Kant. Invoking that concept, however, does not as it turns out represent Velleman's best attempt to carry this

²⁰ *Ibid.*, p. 136.

²¹ *Ibid.*, p. 137.

²² 'A Rational Superego', p. 140.

²³ In 'Modern Moral Philosophy', *Collected Philosophical Papers vol. 3: Ethics, Religion and Politics* (Oxford: Blackwell, 1981).

²⁴ 'A Rational Superego', p. 131.

²⁵ *Ibid.*, p. 131; my italics.

project through. For at this point in the dialectic Velleman simply accepts that the superego is not 'a final or ultimate authority in the ... psyche', but owes its authority to the ego's 'evaluative faculty', 'an independent faculty of judgment as to whom or what to admire'.²⁶ For Velleman, this is not to give up the possibility of a developmental story about moral self-regulation, for now the interesting developmental story will be that of the development of the ego's evaluative faculty itself. So can Velleman's account of *this* development explain how the demands of morality can be both authored by and authoritative for us?

Velleman's answer to his question 'how does a child acquire ... norms [of rationality]?' seems to be that *because* love is a response to the value others have as 'specimens of [rational] humanity', a child's love for its parents is an exercise of its evaluative faculty – it's just by loving its parents that the child discriminates between the dross of their size or beauty and the 'pure gold' of their rationality. So, since we internalize what we love, the child ends up internalizing their rationality as an ideal and thereby 'acquires reason and good sense'.²⁷ That, at least, is *part* of the picture. However, when we peer further into the mechanisms by which the child acquires an 'ideal of practical reason',²⁸ things seem to work differently. When parents 'care for [their children] with a wise good will', Velleman explains, they subordinate their own interests to the needs (etc.) of another. But this is to say that they exemplify taking another person as an end. Now the Categorical Imperative 'just *is* a description of the capacity to take persons as ends'. So the child need do no more in order to 'internalize the Categorical Imperative' than to make his loving parents his ego-ideal.²⁹

As regards the authority of the ego-ideal, this late chapter in Velleman's account of moral development seems to leave us right where we were. Velleman says that the ideal thus internalized 'carries genuine moral authority'.³⁰ But why? The only answer I can find is that it, and the parental injunctions to live up to it, embody the Categorical Imperative.³¹ That account of the authority of the ideal is as good, though only as good, as a Kantian account of the authority of morality. Disappointingly, however, the psychology of internalization apparently plays no role in explaining it.

How plausible is the developmental account in its own right? To say that the child loves its parents as specimens of rational humanity does not sound like an answer to the question 'how is the capacity for rational evaluation acquired?'. Suppose that to love another, or at least to love another rational being, is to exercise one's evaluative faculty. Then the acquisition question demands an account of how we acquire the capacity to love, but Velleman's answer to the question focuses on children at a point at which they already have it.

But perhaps there is more to Velleman's answer. It's noteworthy how prominent a part is played at this stage of Velleman's account of moral development by the nature of the parent's love for the child ('wise good will' etc.: roughly, selfless care), as compared to the part played by the child's love for the parent, i.e. by the highly distinctive account

²⁶ *Ibid.*, p. 152.

²⁷ All quotations from 'A Rational Superego', p. 154.

²⁸ *Ibid.*, p. 155.

²⁹ All quotations from *ibid.*, p. 154.

³⁰ *Ibid.*

³¹ *Ibid.*, p. 155.

according to which love is a response to the very same feature that elicits respect. Given that Velleman spends so much time developing the distinctive account, the appeal to the parent's 'wise good will', neat as it is, comes as a surprise.

It's also, I think, problematic. First, the notion of love deployed in modelling the parents' love for the child is a *different* notion from the one marked out by the distinctive account. It has to be, because the distinctive account applies only to cases where the object of love is an instance of rational nature and, plausibly, children aren't instances of that (yet). Moreover, the appeal to the character of the parent's love for the child may seem to work so well that it's not clear whether we also *need* the distinctive account to explain the child's internalization of the Kantian ego-ideal. Mightn't the child internalize the Categorical Imperative *only* because the parents exemplify it in their conduct towards the child – not because love is (supposedly) a response to the exemplification of rational nature, and the child has *that* attitude to its parents?

Velleman might respond that his account of parents' love for children is intended to apply only to older children, who *are* instances of rational nature. So no new conception of love is involved after all. But parents do love young children as well as older ones, and we had better not overlook this if we are to get a plausible account of the involvement of love in the development of the internal mechanisms of self-regulation.

It could still be that the child's love for the parents is as important in the developmental story as the parents' love for the child: the parents exemplify the capacity to take persons as ends, and thus the Categorical Imperative, insofar as they manifest love for the child, and this ideal having been made available to the child, the child internalizes it not for any old reason but precisely *because* he loves the parents. But if so, it's striking that, where young children are concerned – that is, where a developmental story is most needed – Velleman's distinctive account of love fits loving relations between parents and children in *neither* direction. It doesn't fit the parents' love for the child because young children are non-rational. But it doesn't fit the child's love for the parents either. For, as Velleman says, 'respect for parental discipline as embodying the Categorical Imperative is a sophisticated attainment, which cannot be expected of a younger child'.³² Since love and reverence, on the distinctive account, are responses to the same property, it follows that incapacity for reverence is sufficient for incapacity for love.

How does Velleman fill the gap? Internalization is explained, even in younger children, by the child's love and admiration for the parents. But like love and admiration of more mature kinds, these primitive attitudes don't explain the ego's capacity for evaluation but constitute an exercise of it, and so presuppose it. How, then, does this immature capacity for evaluation get there? And what account is to be given of immature love and admiration, since they surely cannot be responses to the value parents possess as instances of rational nature? If there's going to *be* a developmental story about the more mature capacities for evaluation, the fact that we have the immature capacity must presumably play a large part in it. Velleman says that the primitive form, and the kind of idealized internal others that go with it, are a 'prototype'³³ of the mature form of

³² 'A Rational Superego', p. 155.

³³ *Ibid.*, p. 155.

admiration and the Kantian ego-ideal. Maybe, but what's needed is a realistic picture of the incremental steps that connect the two.

A final comment about Velleman's developmental account connects with the fact that not all parents 'care for [their children] with a wise good will'.³⁴ If what the children internalize is explained by the children's love for the parents *and* love is what Velleman says it is – a response to rational humanity etc. – then we face two equally unattractive alternatives. On one alternative, bad parents can't be loved by their children. But that's not true. On the other, the loving child's gaze passes through the actuality of his parents' foul temper, neglect, excessive entanglement etc. and lights upon a capacity - rational autonomy - that they share with all other mentally non-defective mature human beings (if that is indeed how this capacity is distributed). But if *this* is the feature that underlies the child's love for his parents and what he loves in them is what he internalizes as his ego-ideal, we'd expect children of bad parents to have the same ego-ideals as the children of good ones. But do they? Some Kantians might say every human being endowed with the capacity for rational autonomy has the same 'true and better self'. In my view, however, to have a capacity is one thing and to hold the exercise of that capacity as an ideal, even one more honoured in the breach, is another: children of violent parents, for example, may idealize violence, but this would surely fail to get through the filter of the Categorical Imperative. Here Velleman's unannounced departure from the distinctive account of love is – whether or not Velleman sees it like this – a welcome way out of the difficulty. What the children internalize depends on what's on offer from the parents. Of course children who aren't wisely cared for won't internalize the same kind of thing as those who are, but then we don't want an account that implies that *all* (non-defective etc.) human beings grow up with rational humanity as their ego-ideal, but at most one which implies that the good (if imperfect) ones do.

For all these disagreements with Velleman, he is surely right to emphasize, as few others in philosophy have, that the love of another is a path, if not the only one, to finding and living up to one's 'true and better self'; that others are more likely to appear as authorities to us, and so to promote mechanisms of self-regulation that don't just feel like bullies squatting inside our heads, if they love us and we love them; and – a point of which the two previous ones may be thought of as illustrations - that there's a deep connection between giving and receiving love and the development of our powers of rational agency. If Velleman's moral psychology doesn't succeed in explaining these observations, we are nonetheless in his debt for the depth and sophistication of the attempt.

³⁴ 'A Rational Superego', p. 154.