

Rational Partiality and Objective Value

§1 Introduction. As truth should guide belief and credence, so value should guide desire and preference. More specifically, *objective, all-things-considered value* should guide desire and preference.¹ On supposing that one possibility is objectively, all-things-considered better than another, it is irrational to fail to prefer that possibility to the worse one. Call this principle the *Preferential Principal Principle* (or PPP).² I take PPP to be an important principle of practical reason, and at any rate intuitively plausible and interesting enough for it to be worth exploring its implications.

A natural way of spelling out PPP is as follows. Both objective, all things considered betterness, \succeq° , and a rational agent's preference relation, \succeq , are orderings of possible worlds.³ Analogous to conditional credence, agents also have conditional preferences, \succeq_ϕ , functions from propositions to preference orderings—what they prefer on supposing some proposition ϕ is true. PPP, on this understanding, says that for all rational conditional preference functions,

$$w \succeq_G w',$$

where w and w' are worlds and G is the proposition $w \succeq^\circ w'$. In other words, a rational agent, on supposing that world w is at least as good as world w' , weakly prefers w to w' . Call this formulation of PPP the *worlds formulation*.

One familiar objection to this and similar principles is that it is rational to have *partial* preferences. You can rationally prefer to save your own child to saving two others, despite this not being the objectively better outcome. One way of replying to this kind of objection involves gerrymandering the betterness ordering of worlds to match what we take to be rational preferences. In this paper, I formulate the partiality objection in a way that is unsusceptible to this kind of axiological gerrymandering, then offer what I take to be a more attractive reply in its stead: we should replace the worlds formulation of the PPP with a *centered worlds formulation*, interpreted in a particular way.

§2 The Problem of Intra-World Partiality. A defender of PPP might reply to cases of apparently rational partiality by observing that there are differences beyond the obvious ones in the outcomes that might be reflected by objective betterness. Besides the number of children saved, the relevant worlds differ in, among other things, whether a parent

¹A kind of value is objective iff what has how much of it does not vary from perspective to perspective. And a kind of value is all-things-considered iff there is not some distinct covering value (in the sense of Chang (2004)) that subsumes it and, furthermore, it subsumes any distinct value. This is the kind of value that, if you're a certain kind of consequentialist, you think determines what one all-things-considered ought to do.

²C.f. the Principal Principle linking credence and objective chance from Lewis (1980). For reasons I won't go into here, we'd likely want to formulate the principle in a more general way and derive this more specific principle from it. But the extra generality involves unnecessary complication as far as our current topic is concerned.

³For greater generality, but more complication, we could make these orderings of sets of worlds.

saves her own child. Perhaps there is something impersonally valuable about acting or preferring in ways that look partial. But if so, the world where you save your own child may be impartially better than the one where you save the two. The opponent can then construct more complicated cases where sacrificing one's own child prevents two others from sacrificing theirs, but it seems the PPP defender can respond with more gerrymandering.

There are cases, though, where gerrymandering the betterness ordering must fail. Consider

Preferring to be Myself: My life has gone pretty well so far. But many people have led truly terrible lives. There have been plenty of morally vicious tyrants who have deliberately caused much needless suffering to innocent people, did nothing to redeem themselves, and were moreover miserable. Pick one of these tyrants, let's call him Terry. I strongly prefer living my life as it actually is (and will be) to living Terry's life as it actually was.

This preference is rational, but this cannot reflect any difference in the worlds of the outcome—after all, the worlds are the very same! Terry's life and mine are both in the actual world.⁴ The gerrymandering reply can find nothing to grip onto here.

§3 Centered Objective Value. Preferences like this, as well as other *de se* attitudes (like believing that I am [author's name] and not Terry) have long been known to cause problems for theories that represent contents of attitudes as sets of worlds. One standard solution⁵ is to move from worlds to *centered worlds*—world, individual-in-that-world pairs. Taking on this approach, we represent my preference as $\langle @, \text{Terry} \rangle < \langle @, \text{the-Author} \rangle$. The central proposal of the paper is to carry over this approach to objective value and provide an interpretation that allows rational partiality to be compatible with PPP. My preference is rational because I recognize that $\langle @, \text{Terry} \rangle <^\circ \langle @, \text{the-Author} \rangle$.

The move to centered-worlds (or something similar) for value is not new. Agent-relative forms of consequentialism, for example, rely on something along these lines. But this agent-relative value has generally been understood in a way that makes the value non-objective in the relevant sense.⁶ On this interpretation of centered-world value, the *perspectival* interpretation, what's in question is the value of *the world from the perspective of the individual*. It might be true from my perspective that the world where I save my child is better than the one where I save yours, but this is false from your perspective. This interpretation is thus not objective in the sense required to uphold PPP.

⁴It is arguable that the world where I am Terry, though, is a different world. But even if we accept this, to account for the rationality of this preference the impartialist must reject that value supervenes on worlds' qualitative features. In the paper I argue this is not a position an impartialist can tenably hold.

⁵Advocated by Lewis (1979).

⁶An early statement of the view is in Sen (1982). For more recent discussion, see Smith (2003), Portmore (2005), Schroeder (2007), Brown (2011, pp. 761–763), Suikkanen (2009), and Cullity (2015).

I propose a new interpretation of centered-world value, the *individual* interpretation, which is compatible with PPP. On this view, the bearers of value are not worlds (or worlds-from-a-perspective), but individuals embedded in particular worlds. What it means, then, for one individual-world pair to appear higher up in the betterness ordering than another individual-world pair is for it to be better to be the former individual in the former world than it is to be the latter individual in the latter world. Being me in the actual world is better than being Terry in the actual world, and this is true from any point of view. Since I recognize this, preferring to be me to being Terry is rational, according to PPP. Hence we see that PPP on the centered-worlds interpretation is, given the individual interpretation of centered value, compatible with rational partiality.

References

- Brown, Campbell (2011). "Consequentialize This". In: *Ethics* 121, pp. 749–771.
- Chang, Ruth (2004). "Putting Together Morality and Well-Being". In: *Practical Conflicts: New Philosophical Essays*. Ed. by Peter Baumann and Monika Betzler. Cambridge University Press, pp. 118–158.
- Cullity, Garrett (2015). "Neutral and Relative Value". In: *The Oxford Handbook of Value Theory*. Ed. by Iwao Hirose and Jonas Olson. Oxford University Press, pp. 96–116.
- Lewis, David (1979). "Attitudes *De Dicto* and *De Se*". In: *The Philosophical Review* 88, pp. 216–212. Reprinted in Lewis (1983, pp. 133–159).
- (1980). "A subjectivist's guide to objective chance". In: *Studies in inductive logic and probability*. Ed. by Richard Jeffrey. Vol. II. University of California Press, pp. 224–247. Reprinted with postscripts in Lewis (1986, pp. 82–132).
- (1983). *Philosophical Papers*. Vol. I. Oxford: Oxford University Press.
- (1986). *Philosophical Papers*. Vol. II. Oxford University Press.
- Portmore, Douglas W. (2005). "Combining Teleological Ethics with Evaluator Relativism: A Promising Result". In: *Pacific Philosophical Quarterly* 86, pp. 95–113.
- Schroeder, Mark (2007). "Teleology, Agent-Relative Value, and 'Good'". In: *Ethics* 117.2, pp. 265–295.
- Sen, Amartya (1982). "Rights and Agency". In: *Philosophy & Public Affairs* 11.1, pp. 3–39.
- Smith, Michael (2003). "Neutral and Relative Value after Moore". In: *Ethics* 113.3, pp. 576–598.
- Suikkanen, Jusi (2009). "Consequentialism, Constraints and the Good-Relative-to: A Reply to Mark Schroeder". In: *Journal of Ethics and Social Philosophy* 3, pp. 1–9.