

Value Receptacles

(Forthcoming in *Noûs*)

Richard Yetter Chappell*

University of Pennsylvania / Bowling Green State University

Abstract

Utilitarianism is often rejected on the grounds that it fails to respect the separateness of persons, instead treating people as mere “receptacles of value”. I develop several different versions of this objection, and argue that, despite their *prima facie* plausibility, they are all mistaken. Although there are crude forms of utilitarianism that run afoul of these objections, I advance a new form of the view—‘token-pluralistic utilitarianism’—that does not.

Introduction

Consider the traditional utilitarian view that an act is right if and only if it produces at least as much value as any other that the agent could perform at that time, where the amount of value present in an outcome is fixed by the amount of happiness in that outcome.

It’s often thought that utilitarianism treats persons as mere ‘receptacles’ or repositories for whatever happens to be of value. On this picture, a genuinely utilitarian agent would fail to recognize other individuals as valuable

*Thanks to Nate Gadd, Daniel Greco, Elizabeth Harman, Hrishikesh Joshi, Eden Lin, Brennan McDavid, Tim Mulgan, Philip Pettit, Peter Singer, Michael Smith, Pablo Stafforini, and Helen Yetter-Chappell, for helpful comments and discussion.

ends in themselves; instead, persons are seen as *mere means* or instruments to the end of realizing happiness in the world.

This objection has been formulated in several different ways. Rawls (1999, 24) famously objected that “Utilitarianism does not take seriously the distinction between persons.” Singer (1993, 121) writes, “It is as if sentient beings are receptacles of something valuable and it does not matter if a receptacle gets broken, so long as there is another receptacle to which the contents can be transferred without any getting spilt.” The common thought here is that there’s an important sense in which utilitarianism fails to treat us *as* individuals. It takes our interests into account, but not in a way that appreciates the normative *distinctness* of my interests and yours. We remain, in an important sense, entirely replaceable.

These formulations are evocative, but imprecise. In this paper, I will develop and assess several anti-utilitarian objections along these lines. I begin, in § 1, by considering two versions of the ‘replaceability’ worry that are merely axiological in nature, and easily dealt with.¹ The first concerns death and replacement; the second, whether we can precisely compare the welfare of distinct persons. § 2 addresses the objection that utilitarians care about utility *rather than* caring about people’s interests as such. Finally, § 3 develops and rebuts what I take to be the strongest form of the “value receptacle” objection: that utilitarians treat individuals’ interests as fungible

¹ We will see that there is a subtle sense in which even the later objections qualify as ‘axiological’, but they will concern the general *structure* of our value theory rather than its particular *contents*.

means to the aggregate good.

In each case, I argue that although there are versions of utilitarianism that would fall victim to the objection in question, there is also a natural and compelling version of utilitarianism that escapes the objection. (Curiously, while one might have expected that the *maximizing* component of the view was responsible for the problems here, we'll see that the issue rather turns on the utilitarian's theory of value.) The availability of this response is, if correct, an important result. The idea that utilitarianism neglects the separateness of persons is a common reason for rejecting the theory. If, as I conclude, there is no non-trivial sense² in which this objection applies to the best form of utilitarianism, then we should not reject the theory on these grounds. Nonetheless, it is a useful objection to explore, precisely for its ability to steer us towards a new—more appealing—interpretation of utilitarianism.

² A 'trivial' version of the objection would, for example, be to *presuppose* (rather than argue) that the separateness of persons entails fundamental individual rights not to be sacrificed for the greater good. It is, of course, trivial that such fundamental rights are incompatible with utilitarianism, but this is not something that any utilitarian is likely to be concerned about. The 'non-trivial' objections I go on to explore differ in that they understand the separateness of persons as an independent idea with substantive content, rather than as mere shorthand for an extensional objection to utilitarianism's verdicts.

1 Axiological Refinements

1.1 *Death and Replacement*

One respect in which classical utilitarianism may seem to neglect the separateness of persons is that it attributes no significance to the ways in which experiences are packaged together into distinct lives. In particular, the badness of death is seen entirely in terms of its causing there to be fewer good experiences in future. We may worry that this does not do justice to the significance of death for the individual whose life is cut short. The premature death of an individual is bad in a way that goes beyond the mere failure to create future goods. Most of us would not think it a good thing (all else equal) for someone to be struck down in the prime of life and replaced with a marginally happier substitute. Death is not equivalent, as this view would have it, to the failure to create life.

But we can accommodate this intuition without abandoning utilitarianism. We merely need to refine our theory of value so as to properly capture the disvalue of death. Here's one possibility: Besides preventing the creation of future goods, death is also positively disvaluable insofar as it involves the interruption and thwarting of important life plans, projects, and goals.³ If such thwarting has sufficient disvalue, it could well outweigh the slight in-

³ Importantly, this account of the positive badness of death avoids the opposite mistake of attributing *constant* and *unconditional* disvalue to death. There may be circumstances in which death is an unmitigated blessing, after all. Instead whether—and to what extent—death constitutes a positive harm to a person will depend on the situation, i.e. what important life projects it cuts short.

crease in hedonic value obtained in the replacement scenario. Utilitarians are thus fully able to attribute significance to the packaging of experiences into lives, and to acknowledge the positive disvalue of death—they just need the right theory of value.

1.2 *Imprecise Values*

The traditional utilitarian practice of assigning exact numerical values might also seem inconsistent with respecting the separateness of persons, insofar as it makes it too easy to break ties and mandate saving one person over another. Suppose we begin with two people, neither of whom has a more valuable life than the other, and you can save only one. It doesn't seem that mildly "sweetening" one of the options, with a dollar bill or the like, should break the tie or make the choice any easier or less arbitrary.

Utilitarians may accommodate this phenomenon of *resistance to sweetening* by—once again—appropriately complicating their value theory. Rather than holding the two lives to be precisely equal in value, they may be merely *roughly* equal (Parfit 1984, 431), or 'on a par' (Chang 2002), such that sweetening one option does not necessarily make it of greater total value than the other (despite being better than it was prior to sweetening).

While there is some intuitive support for the thought that resistance to sweetening is often appropriate, I don't think that it would be outright *immoral* to insist on precise values.⁴ As we will see, it's a mistake to think

⁴ While perhaps a "moral error" in some abstract sense, it is not *disrespectful* of an-

that treating people's lives as *comparable* in value entails treating them as *fungible* or interchangeable in the way that we treat money, for example, as being. I might be genuinely torn between two distinct but equal intrinsic values, recognizing the separate force of each, even as my decision hangs in the balance such that the slightest inducement to either side would sway my decision. The sensitivity of my decision to further incentives does not in any way imply a failure to appreciate the distinct and irreplaceable conflicting values in play. So the separateness (or non-fungibility) of values cannot be understood merely as a matter of their being not precisely comparable. We need a better account. In the following sections—most notably § 3—I advance a positive account of what it really takes to disrespect a person by treating them as fungible, and how consequentialists (even utilitarians) can avoid this fate.

2 Incidental Interests

It's widely agreed that we have reasons to help other people. But we may ask about the deeper structure of these reasons: *why* do we have this reason? On whose behalf does this reason exert its normative force or make claims on us? The commonsense answer is that these normative reasons speak on the behalf of the individuals who need our help. It is *for their sake* that we have reason to relieve their suffering. This much seems clear.

Yet utilitarians might be thought to deny this datum. As [Singer \(1993\)](#),

other's individuality as a person in the sense discussed in later sections.

121) puts it, “The total version of utilitarianism regards sentient beings as valuable only in so far as they make possible the existence of intrinsically valuable experiences like pleasure.” There is no mention here of the interests of the beings experiencing these pleasures. If the utilitarian’s theory simply tells her to maximize net happiness, it may seem natural to reconstruct the fitting utilitarian’s thought-process as follows: *Bob is in agony. My goal is to maximize utility, i.e., the balance of pleasure over pain. There is some agony (namely, Bob’s) that I am in a position to relieve. Doing so would serve my goal. So I will act to relieve Bob’s suffering.* But now note that the interests of *Bob himself* seem to have dropped out of the picture for our imagined utilitarian agent. She is merely concerned to minimize pain and suffering. The fact that doing so is *good for Bob* (or anyone else) is not a relevant consideration to her way of thinking, or so we might imagine. Helping people is incidental, a mere side-effect to her real goal of patterning the universe with a particular class of experiences. Call this view *Utility Fundamentalism*.

By taking the value of pleasure (and disvalue of pain) as fundamental, and not to be explained in terms of their value *for* individuals, Utility Fundamentalism seems objectionably fetishistic. It treats individuals as intrinsically valueless ‘receptacles’, of moral interest only insofar as they provide a space or habitat for what (supposedly) really matters: the brute promotion of pleasure over pain. This moral perspective strikes us, I think rightly, as perverse.

If this is how we are to understand the ‘value receptacle’ objection, then utilitarians (and consequentialists more broadly) may escape it simply by rejecting Utility Fundamentalism. After all, there is a very natural alternative account, according to which pleasure (say) is good precisely *because* it is good *for* the individual who experiences it, and suffering is bad because it is bad for the suffering individual (Wilson 2006). On this view—call it *Welfarism*—the interests of individuals play an essential explanatory role in our value theory. When the welfarist utilitarian relieves Bob’s suffering, the fact that this benefits Bob is not merely incidental to her reason for acting. It is, on the contrary, the source or ground of her reason. She has reason to relieve suffering precisely because this is good for someone.

We may demonstrate the difference between these two views by way of a fanciful counterfactual: If the welfarist utilitarian became convinced that some pain was, for some reason, intrinsically good for Bob, she would no longer take herself to have non-instrumental reason to rid Bob of it.⁵ The utility fundamentalist, by contrast, has a fixed goal that makes no mention of the interests of individuals *as such*. She cares about experiences, not *experiencers*. So even if she too believed pain to be good for Bob rather than bad for him, this would be of no intrinsic interest to her: she just wants to minimize pain, no matter whether this helps or harms the individuals

⁵ Note that we needn’t think it genuinely possible for pain to be intrinsically good for a person. The possibility I’m imagining here is merely that a well-meaning agent *believes* this to be so. Any readers who remain suspicious of the example will hopefully still be able to grasp the difference in normative structure between valuing pleasure brutally vs. valuing pleasure *because* it’s good for the person who experiences it.

experiencing the pain in question.

We thus see that only the utility fundamentalist is liable to the ‘value receptacle’ objection, understood as the failure to recognize that happiness (or whatever) is good just because it’s good *for* individuals. This fetishistic perspective is by no means endemic to utilitarianism. Indeed, it is entirely natural for utilitarians to instead take the welfarist route of specifying that happiness is good precisely *because* it’s good for the individual who experiences it. Our current interpretation of the value receptacle objection is then simply inapplicable to this welfarist form of utilitarianism.

3 Are Persons Fungible?

Even given an appropriately welfare-based explanation of why happiness matters, there remains a second interpretation of the ‘value receptacle’ objection that might be leveled against the utilitarian. The remaining objection is that utilitarians treat particular individuals not as ends in themselves, but merely as fungible or replaceable means to the end of promoting *aggregate* welfare.

Recall [Singer \(1993, 121\)](#)’s evocative explanation of the ‘value receptacle’ metaphor: “It is as if sentient beings are receptacles of something valuable and it does not matter if a receptacle gets broken, so long as there is another receptacle to which the contents can be transferred without any getting spilt.” The worry here is that there’s an important sense in which utilitarianism fails to treat us *as* individuals. It takes our interests into account, perhaps even *as*

interests, but not in a way that appreciates the normative *distinctness* of my interests and yours. We are all melded together, into a kind of unstructured, undifferentiated welfare soup.

To make the problem vivid, imagine that Connie the Consequentialist is faced with two poison victims, and just enough anti-venom to save one of them. And suppose that, faced with their pleading faces, but realizing that it makes no difference to the total utility which person she saves, Connie finds herself feeling completely *indifferent* about her choice. It's as if she had to choose between a \$20 bill or two tens. Now, it seems that Connie making a deep moral mistake here. She's treating the two people's interests as completely fungible, like money, and neglecting what we might call the "separateness of persons"—the fact that each person is of distinct intrinsic importance, in their own right, and not merely a fungible means to aggregate welfare.

As this case illustrates, we often imagine the utilitarian agent as having but a single ultimate desire: to maximize aggregate welfare. They thus see different individuals as interchangeable. It makes no difference, to such an agent, which of several people is helped (or indeed whether one person is helped a lot or several people each helped a little), so long as the impact on aggregate welfare would be the same in either case.

To bring out why this is so objectionable, note that fungibility is, in general, the mark of the instrumental. Money is fungible precisely because we do not value the possession of *particular* bills: replacing two tens with a

twenty would serve my ends just as well. For another example, if my sole ultimate desire is to slake my thirst, then I will be indifferent between two equally effective means to satisfying this goal. If someone switches my glass of water for another that's qualitatively identical, this is not a change that's normatively significant to me. I do not desire *that* glass in particular, so it may just as well be replaced by any other that would do the job. On the other hand, if I *had* (bizarrely) desired the original glass in its particularity, then the substitution would be of significance to me: it would thwart one of my non-instrumental desires.

This connection between fungibility and merely instrumental valuation explains why the above objection to utilitarianism seems so forceful. It seems perverse to treat individuals as replaceable or fungible, because such treatment constitutes a failure to intrinsically value individuals in their particularity. The correct moral theory, we feel, must attribute intrinsic value to particular individuals and not just to the general welfare (cf. [Cohen 2011](#)).

How is a theory to satisfy this requirement? We can clarify the matter by reference to what kind of psychology would seem to embody or exemplify an accurate moral perspective (we may call this the 'fitting' psychology for an agent to possess). We have seen that it's morally perverse for an agent to be *indifferent* between options that equally benefit distinct people, for that is to disrespect the individuals by treating them as fungible means to the aggregate welfare. But of course we do not want to favour either person over the other, since such bias would constitute disrespect for the person whose

equal benefit we counted for less. Instead, I propose, the fitting response to a tradeoff between two distinct but equally weighty values is to feel *ambivalent* about the choice. There are distinct reasons pulling you in either direction, corresponding to the distinct values served by either choice. But these reasons are equally weighty, so the agent is *torn* rather than pulled without resistance towards one choice over the other.

This is a distinction we should want our theories to be able to make. Whatever substantive disputes we may have about what is of value, we should all acknowledge the formal difference between (i) a pair of options serving distinct but equally weighty final values, and (ii) a pair of options that serve literally one and the same final value. For example, assuming that token artworks have final value, a choice between saving a great painting or an equally great sculpture is importantly different from a choice between saving the same painting in either of two different (but equally effective) ways. In the latter case, the two options are seen to serve the same token value in virtue of saving the same token artwork. Other cases of this may be more subtle, as even two distinct concrete objects may serve as vessels for one and the same token value. An intuitive example of this is pleasure: I'm completely indifferent between the prospects of a massage for my left foot or my right, assuming that either would be similarly pleasant.⁶ I take this to suggest that left-foot-pleasure and right-foot-pleasure are not distinct final values, the way that the painting and the sculpture (or my welfare and your welfare)

⁶ Thanks to an anonymous referee for prompting me to discuss this case.

are. Instead, it seems, I ultimately value pleasures of a certain qualitative kind *in the aggregate*, and particular instances of such pleasures are thus, in an important sense, of merely ‘instrumental’ value to me. Of course this is not to say that they are causally instrumental to some downstream effect. We may instead call it *constitutive instrumentality*, as each token of pleasure is a constitutive, rather than causal, means to the end of aggregate pleasure.

With this understanding in hand, we may now characterize the fungibility objection as alleging that the utilitarian perspective must likewise treat *individual persons* as constitutive means to the aggregate welfare, rather than as distinct ends in themselves. Given that individual persons have final value, such instrumental treatment constitutes a distinctive kind of *disrespect* or failure to respond appropriately to the value that persons have in themselves.

The reader should now have an intuitive grasp of the distinction between (equally-weighty) distinct final values and (equally effective) mere means to a single final value. I’ve suggested that one way this distinction might play out is that in the second case the two options are perfect substitutes, and hence the fitting attitude for an agent to take towards them is indifference. In the former case, by contrast, the two options are not *substitutes*; they serve different ends, albeit equally worthy ones. This naturally suggests that the fitting attitude to take is ambivalence, rather than indifference.

Another way to support this conclusion is via the idea that it’s fitting to intrinsically desire *each* intrinsic good, with strength proportional to the

magnitude of the object's value. If, and only if, a pair of options serve distinct intrinsic values, will the two options differentially satisfy the intrinsic desires of the morally fitting agent (and hence strike her as significantly distinct). Insofar as the agent has conflicting desires, we can say that she manifests ambivalence rather than indifference over the options.

We are now in a position to evaluate the objection that utilitarianism treats people, and their interests, as fungible. This is, as we have seen, equivalent to interpreting utilitarianism as the view that only one token thing, namely aggregate welfare, has intrinsic value. Call this view *token-monistic utilitarianism*. This view really does neglect the separateness of persons, for it attributes intrinsic value merely to the whole, and not to each of us in our particularity. As a consequence, the token-monistic utilitarian mindset involves but a single desire—to maximize welfare—and treats our individual interests and concerns as mere (constitutive) means to the satisfaction of this more global goal. This is, I agree, morally perverse.

But there is no reason why utilitarianism must take this monistic form. There is a very natural alternative view, call it *token-pluralistic utilitarianism*, on which *each* particular person's interests are (separately) accorded final value.⁷ There is not just one thing, the global happiness, that is good. Instead, there is my happiness, your happiness, Bob's, and Sally's, which

⁷ The view may still be monistic in the sense that there's just one *type* of thing that's good (cf. Hurka 1996). But the crucial point for present purposes is that there are a plurality of *token* final values. The separateness of persons merely requires that we each be valued separately. There's nothing obviously objectionable about it turning out that we are valuable in the same kind of way.

are all equally weighty but nonetheless distinct intrinsic goods. What this means is that the morally fitting agent should have a corresponding plurality of non-instrumental desires: for my welfare, yours, Bob's, and Sally's. Tradeoffs between us may be made, but they are acknowledged as genuine tradeoffs: though a benefit to one may outweigh a smaller harm to another, this does not *cancel* it. The harm remains regrettable, for that person's sake, even if we ultimately have most reason to accept it for the sake of more greatly benefitting another.⁸

⁸ This solution also addresses Otsuka (2012)'s proposal that a theory respects the separateness of persons only when it is sensitive to "competing claims" and so treats "non-identity" cases differently.

To illustrate, suppose we must choose between benefitting the already-existing Jane or a future person. Otsuka claims that it makes a moral difference whether the latter choice would be *identity-affecting*. If it would cause a *different*, happier person to come into existence, distinct from the person (call him 'Jim') who would otherwise come into existence, then there is no "competing claims" justification for benefitting the future person. If you instead help Jane in this case, Jim has no grounds for complaint, since there is nothing else you could have done to help *him*—bestowing the future benefit would (ex hypothesi) have instead brought someone *else* into existence. So, it is easier to justify helping Jane in this identity-affecting case than it would be in the constant-identity case where Jim also has a "competing claim" to be benefitted.

I'm sympathetic to the idea that there is a moral difference between such cases of constant-identity and non-identity welfare tradeoffs. So I'm willing to grant that our moral theories should be sensitive to this difference. But it's a mistake to think that the only form such sensitivity could take is through differences in its implications for right *action*. According to token-pluralistic utilitarianism, there may be a (non-quantitative) difference in the values at stake in the two choice situations. In the constant-identity case, there is a particular token value—future Jim's welfare—that is not being served as well as it might have been, when you instead benefit Jane. In psychological terms: there is something (pro tanto) *regrettable* about this decision, for Jim's sake, even though bestowing a greater benefit on Jane would in fact be the right choice to make. In the non-identity case, by contrast, there is no person-affecting opportunity cost to helping Jane. No existing person could have been helped by the alternative choice. So, there is not the same grounds for pro tanto regret.

(It's an interesting question, beyond the scope of the paper, whether we should prefer to benefit actual people rather than bringing into existence new, happier people. I've suggested that this isn't necessary in order to respect the separateness of persons, but it

Contrast this with the case of money: If you have to invest \$5 to earn \$10, there is nothing to regret. The \$5 is a “cost” merely in the sense that it would have been *even better* if you could have attained the \$10 payoff without having to pay the \$5. But given that this is not an option, there is nothing regrettable about the deal *as a whole*, the way that there is something regrettable about benefitting one person greatly at lesser cost to another. We can explain the difference, in the cash case, as a matter of both sums of money being mere components or constituents of the single token value, or desirable end, of aggregate wealth. This is very different from how the token-pluralistic utilitarian conceives of welfare tradeoffs between distinct persons.

We thus find that utilitarianism is well able to reflect the normative separateness of persons, and to avoid treating people as fungible, replaceable receptacles of value. This is, if correct, an important result: It’s commonly thought that the utilitarian’s willingness to weigh harms to one person against benefits to another essentially involves treating the one as a “mere means”. But my above analysis suggests that this traditional thought is simply confused. One may have thoroughly *non-instrumental* desires for each of two distinct intrinsic goods, and make reluctant tradeoffs between them in a way that is importantly different in kind from the tradeoffs one makes with fungible goods like money. The mere willingness to balance conflicting values is not itself constitutive of instrumental or fungible treatment. Critics may

might be plausible on independent grounds, in which case Welfarist Consequentialists may wish to modify their value theory accordingly. For a good discussion of such non-identity tradeoffs, see [Coons \(2013\)](#).)

still insist that utilitarianism is just *extensionally incorrect* in its prescriptions for morally right action, but those wanting to make stronger claims about ‘value receptacles’ need to back up their claims with a rival account of instrumental valuation—as such rhetoric is seen to be baseless if the present account of instrumental valuation is correct.

An interesting implication of my account is that we may find that we actually treat our interests-at-a-time as fungible.⁹ While we might initially have assumed that our momentary interests have final value, we may find on reflection that we consider our interests across time, unlike interests across people, to be properly fungible. As in the case of fungible pleasures, this view can easily be incorporated into my framework by positing that individuals’ interests-at-times are mere constitutive means to the final good of their timeless welfare. Alternatively, you might opt for the view that it’s fitting to consider tradeoffs between timeslices to be just as emotionally fraught as tradeoffs between persons, and so assign final value to each momentary self individually. For purposes of this paper, I can remain neutral on this question of whether to attribute final value to momentary welfare, or only to timeless welfare.

3.1 *Objections*

I have argued that a utilitarian agent could respect the distinctions between persons by separately desiring the good of *each* person’s welfare, rather than

⁹ Thanks to an anonymous referee for bringing this to my attention.

having a single, totalizing desire for the aggregate good. But a difficulty arises when we consider goods that the agent is unaware of. Consider some particular unknown person, Harry. Our utilitarian cannot have a particularized desire for Harry's welfare, since she cannot even refer to Harry in particular. But her values must extend to others somehow: It's not as though she'd accept an offer to improve the welfare of her neighbour Jones at greater cost to some unknown other. So it seems that we need something like a generic desire for aggregate welfare to step in and fill the gap. (To avoid double-counting, we'd probably need to exclude Jones—and any others for whom the agent already has a particularized concern—from the remaining aggregate.)

Is this a problem? Perhaps not. It doesn't seem so objectionable to treat people you've never even heard of as faceless members of the aggregate. How could they be *other* than faceless and generic to you? Moreover, the agent's attitude here is not merely instrumental. It's not as though our utilitarian thinks that unknown people fundamentally matter only in respect of their being members of the unknown aggregate. Rather, her concern for unknown people's aggregate welfare is a stop-gap measure that reflects, in the only way possible, her appreciation of the fact that each of those individual unknown persons fundamentally matters in their own right. She knows that, if she knew more, she would form particularized desires for the welfare of each; but in the absence of the requisite identifying information, the best she can do to respect these unknown values is to fall back on the generic desire for aggregate welfare, as a kind of placeholder.

So far, so good. But what about *merely possible* future persons? (Compare Parfit (1984)'s 'Non-Identity Problem'.) Here the placeholder strategy seems dubious. Before, we were holding the place for the particularized desires we would have if fully informed—and it seems reasonable for an agent to differentially ascribe normative authority to her fully-informed desires. But in case of merely possible persons, the barrier to particularized reference is metaphysical, not merely epistemic: There *is* no such particular person to refer to. The most we can appeal to is the *counterfactual* desire that we (ideally) would have had if someone else had existed. We would have formed a particular desire for that someone's welfare. But so what? As things stand, there is no such person, and hence no valuable entity for us to respect as best we can. We cannot have a 'second-personal' reason (Darwall 2006), grounded in the normative authority of the non-existent individual himself, to take his possible welfare into account. Our concern must instead be, in a sense, 'impersonal'.

Even so, this needn't return us to any single, totalizing desire that the world be thus-and-so. We may instead have distinct desires for each possible generic good.¹⁰ We still need to distinguish between indifferent and ambivalent pairs of prospects, after all. For example, I may desire both that Bill has a happy child rather than none at all, and that Max has a happy child rather than none at all. Perhaps I cannot coherently desire these things *for* the sakes of the respective children (especially if they never actually exist),

¹⁰ This also applies to the previous case of actual-but-unknown individuals.

but I can desire them—for the sake of the world, perhaps. And in so doing, I recognize that the prospective persons are not fungible, in the following sense: Despite being of equal value, there is a morally relevant difference between a world where only Bill has a child, and a world where only Max has a child. The comparison calls for ambivalence, rather than indifference, since they serve distinct (though equally weighty) ends or ideal desires. If Bill's child would have a better life, then I could prefer that she be the one to come into existence, even while I regret the absence of Max's possible child, whose life would have been (distinctly) intrinsically valuable in its own right.

The objector might respond by suggesting that it's only because of the differential impact on the existing individuals Bill and Max that we see a significant difference here. If we imagine some more thoroughly generic question—say, whether the 100th child born in the year 2500 is a boy or a girl—indifference may seem the only appropriate response. I'm not sure about this, as our lack of response may just be due to our contingent failure to really vividly appreciate what a significant intrinsic difference the identity of each individual makes. But even if the critic is right here, it's not clear that this is any objection to consequentialist theories in particular. If it turns out that distant future people cannot *but* be thought of as fungible, in the noted sense, then this limitation will presumably apply to all moral theories.

So I think the objection ultimately fails. Even in the toughest case—that of merely possible persons, who cannot be the ultimate ground of our concern for their welfare—utilitarians can plausibly still desire each good separately,

and hence refrain from treating people as fungible. And even if it turns out that I'm mistaken about this, and in fact merely possible persons *are* fungible, then that is no fault of utilitarianism. It would instead be a constraint that any moral theory must work within. So the remaining challenge for a theory would just be to ensure that it doesn't inappropriately extend the domain of the fungible to include actually existing persons. Utilitarianism can, as I have shown, meet this challenge.

Conclusion

This paper has developed various interpretations of the “value receptacle” and “separateness of persons” objections, and shown how utilitarians can best respond to each one. We've seen that: (1) Our theory of welfare should ascribe positive disvalue to the premature thwarting of important life projects, or otherwise explain how death is (typically) bad in a way over and above preventing the realization of future goods. (2) The intuition that the value of distinct lives is not precisely comparable can be accommodated by the utilitarian's value theory if desired, but there's nothing deeply objectionable about allowing for precise comparability. (3) Utilitarians should be “welfarists” in the sense that they see happiness (etc.) as good precisely *because* it is good for the people who experience it. And, most importantly: (4) To avoid treating people as replaceable, utilitarians should be *token-pluralists* who see each person's welfare as a distinct final good, rather than

(separateness-violating) *token-monists* who see the aggregate welfare as the only final good. If my analysis of instrumental valuation in § 3 is correct, then the common thought that utilitarian tradeoffs inherently involve treating someone as a “mere means” is simply confused. Willingness to balance conflicting values is not itself constitutive of instrumental or fungible treatment. Whatever *extensional* objections one may have to utilitarianism and its verdicts, it does not have the objectionable *intensional* character that’s often attributed to it.

References

- Chang, Ruth. 2002. "The Possibility of Parity." *Ethics* 112:659–688.
- Cohen, G.A. 2011. "Rescuing Conservatism: A Defense of Existing Value." In R. Jay Wallace, Rahul Kumar, and Samuel Freeman (eds.), *Reasons and Recognition: Essays on the Philosophy of T.M. Scanlon*. Oxford University Press.
- Coons, Christian. 2013. "The Best Expression of Welfarism." In Mark C. Timmons (ed.), *Oxford Studies in Normative Ethics*, volume 2. Oxford University Press.
- Darwall, Stephen. 2006. *The Second-Person Standpoint*. Cambridge, MA: Harvard University Press.
- Hare, Caspar. 2010. "Take the Sugar." *Analysis* 70:237–247.
- Hurka, Thomas. 1996. "Monism, Pluralism, and Rational Regret." *Ethics* 106:555–575.
- Otsuka, Michael. 2012. "Prioritarianism and the Separateness of Persons." *Utilitas* 24:365–380.
- Parfit, Derek. 1984. *Reasons and Persons*. Oxford: Oxford University Press.
- Rawls, John. 1999. *A Theory of Justice*. Belknap Press of Harvard University Press, revised edition.
- Singer, Peter. 1993. *Practical Ethics*. New York: Cambridge University Press, second edition.
- Wilson, Scott. 2006. "Respect for Utilitarianism: A Response to Regan's 'Receptacles of Value' Objection." *Proceedings of the Ohio Philosophical Association* 3.