

SYMPOSIUM

MORAL AND POLITICAL PHILOSOPHY FOR A BROKEN WORLD?



REPLIES TO CRITICS

BY

TIM MULGAN

[THIS PAGE INTENTIONALLY LEFT BLANK]

Replies to Critics

Tim Mulgan

I am grateful to Ben Saunders, Tim Chappell, and Jesse Tomalty for their careful readings of my book, their generous comments about it, and their engagement with its themes and ideas.¹ In my reply, I will address each commentary in turn, also drawing on my own more recent work, where I apply lessons from the broken world to contemporary moral philosophy.

¹ I am also very grateful to Gianfranco Pellegrino for editing this issue, and to Sarah Broadie for organising a workshop on my book at the University of St Andrews in April 2012, where earlier versions of the three commentaries were presented.

I

Ethical Lessons From the Broken World

I begin with some general comments about the moral significance of thinking about a broken future.² In my book, I focus on presenting ethics *within* a broken world—asking how philosophers in that possible future might respond to, and perhaps rethink, our current moral and political theories. However, the broken future also has lessons for contemporary philosophy.

If we encountered an isolated population, perhaps on some distant planet, living without favourable conditions and operating survival lotteries, that would be unsettling enough. But because the broken world may be *our future*, it also has a significant impact on our *current* ethical thinking.

The broken world is a *credible* future. No-one can reasonably be confident that it won't happen. It involves no outlandish claims, scientific impossibilities, or implausible expectations about human behaviour. Climate change—or some other disaster—might produce a broken future. This is not to say, of course, that the broken future will happen. Many other futures are also credible. Some are much better, others are much worse. Our epistemic situation does not allow us to make confident predictions either way. But the broken world is one very real possibility.³ As all three commentators accept the credibility of

² This section draws especially on T. Mulgan, "Ethics for Possible Futures," *Proceedings of the Aristotelian Society*, volume 114 Part 1 (2014), 1-17 and "Utilitarianism for a Broken World," *Utilitas*, forthcoming.

³ For what it is worth, my own (inexpert) reading of the empirical evidence is that we can be confident neither that the future will be as bad as my broken world, nor that it will not be much worse. A particular source of uncertainty is the inability of even the most informed observers to attach meaningful

the broken future, I will not discuss this issue further here. Instead, I shall take the credibility of this future as given, and explore its implications for moral philosophy.

A credible broken future teaches us four main ethical lessons. First, it undermines our tendency to ignore our obligations to distant future people. Philosophers have traditionally marginalised intergenerational issues, because they were confident that they could set the future aside. If we create a stable liberal democratic society in our own generation, then our descendants will inevitably be better-off than us, and therefore their interests do not conflict with ours. In John Rawls's liberal society, for instance, the only intergenerational question is the 'just savings problem': how much better-off we should leave our descendants?⁴ The prospect of a broken future undermines this optimistic presentism. We no longer take it for granted that we *will* leave our descendants better-off, or even that we *can*. And many of our most urgent moral dilemmas involve intergenerational conflict. We now realize that future people might be worse-off *because* we have looked after ourselves.

A second lesson is that the broken future alters the *comparative* plausibility of competing moral and political theories, simply because some theories cope better than others with obligations to future people *in general*. One example, which will become very relevant in my reply to Saunders in section 2, is the perennial debate between contractualist and utilitarian accounts of the foundations of morality. I have argued elsewhere that utilitarianism accommodates the future more easily than

probabilities to outlier possibilities where various feedback loops cause the global climate to spiral out of control once some threshold is passed.

⁴ J. Rawls, *A Theory of Justice*, rev. ed. (Cambridge, MA: The Belknap Press of Harvard University Press, 1999), 251-259.

contractualism, which is its main rival in the contemporary intergenerational literature.⁵

Utilitarianism bases all our obligations on the fact that our actions impact on the well-being of sentient beings. Obligations to future people are thus theoretically on a par with obligations to present people. While utilitarians endlessly debate the precise details of our intergenerational obligations, they have no difficulty making sense of them. By contrast, contractualists have great difficulty accommodating any obligations to future people at all.

Contractualist accounts of intergenerational obligation face two barriers. The first is Derek Parfit's non-identity problem.⁶ Contractualists model morality or justice on a bargain or agreement among rational individuals. But how can we begin to imagine contracts, bargains, or cooperative schemes involving future people whose existence and identity depend upon what we decide? Contractualists as diverse as Immanuel Kant, John Rawls, David Gauthier, and T. M. Scanlon all face serious difficulties here. The second barrier is the lack of reciprocal interaction between present people and distant future people. We can do a great deal to (or for) posterity, but, as the saying goes, what has posterity ever done for us? We cannot bargain, negotiate, or cooperate with those who will live long after us. A *contract* with distant future people seems incoherent.

To crystallise the problem, imagine a 'time bomb' that devastates people in the distant future but has no direct impact

⁵ T. Mulgan, *Future People* (Oxford: Oxford University Press 2006), chapter 1 and 2; "Utilitarianism for a Broken World"; "Contractualism for a Broken World," Paper presented to workshop on contractualism, Universite de Rennes (May 2012). (Available from author).

⁶ D. Parfit, *Reasons and Persons* (Oxford: Oxford University Press, 1984), Part 4.

until then.⁷ (Real life analogues might involve the storage of nuclear waste or the destruction of the global climate.) Suppose that the people who will be affected are so far in the future that no-one alive today cares for them at all. Intuitively, most people believe it would still be very wrong to gratuitously plant a time bomb. But can any social contract deliver this result?

Of course, many contractualists do try to accommodate intergenerational justice. They cite the motivations of present people, exploit contracts between overlapping generations, appoint trustees or ombudsmen for the future, or construct imaginary intergenerational bargaining situations where the parties know neither when nor whether they exist.⁸ But these intergenerational contracts all seem troublingly ad hoc. It is hard to escape the conclusion that, for the consistent social contract theorist, intergenerational justice is (at best) an afterthought—an optional extension of a theory of justice designed for contemporaries. Contractualists cannot accommodate the future as easily or as naturally as utilitarians do.

If conflicts between generations were rare, or if we could be confident that future people would be better off, then this comparative weakness of contractualism might not matter. (After all, no theory is perfect, and utilitarianism certainly has problems of its own.) But, if we face a broken future, then our need for a credible account of our obligations to future people is much greater. This doesn't prove that utilitarianism is superior all-things-considered, but it does significantly enhance its comparative appeal.

⁷ I owe the notion of a time bomb to A. Gosseries, "What do we owe the next generation(s)?" *Loyola of Los Angeles Law Review* 35 (2001), 293-354.

⁸ See, e.g., the essay collected in A. Gosseries and L. Meyer (eds.), *Intergenerational Justice* (Oxford: Oxford University Press, 2009).

This brings us to our third general lesson. Some moral theories handle a *broken* future better than others. Working through the ethical implications of the broken future, both while I was writing my book and in my subsequent work, I have been struck by the number of different ways that philosophers help themselves to optimistic assumptions about the future. Consider four disparate examples: those strands of naturalistic meta-ethics that identify moral facts with the end-points of processes of empirical moral inquiry that may turn out to be inextricably linked to an unsustainable way of life⁹; the many strands of contemporary moral philosophy built on *intuitions* about simple cases—intuitions that are very closely tied to our affluent present¹⁰; libertarians who presuppose that initial acquirers can leave ‘enough and as good’ *for all future people*,¹¹ or Rawlsian liberals who insist that ‘justice’ only applies while favourable conditions persist.¹² The recognition of a credible broken future thus counts against naturalist meta-ethics, intuition-based ethics, libertarianism, and (Rawlsian) liberalism. As philosophy is a comparative business, the broken future thus supports non-naturalism, theoretical ethics, and alternative political philosophies such as utilitarianism.

The fourth impact of the broken future is that it raises troubling practical questions about how we should live *now*. Can we reasonably justify a refusal to adopt the ethical outlook of the broken world *for ourselves*? If future people will be worse-off, partly as a result of our actions, should we reduce our aspirations,

⁹ T. Mulgan, *Purpose in the Universe: the Moral and Metaphysical Case for Ananthropocentric Purposivism* (Oxford: Oxford University Press 2015).

¹⁰ T. Mulgan, “Theory and intuition in a broken world,” in M. Di Paola and G. Pellegrino (ed.), *Canned Heat. Ethics and Politics of Global Climate Change* (Delhi/London: Routledge 2014).

¹¹ T. Mulgan, *Ethics for a Broken World Imagining Philosophy After Catastrophe* (Montreal & Kingston-Itacha: McGill-Queen’s University Press, 2011), 18-68.

¹² *Ibid.*, 160-196.

and bring our notion of what is necessary for a worthwhile human life into line with theirs? Can we insist for ourselves on goods and opportunities that will not, as result of that very insistence, be available to future people?

I am exploring these four lessons in my own current work, and I will draw on them in my replies to Saunders, Chappell, and Tomalty.

II

Ben Saunders on Broken World Democracy

Ben Saunders makes a number of important points about the difficulties that surround any attempt to extend present-day democracy to include future people. He also helpfully separates questions of democracy from questions of justice. Here is how Saunders summarises his own position: “I shall argue that there is nothing *undemocratic* about excluding future generations. Whether we treat them unjustly is a more difficult question, which I do not answer.” (p. 12) In my reply to Saunders, I shall address issues of both intergenerational democracy and intergenerational justice.

Saunders begins with the ‘All Affected Principle’—the common thought that everyone who is affected by a decision has a right to participate in making that decision. This principle is often used to defend extensions of the franchise. (To take one example that Saunder himself discusses: If people in distant lands are affected by pollution originating in our country, then perhaps they have a right to influence our deliberations about industrial and environmental policy.)

As Saunders notes, it is not possible for future people to participate directly in our deliberations. (To avoid unnecessary

complications, I shall assume that ‘future people’ refers to people who exist sufficiently far in the future that they will not overlap with any present person.) When we consider extending the franchise to include additional present people, the barriers are typically practical or political. We could enfranchise people in other countries, if we really wanted to. With future people, by contrast, the barriers to enfranchisement are metaphysical. Only those who exist now can participate in current decisions. Future people do not yet exist. Therefore, we cannot directly enfranchise them. (If democratic institutions endure into the future, then perhaps future people will get to vote when their time comes. But that is not the same as enfranchising them now. And, of course, if we make the wrong decision now, we may prevent future people from enjoying *any* democratic rights.)

At least some of our current decisions will inevitably impact on future people. We will affect them. But we cannot directly enfranchise them. Therefore, we cannot avoid violating the All Affected Principle. If we insist that respect for the All Affected Principle is essential for democratic legitimacy, then it follows that democratic legitimacy is an impossible dream. No government is ever legitimate.

This theoretical puzzle for democratic legitimacy is easily overlooked in practice. As I noted in section 1, political philosophers have typically assumed that future people will be better off than present people, and that there are no significant conflicts of interest between present and future people. If we focus on building stable democratic institutions for ourselves, then the future will take care of itself. Our only obligation to future people is to bequeath our democracy. This is one place where the credible threat of a broken future undermines the optimism of contemporary political philosophy. The

intergenerational adequacy of democracy now becomes a very pressing concern.

Historically, there is a strong connection between social contract theory and democracy. Both are often motivated by the following desires: to respect the separateness of persons, to enable individuals to participate in decisions that affect them, and to avoid sacrificing individual interests on the altar of aggregate well-being. However, thanks to these similarities, democracy is also vulnerable to the objections to social contract theory that I sketched in section 1. The barriers facing any intergenerational contract also threaten intergenerational democracy. As Saunders himself forcefully argues non-identity and lack of reciprocity mean that techniques that successfully extend the franchise to include new present people are unlikely to work for future people.

Like social contract theorists, democrats have tended to sideline these difficulties by making optimistic assumptions about the benefits that present democracy can offer to future people. It doesn't matter whether future people are enfranchised now, because they will inevitably be so much better-off than us. The credibility of the broken future threatens this complacent defence of democracy. This raises the possibility that democracy is another moral ideal that does not translate to a broken world. And, if it might lead to a broken world, perhaps democracy is another indefensible affluent luxury that we ourselves should abandon.

In my book, I imagine a broken future where people have abandoned democracy, where the very idea of democratic government is greeted with derision, and where affluent democracy is identified as one principal cause of the broken world. While contemporary philosophers are almost universal in their support for democracy, my broken world philosophers are

universally suspicious of it. I made this rhetorical decision to emphasise the potentially radical impact of the broken world. (My imaginary philosophers' outright rejection of democracy contrasts with their more even-handed treatment of competing theories of justice—libertarianism, egalitarianism, utilitarianism. In my own recent work, I have used the broken world to argue that utilitarianism is superior to its rivals. But the book aims to be neutral about justice.)

Saunders raises a number of very telling questions about my approach to democracy in *Ethics for a Broken World*. He rightly observes that I do not offer any detailed description of the non-democratic society that operates in the broken world. This lack of detail is a general policy throughout the book. I tried to leave the background facts of everyday life in the broken world as open as possible, so that readers would not be distracted by unnecessary and arbitrary details. However, the lack of detail also limits the power of my implicit critique of democracy. Without a credible alternative, we cannot assume that democracy is not the best available system. Even if democracy might yield a broken future, we should only reject it if some identifiable alternative offers a better prospect.

In my book, I stipulate that people living in the broken world have abandoned democracy. This may suggest that I believe democracy *cannot* survive in such a world. However, I am not so pessimistic. In fact, I am inclined to agree with Saunders's suggestion that, despite its problems, democracy may be the best response to the threat of a broken world. As he says: "Democracy may not be perfect, but it is arguably less dangerous to the future than any feasible alternative (p. 29)."

One significant impact of the broken future is that we must reconsider the comparative plausibility of competing arguments for democracy. We can divide arguments for democracy into two

broad types: intrinsic and instrumental.¹³ An *intrinsic* argument claims that democracy is the only legitimate form of government, because democracy alone respects the equal rights, dignity, and autonomy of moral agents. For instance, if everyone has an equal right to participate in decisions that affect them, then people have a right to democracy—even if they would be better-off under a benevolent dictatorship. Democracy is thus the only legitimate system, whether or not it is the most efficient way to promote people’s interests. By contrast, an *instrumental* argument claims that, despite the superficial appeal of benevolent dictatorship, democracy is the most efficient way to promote some independently valuable end.

The divide between intrinsic and instrumental arguments often tracks the divide between contractualist and utilitarian theories of justice. Arguments for democracy within the social contract tradition typically focus on democracy’s intrinsic merits. Democracy is legitimate because it is what rational agents will choose in a fair and equal bargaining position. On the other hand, when utilitarians defend democracy they invariably offer instrumental arguments. For instance, John Stuart Mill famously argued that, under the right circumstances, representative democracy is the system of government that best promotes human well-being.

I argued in section 1 that, whatever their respective merits in relation to justice between contemporaries, utilitarianism is superior to contractualism in the intergenerational case. For similar reasons, while intrinsic and instrumental arguments for democracy both have their defenders, only instrumental considerations are relevant for future people. Future people cannot themselves participate in our present democracy. Therefore, we cannot plausibly claim that our democracy respects

¹³ *Ibid.*, chapter 16.

their right to participate. We can only offer the instrumental argument that democracy best protects the interests of future people. (This might include their interest in having their own democratic institutions.) If democracy is *not* instrumentally justified—if some alternative would do a better job for future people—then perhaps we should endorse that alternative instead.

As Saunders notes, even if democracy is “the only justifiable form of government in our circumstances,” we must weigh it against other values (p. 18). Saunders provides a nice example of the potential for conflict between democracy and other ideals. “Perhaps there is a sense in which it would be more democratic if decisions over my private life, such as what religion I should practice, were taken out of my hands and decided by a vote with universal franchise. But, if this is so, then giving me rights over my own private life is a justified departure from democracy” (p. 18). Here, the clash is between democracy and (other) individual rights. In a similar way, we might reasonably prefer a less democratic option today, if that would make life better for future people. (One intriguing possibility is that a non-democratic government today might be the best way to safeguard the rights of future people. I return to the potential for intergenerational conflicts of rights at the end of my reply to Tomalty in section 4.)

If we take seriously the possibility of a broken future, then we must focus, not on the intrinsic merits of democracy, but on its effects. In particular, we must ask whether democracy best serves the interests of future people. This is not a question that has received sufficient attention from political philosophers. Saunders addresses it perceptively in his commentary. He argues that the broken world does not provide sufficient reason to reject democracy. I am inclined to agree.

I begin with a debate *within* democratic theory that Saunders also discusses. This is the familiar debate between majoritarianism

and constitutionalism. Majoritarians favour an unrestricted democratic legislature with the power to remake all laws; while constitutionalists favour an entrenched bill of rights that places some decisions beyond the scope of the current majority. (These two theories roughly correspond to the constitutional systems of the UK and the USA respectively.) Some commentators have argued that, while majoritarianism is more democratic, an entrenched constitution will better protect the rights of future people. As Saunders notes, I have myself defended majoritarianism against this objection. In my earlier book, *Future People*, I argued on utilitarian grounds that the best protection for future people lies in the *motivations* of present people.¹⁴ If present people care for future people, then we will do what we can to promote their interests and safeguard their rights. And there is no good reason to believe that any minority—rulers, judges, experts, or whoever—will reliably perform better as guardians of future people than the majority. Saunders puts it well: “Either power is held by the masses or it is held by a few. Neither group is accountable to the future, so neither can be assured to treat future generations justly” (p. 27). The best hope for future people lies in the possibility that *all* present people will come to see themselves as guardians of the future.

The threat of a broken future raises the stakes in the debate between majoritarians and constitutionalists, and focuses our attention on instrumental arguments based on the impact of present decisions on future people. But I agree with Saunders that the broken world, in itself, does nothing to support constitutionalism over majoritarianism. There is no evidence that judges are more likely than legislators to defend the rights of future people, or to take a long-term view. As my broken world

¹⁴ T. Mulgan, *Future People*, chapter 8. This chapter also provides references to the wider debate between majoritarianism and constitutionalism.

thought experiment makes plain, there is no guarantee that democracy will avoid a broken future. But no alternative political system offers any guarantee either. In chapter 14 of *Ethics for a Broken World*, I ask the reader to imagine how a Rawlsian Supreme Court judge might act in response to the threat of a broken future. I suggest that such a judge, in such a situation, might take very radical (and very undemocratic) action. But the rhetorical point of that imaginary tale is to highlight the fact that, in the real world, we cannot rely on actual Supreme Courts to save us from the threat posed by climate change.

Saunders also explores various proposals to indirectly enfranchise future people. (These include sub-majority rules, the appointment of special representatives of future people, and so on.) I briefly discussed such measures in chapter 17 of *Ethics for a Broken World*. As with constitutionalism, I do not believe that these specific departures from majoritarianism offer any guarantee that we can avoid a broken future. And the same is true of even less democratic alternatives, such as benevolent dictatorship.

Current liberal democracy presupposes affluence and Rawlsian favourable conditions. Our democratic institutions cannot survive into a broken world. But it does not follow that democracy *per se* cannot thrive there. Broken world philosophers will need to re-imagine democracy, just as they must re-imagine many other central themes of affluent philosophy. In my précis of *Ethics for a Broken World* in this volume, I said that the central focus of broken world philosophy is the development of the just survival lottery. Perhaps the just lottery is also a democratic one. I will return to this possibility at the end of my reply, as it brings together themes raised by all three commentators.

III

Timothy Chappell on The Future-Person Standpoint

Timothy Chappell offers a very thought-provoking reading of my project in *Ethics for a Broken World*. His discussion of the ‘future-person standpoint’ is certainly a ‘fair extrapolation’ of the ideas presented in my book. The idea that we need to justify ourselves to future people is present in the book, but I do not think it is as explicit or systematic as Chappell suggests. Earlier versions of Chappell’s commentary have inspired my own recent work on possible futures, and his extrapolation is most welcome.

Chappell opens with two forms of moral argument that he suggests are central to *Ethics for a Broken World*. The first is as follows:

1. In the desperate circumstances of a broken world, X would be right.
2. But X looks *obviously* wrong to us.
3. Therefore: “maybe we should consider the possibility that our objections to X are less absolute than we taken them to be” (p. 31).

Chappell’s second form of argument goes like this:

1. In the affluent circumstances of our actual world, X looks *obviously* right to us.
2. But X would be wrong in the desperate circumstances of the broken world.
3. Therefore: “maybe we should consider the possibility that our acceptance of X in the actual world is less defensible than we take it to be” (p. 31).

I agree that both forms of argument do operate (below the surface) in *Ethics for a Broken World*. The qualifications that Chappell includes in both his conclusions are important. The fact that people in the broken world may have ethical intuitions or moral practices that differ from our own does not *prove* that *all* our moral beliefs are false or unreliable. I am not suggesting that readers of my book should simply abandon their affluent ethics in its entirety! However, thinking about a credible future where things that strike us as abhorrent are everyday facts of life, and where central elements of our own lifestyle are considered abhorrent, should give us pause for thought. Consider two examples. I argue in my book that future people will accept survival lotteries as the paradigmatic example of justice, and that they will regard our gratuitous consumption of fossil fuels as morally reprehensible. Suppose we accept these speculations about the future. Can we still be confident that we are justified in insisting on guaranteed rights for ourselves, or in defending our affluent lifestyle?

Chappell puts the point well: “Our perspective is, historically speaking, a most unusual perspective. What reason is there to expect it not to be also a *warping* perspective?” (p. 31) We naturally think that favourable conditions are the stable end-point of human history, and therefore that ethics should be adapted to such conditions. We think of our affluent moral sensibility as a sign of moral progress. But what if favourable conditions are merely a transient blip, and future philosophers will come to regard affluent ethics as a temporary aberration?

Chappell then outlines two ‘master-theses’ about moral argument that ‘may also structure Mulgan’s project’ in *Ethics for a Broken World*.

MT1: For an arrangement to be justified, it must be justifiable to the person who does worst out of it (p. 32).

MT2: For an arrangement to be justified, it must be absolutely justifiable—justifiable irrespective of any kind of contingency (such as the contingent circumstances of our own affluent society (p. 32).

These are interesting theses, and I have some sympathy for both of them. I also agree that several discussions in my book do suggest these principles. However, I am not convinced that Chappell's two master-theses, as he interprets them, reflect either the project I undertake in my book, or the way that I currently think about the broken world. I will briefly elaborate our disagreement at the end of this section.

Chappell links my project in *Ethics for a Broken World* to two other ethical traditions. The first is what he calls *Martianism*: the “deliberate adoption of an amazed outsider’s view on human arrangements” (p. 45). Chappell rightly notes that, in this regard, my book stands in a long philosophical tradition dating back to the Ancient Greeks. Chappell kindly describes my project as “a genuinely new variation on this ancient theme of Martianism” (p. 46). Unfortunately, while it may be unusual within academic philosophy, I do not think my use of imaginary futures is as original as Chappell suggests. After all, similar rhetorical devices are very familiar in science fiction. However, I certainly agree that one aim of my book is to use the imaginary reactions of an amazed outsider to draw attention to questionable and arbitrary features of our ethical practices.

As I suggested in section 1 of this reply, my use of the broken future goes beyond conventional *Martianism*, because its ‘amazed outsider’ is neither imaginary nor detached. My imaginary future philosophers are not Martians. They are not visitors from a distant world. Nor can we view their social arrangements with the detachment we might reserve for a broken world unconnected to ourselves. My broken world *is* our future. (Or, at least, it is one

credible future.) This brings us to Chappell's second ethical tradition.

Chappell draws a parallel between my use of the broken world and Stephen Darwall's recent emphasis on the importance of a *second-personal standpoint* in ethics. Our moral decisions often need to be justified, not merely *against* some impartial standard, but *to* the particular individuals who are affected. If I sacrifice your life for the common good, then I must offer a moral justification *that is addressed to you*.

I am sympathetic to Darwall's emphasis on the second-personal dimension of ethics. However, I also worry that this emphasis can lead us to under-estimate the comparative significance of our obligations to future people. It is much easier to offer second-personal justifications to present people than to future people. Our duties to contemporaries naturally engage our moral sentiments, because we must justify ourselves to the actual people whose real-life interests are affected by our actions. By contrast, distant future people are very remote from our everyday concerns. How can I justify myself to some particular individual living five hundred years in the future?

This worry connects with my earlier discussions of contractualism and democracy in sections 1 and 2. Especially in their debates with utilitarians, contractualists and democrats often cite the importance of justification to particular individuals. But this further highlights the barriers to extending contractualism and democracy to future people. Non-identity and the absence of reciprocity seem to make it impossible to offer the necessary second-personal justifications.

I now believe that one important function of my broken world project is to provide these future-directed second-personal justifications. Asking how our actions might impact on actual

future people can help redress the imbalance between present and future people, by giving our obligations to future people the same felt urgency as our obligations to one another. *Imagining* a credible broken future thus raises the *motivational* significance of our intergenerational obligations, as well as their theoretical significance.

Second-personal justification unites the concerns of all three commentators. Saunders argues that the need to justify ourselves to future people goes beyond the question of extending the franchise (p. 15); and the affluent ideal of human rights that Tomalty explores often emphasises a right to have one's complaints heard *and to receive a satisfactory justification*.

I am grateful to Tim Chappell for helping me to frame my own project in this way. One limitation of my book, of course, is that it imagines only one particular credible future. In one sense, this limitation is inevitable. If we are to construct second-personal justifications to future people, then we need to focus on particular futures that contain particular individuals. (We want to ask whether we could justify ourselves *to this particular person*, not merely to some abstract generality of possible future persons.) However, our ethical reasoning is incomplete if we consider only one possibility. In my recent work, I seek to address this limitation by imagining a range of other possible futures, including a virtual future where human beings spend their entire lives in Nozick's experience machine; a digital future where humans have been replaced by unconscious digital beings; and a theological future where the existence of God has been proved.¹⁵ I argue that, like the broken future, these other credible futures also affect our current ethical thinking in surprising ways. While some of the implications of alternative futures are distinctive,

¹⁵ For an overview of my current work here, see T. Mulgan, "Ethics for Possible Futures."

they often point in the same direction as the broken future. In particular, other futures also raise the importance of intergenerational ethics and alter the balance between competing moral theories.¹⁶

There is thus much in Chappell's commentary that I agree with, and much else that I find fruitful. I will close with some points of disagreement. These relate, perhaps unsurprisingly, to our different commitments in normative ethics.

In *Ethics for a Broken World*, I do not argue either for or against any particular moral theory. I believe this neutrality is appropriate in a textbook written for philosophical beginners. However, in my own more recent work, I do use the broken world to argue for utilitarianism and against its rivals, notably contractualism and libertarianism.¹⁷ I also argue for particular theses within utilitarianism, such as the objective list theory of well-being, and rule utilitarianism.

Chappell is not a utilitarian, and he wants to use his two master-theses to defend non-utilitarian conclusions. By contrast, insofar as I am willing to endorse those two master-theses, I read them in a manner that is consistent with utilitarianism. Indeed, I would argue that utilitarianism can help us to unify Chappell's two theses.

Chappell's two theses deploy two distinct notions of justification. In MT1, we justify ourselves *to an individual*; while MT2 seeks an absolute non-contingent justification. MT1 thus suggest a focus on particular individual perspectives, while MT2

¹⁶ A third common feature is that all my possible futures, including the broken world, push morality in a more objective direction. (T. Mulgan, "Ethics for Possible Futures," section 2.)

¹⁷ T. Mulgan, "Utilitarianism for a broken world"; "Contractualism for a broken world"; "Ethics for Possible Futures."

suggests an impartial standard that is independent of individual differences. Chappell himself uses the concept of virtue to unite the two theses. Virtues respond to individuals, but virtues are also justified absolutely because “the traditional virtues do make sense even in [broken world] scenarios” (p. 33). I will now argue that the utilitarian tradition provides an alternative unification.

Utilitarianism need not be antagonistic to second-personal justification. The caricature of the utilitarian is a calculating machine who treats individuals merely as anonymous utility-containers. While there is some justice in this caricature, it does not represent all utilitarians. In particular, the liberal utilitarian tradition of John Stuart Mill emphasises the importance of individuality, freedom, autonomy, and individual moral judgement.¹⁸

The utilitarian tradition encompasses a bewildering variety of themes and theories. To get a manageable discussion in *Ethics for a Broken World*, I artificially divided utilitarians into two main groups: act utilitarians and rule utilitarians. (While these are familiar names, I use them more broadly, to contrast two composite positions.) This artificial division captures two competing strands to the utilitarian tradition—radical iconoclasm and liberal moderation.

Act utilitarians defend a purely impartial moral theory. They evaluate individual actions solely by their impact on aggregate human pleasure; and they accept the resulting verdicts, however extreme or counterintuitive. *Rule utilitarians* favour a moderate morality. They picture morality as a collective enterprise, and evaluate moral codes and political institutions by their collective

¹⁸ I explore Mill’s utilitarianism further in Mulgan, “Mill and the broken world,” *Revue Internationale de Philosophie* (2014).

impact on human well-being. Rule utilitarians endorse many non-utilitarian prohibitions, permissions, rights, and freedoms.

The broken world impacts differently on different versions of utilitarianism. The beauty of act utilitarianism is a simple moral principle that applies to any situation. Act utilitarians hold that, whatever her circumstances, every agent should always perform the action that produces the best consequences. In a broken world, act utilitarians face no significant theoretical difficulties, as this simple principle carries over unchanged. However, the broken world does exacerbate one perennial difficulty for act utilitarianism—its counter-intuitiveness. In particular, act utilitarianism is notoriously demanding even when confined to an affluent present. The broken world greatly exacerbates those demands. (Think of all those distant future people, worse-off than us, whose well-being is so dependent on our actions!)

Act utilitarianism is so extreme because it pictures morality as a project given to a single utilitarian agent who must heroically maximize human happiness in a non-utilitarian world. Unsurprisingly, her life is demanding, alienating, and unattractive. But this individual model seems especially out of place against the backdrop of a broken future—where the most pressing moral issues are collective and intergenerational. (Consider the futility of asking what I should do to avert dangerous anthropogenic climate change!) In this new context, the rule utilitarian picture of morality as a task given, not to each individual agent, but to a *community of human beings* begins to seem much more apt.

For rule utilitarians, the fundamental moral questions are: ‘What if *we* did that?’, and ‘How should *we* live?’ My version of rule utilitarianism draws on Brad Hooker’s recent formulations.¹⁹

¹⁹ B. Hooker, *Ideal Code, Real World* (Oxford: Oxford University Press, 2000). I develop my rule utilitarianism in T. Mulgan, *The Demands of Consequentialism*

We first seek an *ideal moral code*. Acts are then assessed *indirectly*: the right act is the act called for by the ideal code. We imagine ourselves choosing a moral code to govern our community. I operationalize this by asking what would happen if we (the present generation) attempted to teach a given moral code *to the next generation*. This sets aside the cost of changing existing moral beliefs, but factors-in the cost of (for instance) trying to get a new generation to accept a very demanding ethic.

Rule utilitarianism promises an overarching moral theory grounded in the utilitarian tradition—one that bases morality on the promotion of well-being, but avoids the extreme demands and injustices of act utilitarianism. Drawing on arguments made famous by Mill, rule utilitarians champion their ability to accommodate a wide range of common-sense rights and freedoms, and to favour democratic government over despotism, liberal society over its rivals, free markets over command-and-control economics, and so on. One pressing question is whether this liberal moderation can survive into a broken world. I return to that question in my discussion of survival lotteries in section 4.

Act utilitarianism offers a single mode of justification. The right action is whatever, in the circumstances, happens to maximise aggregate well-being. Suppose one individual's interests are sacrificed for the greater good. If this sacrifice maximises aggregate well-being, then it is justified. This is the only justification that is available, and the *same* justification will be offered both to the injured person herself and to any interested bystander. There is no room in act utilitarianism for distinctively second-personal justification. The fact that *my* interests are at

(Oxford: Oxford University Press, 2001), especially chapter 3; *Future People*, chapters 5 and 6; *Ethics for a Broken World*, chapter 7; “Utilitarianism for a Broken World”; “Mill and the Broken World.”

stake has no moral significance, and so there is no need to justify anything *to me*.

Those, like Chappell and me, who are sympathetic to Darwall's second-personal emphasis thus have good reason to reject *act* utilitarianism. But it does not follow that we must abandon utilitarianism altogether. Rule utilitarianism has resources that its simpler sister-theory lacks. Given the importance of the second-personal standpoint *to human beings*, it is quite possible that the ideal moral code will recognise the importance of justifications that are addressed to particular individuals. Within rule utilitarianism, these justifications are ultimately grounded in the promotion of aggregate well-being. But this does not undermine their second-personal sincerity. (In the same way, the rule utilitarian's commitment to keep her promises is not rendered insincere by the fact that its ultimate justification rests on aggregate well-being.)

Of course, there are limits to the flexibility of rule utilitarianism. If circumstances change *too much*, then even the most familiar moral rules may need to give way. In a broken world, many of our most cherished affluent privileges, permissions, and prohibitions may make no sense. (I explore several examples in chapter 7 of *Ethics for a Broken World*, and in my reply to Tomalty below.) This marks a theoretical divide between Chappell and me. Chappell interprets MT2 as the claim that particular virtues can be given some absolute justification that is unrelated to contingency and circumstance. For the rule utilitarian, by contrast, what is absolute is only the *standard* of justification. The measure of aggregate well-being is unvarying, but the particular moral code that it justifies may vary with the circumstances.

On the other hand, my theoretical disagreement with Chappell may not generate any great practical disagreement. Rule

utilitarianism is open to the possibility that there are some very general virtues that carry over into the broken world, and into any other credible future. Indeed, there is good utilitarian reason to expect this. For the rule utilitarian, each virtue is a useful collective response to some central aspect of human well-being. If this is so, then we might well expect to find that, in any actual future, the ideal code includes analogues of generosity, benevolence, truthfulness, kindness, courage, prudence, and the like.

Rule utilitarianism offers another link between Chappell's two master-theses. Following Mill, modern rule utilitarians often defend liberal democratic institutions. Within those institutions, decisions are made by following rules that respect individual rights and liberties, not by the calculation of aggregate utility. Real-life second-personal justifications often cite these intermediate institutions. If someone is affected by my actions, then the following is often a perfectly good justification to offer her: 'I have treated you in accordance with the just rules that govern our established institutions'. In a broken world, rule utilitarians will continue to offer such institution-based justifications, even though the precise details of their utility-maximising institutions may differ.

IV

Jesse Tomalty on human rights

Jesse Tomalty offers a very thoughtful discussion of the place of human rights in a broken world. She also asks how the prospect of a broken future might impact on our thinking about rights in our (comparatively) affluent present. I have been

thinking about rights in the broken world a great deal since I wrote my book, and I have found Tomalty's comments extremely helpful.

As Tomalty notes, *Ethics for a Broken World* presents rights as an *affluent* obsession. This suggests that people in the broken world have outgrown this obsession, and they are no longer interested in rights. As with democracy, this is a deliberate rhetorical decision, designed to generate what Chappell might call a Martian outsider's perspective on our own affluent rights-talk. As I noted in my précis, affluent conceptions of rights do not translate easily to the broken world. However, I personally believe that broken world dwellers may still find a use for the general concept of rights.

In my book, I do not address *human* rights per se. I focus instead on general theories of rights (libertarian, utilitarian, Rawlsian), and on rights to self-ownership, property, liberty, or democratic participation. This omission was partly owing to constraints of space, and partly because, unlike Tomalty, I have no particular expertise in human rights theory. However, this omission does not mean that human rights have no place in the broken world. On the contrary, I agree with Tomalty that, if they find any use for the concept of rights, people in the broken world will be interested in human rights.

Tomalty notes that some approaches to human rights that work for us would make no sense in a broken world. Without favourable conditions, it is not possible to respect all human rights (especially if we retain our modern conception of what those rights entail). Nor will it make sense for future people to maximise human rights, as there would be no room for any other human activity. Every resource would be devoted to mere survival.

Tomalty offers an interesting discussion of the different categories of rights. Scarcity has an obvious impact on ‘socio-economic’ rights, such as rights to food, shelter, education, health care, and so on. In a broken world, where resources are not sufficient to meet all basic needs, these rights cannot possibly be guaranteed to everyone. This difficulty is especially significant for utilitarians. As I noted in my précis of *Ethics for a Broken World*, utilitarians insist on socio-economic rights. Without these positive rights, utilitarians cannot reasonably give rights the priority traditionally associated with them.

This might suggest that a broken world is only problematic for defenders of socio-economic rights. However, as Tomalty insightfully points out, other rights may also be under threat in a broken world. In particular, many civil and political rights rely on legal institutions. The practical protection of these rights thus depends on the continued existence of those institutions. The maintenance of just legal institutions is very costly. If resources are scarce, then legal institutions may be stretched to breaking-point, and even the most negative libertarian rights may be under threat. (In Part One of *Ethics for a Broken World*, I discuss a number of other ways that libertarian rights might collapse under the prospect of a broken future.)

There is nothing in Tomalty’s discussion of human rights that I disagree with. And I have learnt much from studying her account. In the remainder of this reply, I want to explore in more detail the possible roles that human rights might play in a broken world, drawing on several of Tomalty’s own observations.

Tomalty notes that some human rights translate relatively easily. Rights to non-discrimination, equal treatment, and due process are as compelling in a broken world as they are in our affluent present. There is no guarantee that these human rights *will* survive into a broken world. One very real threat is that, as

the world breaks, illiberal anti-human-rights regimes may emerge. But the *justification* for anti-discrimination rights remains intact, as it does not depend on favourable conditions.

Tomalty highlights one possible future role for human rights: “inhabitants of the broken world could articulate the injustice of their situation in terms of human rights” (p. 55). Future people may express their own moral outrage at the conduct of their affluent ancestors in terms of the violation of their own human rights. Drawing on influential work by Simon Caney, Tomalty points out that, on most accounts of human rights, our failure to avoid a broken world will be a major human rights violation unless we can prove that we could not avoid this result without violating some present human rights.

I agree that future people may find the current philosophical debate over climate change and human rights helpful. I would also suggest that they will re-interpret some key terms in that debate. For instance, one popular theme is the distinction between *luxury* emissions and *subsistence* emissions.²⁰ While all present people share a common responsibility to reduce their CO2 emissions to protect the human rights of future people, each present individual also has a basic human right to engage in essential subsistence emissions. (There is a clear difference between the British hobby farmer who raises cattle to take advantage of lucrative EU subsidies and the sub-Saharan subsistence farmer who must raise cattle to survive.) The obligation to fight climate change by reducing emissions does not extend to subsistence emissions. Our common responsibility is thus highly differentiated in practice, because anyone whose emissions are entirely subsistence-based has no responsibility to do anything.

²⁰ See, e.g., H. Shue, “Subsistence emissions and luxury emissions,” *Law and Policy* Vol. 15 (1993), 39-60.

In a broken world, where universal survival is impossible, there can be no guaranteed human right to subsistence. Broken world philosophers have two choices. They could regard this human right as something that should have guided affluent decisions, but no longer applies in their broken world. They would then abandon human-rights-talk altogether. Alternatively, as with other aspects of affluent rights-talk, broken world philosophers might try to adapt this human right to their own situation.

To illustrate the ways that rights might evolve in a broken world, let us imagine how a future broken world rule utilitarian, inspired by J. S. Mill, might design a just survival lottery.²¹

Any survival lottery is a very striking departure from contemporary liberal democratic ideals, as it must violate the basic rights and freedoms that affluent liberals claim *for all*.²² Many liberals will refuse to countenance any survival lottery, and thus deny that justice has any place in a broken world. At the other extreme, act utilitarians will embrace the survival lottery as the obvious and morally unproblematic solution in a dire situation. The Mill-inspired utilitarian liberal is both (as a utilitarian) willing to think the unthinkable, but also (as a liberal) anxious to preserve our rights and freedoms as far as possible.

Act utilitarians will favour a simple hedonist survival lottery—an impersonal aggregative procedure that maximises total pleasure across generations. But my Mill-inspired rule utilitarian, like Mill himself, is not a simple hedonist. Well-being is not simply a matter of counting pleasures. A good human life also

²¹ This section draws on T. Mulgan, “Mill and the Broken World”; and my various discussion of survival lotteries throughout *Ethics for a Broken World*.

²² This is why Rawlsian liberals have such difficulty with the allocation of health care, where the scarcity of resources relative to needs mirrors a broken world.

requires the pursuit of freely-chosen objectively worthwhile goals. For Mill, individuality, self-realization, and freedom are central to well-being. While freedom without survival makes no sense, survival without freedom has no value. Freedom is one of the things that makes survival valuable. Our rule utilitarian might thus rationally accept a lower chance of individual survival in exchange for greater freedom for those who do survive.

Once we admit values other than pleasure or survival, we are confronted by a potentially infinite array of possible survival lotteries—offering different chances of survival, and different opportunities and goods for survivors. The traditional hedonist ranks these lotteries using aggregate pleasure. Mill offers no simple impersonal metric. He would rely instead on the verdicts of his famous *competent judges*.²³ If everyone would prefer a lottery with more freedom and less survival, then this is very good evidence that such a lottery best promotes human well-being. Indeed, it is the only possible evidence.

How will Mill's competent judges, dwelling in their broken world, design a just survival lottery? We know they must depart from our current notion of rights. Scarcity of resources requires a shift from guaranteeing everyone's survival to managing a fair distribution of chances to survive. Instead of vainly trying to guarantee a worthwhile life for all, a just lottery will guarantee everyone a fair and equal *chance* of living a worthwhile life. This is one place where human rights associated with procedural fairness and non-discrimination come to the fore. Indeed, in a broken world, these rights become even more important. Consider the contrast between a survival lottery that selects individuals on the basis of membership in a privileged group, and one that selects its

²³ J. S. Mill, *Utilitarianism*, in *Essays on Ethics, Religion, and Society*, J.M. Robson (ed.) (Toronto/London: University of Toronto Press/Routledge & Kegan Paul, 1969), chapter 2.

survivors at random. Even if the former lottery seems more ‘efficient’, it clearly violates our sense of procedural fairness. Under extreme scarcity, only a random procedure can respect people’s procedural human rights. The survival lottery must literally be *a lottery*. Tomalty puts the point well: “Denying someone access to the means for subsistence might not constitute a failure to respect a person’s moral status in the broken world, but denying him access to the means for subsistence simply because of his ethnicity plausibly would could as such a failure of respect” (p. 57).

Act utilitarians cannot accommodate these procedural human rights. But, for reasons similar to those outlined in my reply to Chappell in section 3, I believe that rule utilitarians can respect them. Given the intimate connections between self-respect and well-being, the moral code that maximises human well-being will be one that respects the moral status of persons. And, as Tomalty herself argues, this respect lies at the heart of our ideal of human rights (p. 56).

Another key difference between act and rule utilitarians is their attitude to disagreement. For the act utilitarian, moral disagreement is a sign of error, prejudice, or caprice. The moral truth is identified with the impersonal unchanging utilitarian calculus. For the Mill-inspired rule utilitarian, by contrast, moral disagreement is a positive and dynamic force. Given the complexities of human well-being, and the many different ways that individuals can reasonably respond to different values, it is natural to expect rational individuals to disagree.

Mill’s attitude to disagreement is linked to his faith in moral progress. It may seem odd to speak of progress in relation to the broken world. But our concern here is with improvements in moral knowledge. Those who dwell in the broken world are best-placed to adjudicate between competing survival lotteries. Their

verdicts are more reliable than ours. Each new generation in a changing world is better-placed than its parents to make moral judgements pertaining to its own situation.

These rule utilitarian commitments to reasonable disagreement and moral progress have several implications. One implication is that, even within a given generation, different individuals may reasonably reach different verdicts—depending on their personal evaluation of competing values and risks. Perhaps no possible *uniform* lottery will suit all competent judges. In *Ethics for a Broken World*, I explore the possibility that broken world philosophers may develop *flexible* lotteries, where each participant selects her own mix of risk and reward.²⁴ (Tomalty picks up on this suggestion in relation to the range of reasonable future interpretations of human rights.)

We cannot separate the content of the rights allocated within the lottery from the procedural right to participate in choosing the lottery. This is why no central bureaucracy or code of rules could hope to either design the best lotteries, or choose between them. The best way to ‘design’ a lottery for future people is to enable them to design it for themselves. Mill offers inspiration here. His own ideal of the utilitarian reformer is not a single-minded bureaucrat, nor a slavish follower of some simple utilitarian principle or code, but rather a creative individual engaged in a moral ‘experiment in living’ who invites her fellow citizens to follow her example. In the broken world, moral entrepreneurs might imagine new ways of life, new moral ideals, and new survival lotteries.

Rule utilitarian attitudes to disagreement also support democracy and majoritarianism. Once several candidate lotteries are on the table, the choice between them should be as

²⁴ T. Mulgan, *Ethics for a Broken World*, 185-196.

democratic as possible. There are three reasons for this: (1) Each individual has a right to participate in decisions that affect her; (2) Each individual is best-placed to know what will best promote her own interests; and (3) At any point in time, present people are better-placed than their predecessors to decide how competing interpretations of human rights should be balanced against one another. If we seek to entrench *our* interpretation of human rights, we will be imposing our partial opinions on future people. This is neither fair nor efficient.

At the end of section 1 of this reply, I suggested that a credible broken future raises troubling practical questions about how we should live *now*. Can we reasonably justify a refusal to adopt the ethical outlook of the broken world *for ourselves*? The removal of favourable conditions raises an even more disturbing question. Suppose we conclude that, while we can guarantee our own basic needs, our descendants will need to run a survival lottery. Can we still insist on guaranteed survival for ourselves, or should we move in their direction—operating a survival lottery across the generations? If future people must rethink *their* human rights, and replace guarantees with lotteries, then perhaps we should rethink *our* human rights too. (Otherwise, we are simply asserting that our rights are more important than theirs.) Perhaps, faced with a broken future, even rule utilitarians must embrace a more demanding and austere morality.

The survival lottery strikes us as morally unthinkable. But if we leave future people in a place where *they* must think the unthinkable, then perhaps we should think it too. Perhaps the design of a just survival lottery should be *our* central philosophical concern as well.

If the future is too broken—if no fair lottery offers everyone a decent chance of survival—then liberal institutions cannot survive. But in a less-broken future, lessons from the utilitarian

tradition may keep alive the hope of safeguarding fairness, security, and individuality. Human rights and democracy, suitably re-imagined, may have a central role to play in the resulting utilitarian lottery.

University of St Andrews

&

The University of Auckland

If you need to cite this article, please use the following format:

Mulgan, Tim, "Replies to Critics," *Philosophy and Public Issues (New Series)*, Vol. 4, No. 2 (2014), 59-92, edited by S. Maffettone, G. Pellegrino and M. Bocchiola