

## 6

### Ideal Theory as Faith

Chapter 2 suggested that we can understand apocalyptic thought as a form of ideal theory, since it identifies an ideal and theorizes a path to it. Apocalyptic thought and ideal theory, it turns out, share more in common than is often assumed. This chapter builds on that idea. In doing so, it doesn't cast aside earlier methodological recommendations and suggest, without explicit evidence, that apocalyptic thought exercises insidious influence over ideal theory today. But the chapter does explore parallels between ideal theory and apocalyptic thought, with a focus on what grounds people have for believing them.

The most influential understanding of ideal theory comes from John Rawls, who explains it as offering principles of justice that members of a liberal democratic society have reason to accept. Ideal theory, according to this view, has navigational value: it outlines a shared goal – what Rawls calls a “realistic utopia”<sup>1</sup> – for those in society to strive toward. Apocalyptic and other religious beliefs, on the other hand, are not based on reasons all can be expected to accept. Individuals can hold such beliefs on faith, but unlike ideal theory, it would be unreasonable to demand the rest of society to adopt and pursue their goals. For Rawls, then, there is a neat distinction between ideal theory and religious belief: the former is based on plausible reasons that others should accept, whereas the latter is unsuitable to guide society as a whole.<sup>2</sup>

This distinction proves too neat. The grounds for ideal theory turn out to be shakier than ideal theorists tend to admit. Indeed, in recent years, there have been growing concerns over ideal theory. What has resulted is intense debate over the topic in political philosophy, which at times has grown insular and

<sup>1</sup> John Rawls, *The Law of Peoples* (Cambridge, MA: Harvard University Press, 1999), 12. See also Rawls, *Justice as Fairness: A Restatement*, ed. Erin Kelly (Cambridge, MA: Harvard University Press, 2001), 13.

<sup>2</sup> Rawls, *Justice as Fairness*, 26–29.

arcane, as understandings of ideal theory have multiplied.<sup>3</sup> This chapter suggests that a less insular approach, in dialogue with social science and other research, stands the best chance of advancing the current debate over ideal theory and its limitations.<sup>4</sup>

Specifically, my argument builds off criticisms of ideal theory by Gerald Gaus,<sup>5</sup> and draws on social science research to put them on firmer ground. To plausibly defend an ideal theory, it is necessary to show that its principles would have normative force in the future society it envisions. But research on human prediction undermines the claim that we can plausibly know societal conditions in the distant future and what principles of justice would look like under those conditions. The immense complexity of social phenomena and the occurrence of Black Swans – rare, difficult-to-predict events with transformative effects<sup>6</sup> – severely constrain human predictive capacities. Evidence for this point comes from forecasting studies, which suggest no reason for confidence in predictions about society for the distant future.<sup>7</sup> Since defenses of ideal theory *depend* on such predictions, they necessarily fail. Similar to religious and apocalyptic belief, ideal theory lacks plausible grounds and ultimately rests on faith. So contrary to Rawls's view, people do not have compelling reason to accept any proposed account of ideal theory.

Though Rawls's approach runs into insurmountable problems, it is important to appreciate its normative appeal. Too often, critics fail to grasp the moral instincts that motivate ideal theory. When ideal theory aims to identify the

<sup>3</sup> See, e.g., Alan Hamlin and Zofia Stemplowska, "Theory, Ideal Theory and the Theory of Ideals," *Political Studies Review* 10, no. 1 (2012): 48–62; Zofia Stemplowska and Adam Swift, "Ideal and Nonideal Theory," in *The Oxford Handbook of Political Philosophy*, ed. David Estlund (New York: Oxford University Press, 2012), 373–88; Laura Valentini, "Ideal vs. Non-ideal Theory: A Conceptual Map," *Philosophy Compass* 7, no. 9 (2012): 654–64; and Kwame Appiah, *As If: Idealization and Ideals* (Cambridge, MA: Harvard University Press, 2017), 112–72.

<sup>4</sup> A similar recommendation comes from Lisa Herzog, "Ideal and Non-ideal Theory and the Problem of Knowledge," *Journal of Applied Philosophy* 29, no. 4 (2012): 271–88. The goal here is to put this recommendation into practice by examining findings from empirical research, specifically on prediction, to better understand limits to ideal theorizing. This analysis leads to a conclusion far more skeptical than Herzog's on ideal theory's potential to guide collective action.

<sup>5</sup> Gerald Gaus, *The Tyranny of the Ideal: Justice in a Diverse Society* (Princeton, NJ: Princeton University Press, 2016). See also David Wiens, "Against Ideal Guidance," *Journal of Politics* 77, no. 2 (2015): 433–46; and "Political Ideals and the Feasibility Frontier," *Economics and Philosophy* 31, no. 3 (2015): 447–77.

<sup>6</sup> Nassim Taleb, *The Black Swan: The Impact of the Highly Improbable* (New York: Random House, 2010).

<sup>7</sup> Philip Tetlock, *Expert Political Judgment: How Good Is It? How Can We Know?* (Princeton, NJ: Princeton University Press, 2005); and Philip Tetlock and Dan Gardner, *Superforecasting: The Art and Science of Prediction* (New York: Crown Publishers, 2015).

most just society possible, it seeks information relevant for guiding action in a complex world where the path to greater justice is far from straightforward. Knowing the ideal helps avoid paths that, though appealing in isolation, lead away from the ideal. Without a plausible ideal to guide action, political philosophy is left with deep uncertainty over how best to advance justice long term. One never knows if certain actions move society closer to or further from its most just form possible. Even efforts that advance justice *now* risk taking society down paths that close off greater justice *later*.

Regrettably, ideal theory finds itself unable to escape this dilemma. There is perhaps, though, still a role for it in advancing justice. One function of ideal theory is to offer hope in the possibility of a just society.<sup>8</sup> Both ideal theory and apocalyptic thought offer *utopian hope*, which gives meaning to imperfect, partial efforts to advance justice by understanding them as steps toward the ideal within the long arc of history. Utopian hope goes beyond desiring and believing in the possibility of justice in the short term, which even if realized can often be fleeting in the face of new political developments that threaten to overturn progress. Utopian hope sets its sights on a far loftier goal: a future that *ultimately* proves hospitable to justice and the ideal society. Ideal theory offers such hope to sustain people when the immediate prospects of justice seem bleak.

Utopian hope offers psychological benefits and, understandably, some embrace it. For those who do so by turning to either religion or ideal theory, it is important to be honest that such beliefs rely on faith. The chapter closes with a look at Rawls's writings to show that, though its role is often downplayed, faith has been an inextricable part of contemporary ideal theory from the start. Once we recognize that point, it becomes clear that political philosophy must rethink ideal theory's role in advancing justice.

### IDEAL THEORY'S NORMATIVE APPEAL

Before getting into ideal theory's limitations, let's first look at what draws people to it. Ideal theory takes different meanings, and here the focus is on

<sup>8</sup> Rawls, *The Law of Peoples*, 128; *Justice as Fairness*, 37–38; and *Political Liberalism*, exp. ed. (New York: Columbia University Press, 2005), ix. See also Paul Weithman, *Why Political Liberalism? On John Rawls's Political Turn* (New York: Oxford University Press, 2010), 367–69; Benjamin McKean, "Ideal Theory after Auschwitz? The Practical Uses and Ideological Abuses of Political Theory as Reconciliation," *Journal of Politics* 79, no. 4 (2017): 1177–90; and Dana Howard, "The Scoundrel and the Visionary: On Reasonable Hope and the Possibility of a Just Future," *Journal of Political Philosophy* 27, no. 3 (2019): 294–317.

what I call *navigational ideal theory*.<sup>9</sup> This conception of ideal theory seeks to outline the best and most just society with the potential of being realized at some future point.<sup>10</sup> I focus on this conception because, among competing understandings of ideal theory, it is the best candidate to serve as a normative guide to action.

To explain navigational ideal theory's normative appeal, it helps to contrast it with a conception less suited to serve as a moral guide. Sometimes ideal theory refers to *idealization*, meaning that it assumes ideal conditions that are rare or impossible in order to isolate and explain a concept. This manner of theorizing is common in science. Isaac Newton explains gravity by introducing an ideal pendulum, which experiences no friction or air resistance. This pendulum does not represent the perfect pendulum engineers should aim to build. Rather, the term ideal conveys that the pendulum functions under simplified conditions, which put the focus on gravitational force.<sup>11</sup> A similar rationale motivates thought experiments in philosophy that assume ideal conditions to better understand our intuitions about a normative concept. For instance, though others do not always contribute their fair share, imagining that they do illuminates basic intuitions about what fairness demands.

Ideal theorists sometimes construct idealizations with bolder normative ambitions: to present an end goal, ideal justice, to strive for. This approach specifies a goal under conditions that rarely if ever hold in reality. Constraints in the real world, but absent from an idealization, can make the ideal impossible. Ideal theory thus can send people chasing after utopia in one sense of the word – “nowhere,” an ideal that doesn't exist now and is impossible to ever realize. Ideal theory that sets forth an impossible ideal raises concerns because it risks having perverse normative effects when the ideal looks much different from the best possible option.

To illustrate, imagine a perfect society. Many envision a place free from injustice where everyone always acts justly. But suppose weakness of will and other shortcomings prevent this ideal from ever being realized regardless of what society does – there always will be some who act unjustly. The perfect society imagined is what some call a “hopeless” ideal, meaning that the ideal

<sup>9</sup> Some use the term “realistic utopianism,” which comes from Rawls, to communicate a similar idea. See Ben Laurence, “Constructivism, Strict Compliance, and Realistic Utopianism,” *Philosophy and Phenomenological Research* 97, no. 2 (2018): 433–53.

<sup>10</sup> This understanding of ideal theory focuses on what is possible indefinitely into the future, not just now, so as to determine an ultimate goal. For a defense of this assumption, see the Appendix and specifically the response to Objection 1.

<sup>11</sup> Jenann Ismael, “A Philosopher of Science Looks at Idealization in Political Theory,” *Social Philosophy and Policy* 33, nos. 1–2 (2016): 11–31.

requires action that each individual has the ability to do, yet collectively there is virtually no chance that everyone will do their part to realize the ideal.<sup>12</sup> If a society characterized by perfect compliance is indeed a hopeless ideal, it likely would be unwise to pursue it. Though some measures to increase compliance advance justice, beyond a certain point they backfire – like draconian surveillance to combat noncompliance. Indeed, such surveillance often backfires and leads to racially disparate outcomes.<sup>13</sup> We thus should be cautious of pursuing hopeless and unattainable ideals since they can lead away from the best possible option.<sup>14</sup>

There are legitimate concerns, then, regarding ideal theory's normative value. Yet this common criticism – ideal theory fails as a normative guide because it embraces unattainable ideals – is not a flaw inherent to it. Navigational ideal theory avoids this pitfall by focusing on what is collectively feasible and identifying the most just society possible. In this way, it steers clear of unattainable ideals that would undermine its normative value.

But this approach still has its critics.<sup>15</sup> It is helpful to differentiate common objections against navigational ideal theory so as to be clear why it might fail as a normative guide:

- (a) There is no such thing as the most just society possible.<sup>16</sup>
- (b) Ideal theory seeks information – what is the most just society possible – irrelevant to advancing justice.<sup>17</sup>
- (c) Ideal theory cannot identify the most just society possible.<sup>18</sup>

According to (a) and (b), ideal theorists ask the wrong question – what is the most just society possible? – when formulating a theory of justice. That

<sup>12</sup> David Estlund, "Utopophobia," *Philosophy and Public Affairs* 42, no. 2 (2014): 117–18.

<sup>13</sup> See James Forman Jr., *Locking Up Our Own: Crime and Punishment in Black America* (New York: Farrar, Straus and Giroux, 2017), 185–215.

<sup>14</sup> This concern applies to Rawls's ideal theory, which assumes strict compliance – that is, everyone upholds the principles of justice. See Rawls, *A Theory of Justice*, rev. ed. (Cambridge, MA: Harvard University Press, 1999), 8. This idealization seems incompatible with his intention of offering a *feasible* goal, and thus may result in a theory that directs people toward an unattainable ideal when the best feasible option looks much different. See Colin Farrelly, "Justice in Ideal Theory: A Refutation," *Political Studies* 55, no. 4 (2007): 844–64; and Laurence, "Constructivism, Strict Compliance, and Realistic Utopianism."

<sup>15</sup> See, e.g., Amartya Sen, "What Do We Want from a Theory of Justice?" *Journal of Philosophy* 103, no. 5 (2006): 215–38; Amartya Sen, *The Idea of Justice* (Cambridge, MA: Harvard University Press, 2009); David Schmidtz, "Nonideal Theory: What It Is and What It Needs to Be," *Ethics* 121, no. 4 (2011): 772–96; and Jacob Levy, "There's No Such Thing as Ideal Theory," *Social Philosophy and Policy* 33, nos. 1–2 (2016): 312–33.

<sup>16</sup> See, e.g., Schmidtz, "Nonideal Theory," 774.

<sup>17</sup> See, e.g., Sen, "What Do We Want from a Theory of Justice?" 221–22.

<sup>18</sup> See, e.g., Gaus, *The Tyranny of the Ideal*, 139–44.

question is the wrong place to start because there is no such thing as the ideal society and, even if there were, knowledge of it would prove useless for advancing justice. In contrast, (c) avoids the strong claim that ideal theory is fundamentally misguided, and instead argues that ideal theory cannot answer the question it poses, no matter how valuable the answer would be. We'll examine (c) later in the chapter, but the remainder of this section focuses on the more fundamental critiques of ideal theory, (a) and (b).

Some find (a) compelling, uneasy with the idea that justice takes one perfect, platonic form. These critics believe the focus should be on eliminating injustice rather than climbing toward some illusory peak form of justice. As Amartya Sen puts it, we care about preventing famine, less so about whether a 45 or 46 percent tax rate for top earners best represents justice.<sup>19</sup>

Without question, ending grave injustice deserves priority. But that view is compatible with ideal theory. Ending the world's many injustices is no small feat, and constitutes its own lofty ideal that can serve as an end goal in efforts to advance justice. The interconnected nature of society demands that we look for an optimal approach, since ending one injustice can exacerbate others. As an example, consider Abraham Lincoln who secretly promised government jobs and other perks to Democrats to ensure passage of the Thirteenth Amendment abolishing slavery.<sup>20</sup> The injustice of slavery likely justified such bribery, but most would be uncomfortable with its continuing indefinitely. Ideal theory keeps in view the many aspects of justice worth striving for. Rather than necessarily condemning all compromises, it reminds us to consider how compromises impact efforts to advance toward an ideal goal.

So ideal theory emphasizes that advancing justice demands more than tackling one injustice at a time: it requires a holistic view focused on what social arrangements best eliminate various injustices. For the pursuit of that goal, there is reason to hang on to the climbing metaphor common to ideal theory. Societies are not simply just or unjust, but exhibit degrees of injustice. Efforts against injustice often are intergenerational and build off past successes with an ideal in mind, which is why the metaphor of climbing toward a peak is apt.

Sen is perhaps right that fine distinctions about perfect justice are not critical. Society could take various forms, whose comparative levels of justice vary slightly, but which all eliminate significant injustices. We want to reach

<sup>19</sup> Sen, "What Do We Want from a Theory of Justice?" 223.

<sup>20</sup> See James McPherson, *Battle Cry of Freedom: The Civil War Era* (New York: Oxford University Press, 1988), 839; and John Parrish, "Benevolent Skullduggery," in *Corruption and American Politics*, ed. Michael Genovese and Victoria Farrar-Myers (Amherst, NY: Cambria Press, 2010), 78–79.

one of those ideals and care less about reaching the absolute best one. That point implies revising, not abandoning, ideal theory. In efforts to advance justice, what is important is not necessarily ending up atop Everest, so to speak, but aiming for an eight-thousander – one of those rare peaks over 8,000 meters – free from the injustices of lesser peaks and valleys. This view opens the door to a pluralist approach to ideal theory, where there is a set of most just options and any one is worth striving for. That revision still leaves the core normative function of ideal theory in place: offering a lofty goal to guide action, while recognizing that many injustices must be overcome on the way to it.

Instead of denying existence of the ideal, (b) doubts that knowing it helps advance justice. Sen takes this view: “[T]he existence of an identifiably inviolate, or best, alternative does not indicate that it is necessary (or indeed useful) to refer to it in judging the relative merits of two other alternatives.”<sup>21</sup> Here he overstates his case by suggesting that ideal theory would be irrelevant to guiding action *even if it gained the information it seeks*. Imagine that ideal theorists succeeded in compiling a volume that accurately details the most just society (or societies) possible. Contrary to Sen, there is reason to believe that such a work would offer insights into how best to advance justice.

The uneven and path-dependent nature of social change helps explain why. If the path to advancing justice were always smooth and gradually ascending – like a hike up Mount Fuji, as Gaus puts it<sup>22</sup> – knowing the ideal would be unnecessary. The option that leads to *greater justice* also would lead toward the *most just* possibility. But the path to advancing justice sometimes is more rugged with peaks and valleys: the option promising greater justice now leads to a local peak but away from the most just possibility. Knowledge of the ideal helps identify and avoid paths that, though appealing in isolation, lead away from and can preclude the ultimate goal.<sup>23</sup>

This strategy of bypassing opportunities to advance justice in the short term to do so in the long term does not mean that ideal theory *always* permits an action as long as it takes society closer to the ideal. Some ideal theories categorically prohibit certain actions – say restrictions of basic liberties – even if they represent the only path to an ideal. In this case, the ideal is morally infeasible. We could achieve it, but not through morally permissible actions. This complication does not undermine the normative value of an ideal goal. It rather adds another layer of feasibility to consider – moral feasibility – when determining the most just society to strive for.

<sup>21</sup> Sen, “What Do We Want from a Theory of Justice?” 222.

<sup>22</sup> Gaus, *The Tyranny of the Ideal*, 62.

<sup>23</sup> Rawls, *The Law of Peoples*, 89–90; A. John Simmons, “Ideal and Nonideal Theory,” *Philosophy and Public Affairs* 38, 1 (2010): 5–36; and Gaus, *The Tyranny of the Ideal*, 61–67.

In many contexts, we readily recognize the value of a long-term goal like ideal theory provides. A medical student who wants to serve the poor as a doctor forgoes volunteering full-time at a homeless shelter, even though doing so would lead to greater justice now. Full-time volunteer work would force the student to abandon their studies and goal of advancing justice as a doctor. Leaders of social movements often make similar judgments. Rather than pursue *every* opportunity to advance justice, they strategically dedicate energy and resources to court cases, legislative campaigns, and protests best suited to advance their long-term goal.<sup>24</sup> Likewise, ideal theory offers an objective to orient action away from paths that diverge from our ultimate goal.

Despite these points in ideal theory's favor, perhaps the ideal is too far off to offer meaningful guidance. Admittedly, ideal theory cannot provide complete guidance on how to act by specifying an ideal – knowing a goal is distinct from knowing how to achieve it. Nevertheless, knowledge of a goal is often informative in evaluating paths to it. If a core principle of the ideal society were a ban on nuclear technology, that would tell us that the goal of banning nuclear weapons and power is not a fool's errand but worth pursuing. Though ideal theory cannot provide complete guidance, it offers information that merits consideration and has potential value in advancing justice.

In sum, valid moral instincts lie behind wanting to identify the most just society possible. Navigational ideal theory seeks information relevant to guiding action in a complex world where the path to greater justice is far from clear. The critical question, which we consider next, is whether ideal theory can attain this information.

#### WHAT A PLAUSIBLE DEFENSE OF IDEAL THEORY REQUIRES

This section identifies a necessary condition for a plausible defense of navigational ideal theory as part of an argument for why, unfortunately, such a defense fails. The argument builds on criticisms of ideal theory by Gaus in his book *Tyranny of the Ideal*. There he raises doubts that we have the information necessary to determine the most just society possible. Since the ideal likely lies far in the future, describing it requires predictions about far-off worlds. Gaus assumes we are less accurate in judgments about justice for the distant future than the present.<sup>25</sup> Though a reasonable claim, not all political

<sup>24</sup> See, e.g., David Garrow, *Bearing the Cross: Martin Luther King, Jr., and the Southern Christian Leadership Conference* (New York: HarperCollins Publishers, 2004).

<sup>25</sup> Gaus, *The Tyranny of the Ideal*, 78.



philosophers share it.<sup>26</sup> Some have more confidence in their predictions and descriptions of distant ideal worlds, which Gaus dismisses as “sheer delusion.”<sup>27</sup>

The disagreement between Gaus and ideal theorists reflects a long-standing split among political philosophers: some are deeply suspicious of ideal and utopian theorizing, while others embrace it. Social science research on prediction can help overcome this impasse by shedding light on the uncertainty inherent in ideal theorizing. Given the sharp divide over ideal theory, there is value in detailing this evidence.

Let’s start with premise (1) of my argument drawing on this evidence:

- (1) Defenses of navigational ideal theory are plausible only if they show that the theory’s principles would have normative force in the society it envisions.

This premise identifies a necessary condition to plausibly defend navigational ideal theory. It is modest in that it does not require a defense to show that an ideal theory’s principles would have *more* normative appeal than all other proposals, leaving open the possibility discussed earlier that the ideal society could take various forms. The basic intuition of premise (1) is that, since ideal theory outlines principles for a *future* society, these principles cannot just have normative force now. They must have normative force in the ideal society. We do not want to embrace an ideal theory, pursue its ideal, arrive at it, and then discover its principles of justice are ill-suited for the society we find ourselves in.

Another way of expressing this point is that ideal theory must consider the social realizations of its principles. Ideal theory involves not just theorizing about principles of justice, but also thinking through what those principles would look like in society when implemented.<sup>28</sup> Many ideal theorists are sensitive to this point, even if they do not explicitly say so. Ideal theories usually include background social conditions – that is, they specify the context in which ideal principles of justice would be implemented. Knowing that context helps in imagining what an ideal theory’s principles would look like when realized and whether they would have normative appeal in that form.

Rawls’s theory of justice, for example, includes background social conditions. Its principles are adopted under “favorable circumstances,”<sup>29</sup> where

<sup>26</sup> Gaus, *The Tyranny of the Ideal*, 102.

<sup>27</sup> Gaus, *The Tyranny of the Ideal*, 106.

<sup>28</sup> Sen, *The Idea of Justice*, 18–22; and Gaus, *The Tyranny of the Ideal*, 23.

<sup>29</sup> Rawls, *A Theory of Justice*, 216.

there is moderate scarcity but conditions are not so harsh as to preclude cooperation or democratic institutions.<sup>30</sup> Rawls also takes pluralism as a given in the ideal society, assuming that individuals hold different religious and philosophical views.<sup>31</sup>

Such background conditions play a key role in Rawls's method of defending ideal theory. In his view, simply reflecting on a theory's principles in the abstract is insufficient for evaluating them. He instead recommends evaluating ideal principles against our judgments at all levels of generality – from abstract conceptions of justice to its demands in concrete situations – in search of a reflective equilibrium, where the principles and all our judgments align. This method does not privilege judgments at one level of generality over others.<sup>32</sup> It thus ensures a place for normative judgments on how an ideal theory's principles would fare in potential circumstances where they would be realized. By adopting this method for evaluating ideal theory, Rawls shows an appreciation for the concerns of premise (1).<sup>33</sup>

Other approaches are less congenial to premise (1). G. A. Cohen, for instance, criticizes Rawls for outlining a theory too closely tied to particular social conditions and recommends instead identifying more general principles of justice.<sup>34</sup> According to this view, a theory of justice need not provide background social conditions. But even for that approach, background social conditions remain relevant. Evaluating principles of justice in isolation, detached from their application in social settings, severely limits our ability to form and be confident in our normative judgments. We develop and refine such judgments by considering the application of principles in concrete contexts – a point Cohen himself recognizes.<sup>35</sup> If ideal theory fails to give background social conditions, evaluating it still involves bringing to mind likely conditions in the ideal society so as to determine what the theory's principles would look like in practice.

A more direct challenge to premise (1) comes from the idea that some principles of justice are so self-evident that no social realization of them – no matter how disastrous – could give us reason to rethink them. Like Sen and Gaus, I find this claim implausible.<sup>36</sup> To deny that *any* social realization could

<sup>30</sup> Rawls, *A Theory of Justice*, 110; and *Justice as Fairness*, 47.

<sup>31</sup> Rawls, *Justice as Fairness*, 84.

<sup>32</sup> Rawls, *A Theory of Justice*, 17–18; and *Justice as Fairness*, 29–32.

<sup>33</sup> See Gaus, *The Tyranny of the Ideal*, 21–22. Concern about the social realizations of principles of justice also comes up, at least in passing, in Robert Nozick's account of ideal theory. He leaves open whether we should reject principles of justice that cause “catastrophic moral horror.” See Nozick, *Anarchy, State, and Utopia* (New York: Basic Books, 1974), 30.

<sup>34</sup> G. A. Cohen, “Facts and Principles,” *Philosophy and Public Affairs* 31, no. 3 (2003): 241–42.

<sup>35</sup> Cohen, “Facts and Principles,” 227.

<sup>36</sup> Sen, *The Idea of Justice*, 21; and Gaus, *The Tyranny of the Ideal*, 26–29.

challenge an ideal theory before even considering it betrays a lack of epistemic humility. It is impossible to know the full implications of principles we propose, so it is important to remain open to encountering considerations that could prompt us to revise them.<sup>37</sup> When moral reasoning closes that possibility and holds on to principles despite their troubling and absurd implications, it becomes rigid dogmatism – like Immanuel Kant’s doubling down on the claim that it’s wrong to lie to a murderer at the door looking for a friend.<sup>38</sup> Ideal theorists can claim that their principles of justice are true even if they cause the world to perish.<sup>39</sup> But they will encounter stiff resistance – and for good reason.

### PREDICTION AND IDEAL THEORY

So considering the social realizations of an ideal theory’s principles is key to evaluating and defending them. Now to the next step of the argument:

- (2) Showing that navigational ideal theory’s principles would have normative force in the society it envisions requires reliably accurate predictions about science, technology, economics, and politics for the distant future.

Given today’s pervasive and entrenched injustices, few expect the ideal society to appear any time soon. Ending society’s most significant injustices requires collective efforts that span generations. Since advancing justice is a long-term project, the potential arrival of the ideal society lies in the distant future. The ideal theorist interested in defending their theory must make predictions about society far into the future to show what their theory’s principles would look like and that they would have normative force under those conditions.

Premise (2) perhaps seems misguided since ideal theory’s purpose is to *prescribe* a goal, not *predict* whether it will be reached. That is right, but prescribing a feasible goal – navigational ideal theory’s focus – requires prediction. Specifically, the ideal theorist must predict what is possible in the future, and from that feasible set identify the most just option.<sup>40</sup>

<sup>37</sup> See Elizabeth Anderson, “Moral Bias and Corrective Practices: A Pragmatist Perspective,” *Proceedings and Addresses of the American Philosophical Association* 89 (2015): 21–47.

<sup>38</sup> Immanuel Kant, “On the Supposed Right to Lie from Philanthropy,” in *Practical Philosophy*, trans. and ed. Mary Gregor (New York: Cambridge University Press, 1996), 605–15.

<sup>39</sup> This phrase is altered from a line, not surprisingly, in Kant’s writings. See Kant, “Toward Perpetual Peace,” in *Practical Philosophy*, trans. and ed. Mary Gregor (New York: Cambridge University Press, 1996), 8:378.

<sup>40</sup> There is much debate over how to define possible or feasible. See Juha Räikkä, “The Feasibility Condition in Political Theory,” *Journal of Political Philosophy* 6, no. 1 (1998):

It is not enough, then, for the ideal theorist to predict what their proposed principles of justice would look like if implemented now. Before the ideal has any hope of arriving, much in society could change that would provide legitimate grounds for reconsidering our principles of justice.<sup>41</sup> Most obviously, society could cease to reflect background economic and political conditions assumed by ideal theory and render the theory's principles obsolete. If Rawls's assumption of moderate scarcity no longer held and new technology brought an overabundance of widely distributed resources, previous debates over distributive justice could look quaint, while other issues moved to the fore. Accurate predictions about future economic and political conditions are necessary to avoid a theory that becomes dated and vulnerable to challenges.

Defending ideal theory also calls for accurate predictions about science and technology. Though we use existing normative principles to evaluate innovations, the interaction between the two proves more complex. Technological and scientific discoveries can raise valid reasons to rethink

27–40; Geoffrey Brennan and Philip Pettit, “The Feasibility Issue,” in *The Oxford Handbook of Contemporary Philosophy*, ed. Frank Jackson and Michael Smith (New York: Oxford University Press, 2007), 258–80; Mark Jensen, “The Limits of Practical Possibility,” *Journal of Political Philosophy* 17, no. 2 (2009): 168–84; David Estlund, “Human Nature and the Limits (if any) of Political Philosophy,” *Philosophy and Public Affairs* 39, no. 3 (2011): 207–37; Pablo Gilabert and Holly Lawford-Smith, “Political Feasibility: A Conceptual Exploration,” *Political Studies* 60, no. 4 (2012): 809–25; Anca Gheaus, “The Feasibility Constraint on the Concept of Justice,” *Philosophical Quarterly* 63, no. 252 (2013): 445–64; Holly Lawford-Smith, “Understanding Political Feasibility,” *Journal of Political Philosophy* 21, no. 3 (2013): 243–59; David Wiens, “‘Going Evaluative’ to Save Justice from Feasibility—a Pyrrhic Victory,” *Philosophical Quarterly* 64, no. 255 (2014): 301–7; David Wiens, “Motivational Limitations on the Demands of Justice,” *European Journal of Political Theory* 15, no. 3 (2016): 333–52; Nicholas Southwood, “Does ‘Ought’ Imply ‘Feasible’?” *Philosophy and Public Affairs* 44, no. 1 (2016): 7–45; Zofia Stemplowska, “Feasibility: Individual and Collective,” *Social Philosophy and Policy* 33, nos. 1–2 (2016): 273–91; Emily McTernan, “Justice, Feasibility, and Social Science as It Is,” *Ethical Theory and Moral Practice* 22, no. 1 (2019): 27–40; and David Estlund, *Utopophobia: On the Limits (if any) of Political Philosophy* (Princeton, NJ: Princeton University Press, 2020). For many issues – from the minimal probability level needed to deem an outcome realistically possible to whether psychological factors like weakness of will represent unchangeable feasibility constraints – there is no consensus and I take no position on them here. Doing so is unnecessary for my argument. What stands out in current debates is how much about future feasibility is unknown. Clearly, future scientific, technological, economic, and political developments will alter what is feasible. The ideal theorist is in no position to make reliably accurate predictions regarding those developments and thus what will be feasible in the future. That limitation poses an insurmountable obstacle for making a plausible defense of navigational ideal theory, as I explain later.

<sup>41</sup> Shmuel Nili raises this point and its complications. See Nili, “The Moving Global Everest: A New Challenge to Global Ideal Theory as a Necessary Compass,” *European Journal of Political Theory* 17, no. 1 (2018): 87–108.

moral judgments.<sup>42</sup> If the ideal theorist fails to account for that, they risk specifying principles that appear just today but would be incomplete and mistaken in the ideal society.

Past innovations illustrate this point by expanding the realm of what's possible and giving rise to new rights and liberties. Consider developments in medicine. Previously, society had few options for healing the sick and injured, and thus a right to healthcare made little sense. Now in a world with effective methods for treating many injuries and diseases, access to such care has come to be seen as a right. People disagree on how extensive this right is, but most accept that wealthy societies have *some* obligation to treat the critically injured, regardless of whether they can pay for care. Likewise, the invention of the printing press contributed to the emergence of a freedom now widely recognized – freedom of the press. Today the Internet is having transformative effects, as it becomes more essential for communication, learning, and political engagement. Some believe access to it should be a basic right.<sup>43</sup> Future innovations will further shape conceptions of justice. The ideal theorist who enumerates basic rights and liberties for the ideal society without considering future innovations risks giving an incomplete list that neglects concerns at the center of tomorrow's debates about justice.

Beyond expanding conceptions of justice, scientific and technological discoveries sometimes unsettle them. Consider advances in our understanding of nonhuman animals. Research has overturned the view, most famously defended by René Descartes, that animals are machines lacking a variety of capacities believed to be distinct to humans.<sup>44</sup> Studies show animals to be creatures far more complex than previously believed, which experience pain and emotion, use language and tools, engage in problem solving, cooperate, and aid others. Such discoveries prompt us to rethink animals' place in the moral universe and suggest that they deserve stronger consideration in theories of justice than they traditionally receive.<sup>45</sup>

<sup>42</sup> See Hans Jonas, "Technology and Responsibility: Reflections on the New Tasks of Ethics," *Social Research* 40, no. 1 (1973): 31–54; and Marianne Boenink, Tsjalling Swierstra, and Dirk Stemerding, "Anticipating the Interaction between Technology and Morality: A Scenario Study of Experimenting with Humans in Biotechnology," *Studies in Ethics, Law, and Technology* 4, no. 2 (2010): Article 4, <https://doi.org/10.2202/1941-6008.1098>.

<sup>43</sup> Merten Reglitz, "The Human Right to Free Internet Access," *Journal of Applied Philosophy* 37, no. 2 (2020): 314–31.

<sup>44</sup> René Descartes, *Discourse on Method*, 3rd. ed., trans. Donald Cress (Indianapolis: Hackett Publishing Company, 1998), 55–59 (page numbers refer to Adam and Tannery edition of Descartes's works).

<sup>45</sup> See Lori Gruen, *Ethics and Animals: An Introduction* (New York: Cambridge University Press, 2011).

In response to these examples, some may still resist premise (2). Rather than show the need for long-term prediction in ideal theory, perhaps these examples show the importance of specifying principles of justice that remain valid across contexts. For instance, instead of making a sharp distinction between humans and animals, a normative theory is better off identifying capacities that merit moral consideration, without predicting in advance what life forms possess them. Ideal theorists, in other words, need more theory.

In a sense that is right. Innovations and discoveries highlight blind spots in normative theories, and it would be best to formulate them free from errors to begin with. The problem, though, is that we often recognize errors in our theories only *after* confronting concrete counterexamples. Beforehand, it is difficult to know the specific refinements and qualifications that principles of justice require. For that reason, the ideal theorist cannot afford to neglect major technological and scientific advancements in a future hospitable to the ideal society, which may be radically different from our own. Given how often new discoveries influence conceptions of justice, we have little reason to assume that proposed principles of ideal justice that look appealing today are suited for future worlds. We also need an accurate account of what those worlds could look like.

#### WHY DEFENSES OF IDEAL THEORY (REGRETTABLY) FAIL

To review, the ideal theorist wanting to plausibly defend their theory must explain what their proposed principles of justice would look like when implemented in the ideal society, which is unlikely to arrive soon. That forces the ideal theorist to make predictions for the distant future about science, technology, economics, and politics, since these factors impact what form ideal principles would take when realized. Unfortunately for the ideal theorist, they cannot accurately make these predictions with any consistency, which is my next claim:

- (3) We cannot make reliably accurate predictions about science, technology, economics, and politics for the distant future.<sup>46</sup>

<sup>46</sup> Karl Popper raises a similar concern. See Popper, *The Open Universe: An Argument for Indeterminism*, ed. W. W. Bartley, III (New York: Routledge, 1992), 68–77; and *The Poverty of Historicism* (New York: Routledge, 2002), xi–xiii. I share Popper’s skepticism, but premise (3) does not rely on his claim that it is *logically* impossible to make such predictions, which may be too strong. See E. Lagerspetz, “Predictability and the Growth of Knowledge,” *Synthese* 141, no. 3 (2004): 445–59.

Those who study prediction overwhelmingly agree on this point. Social science research offers nothing to suggest that we come close to having the predictive capacities necessary to formulate a defensible account of navigational ideal theory.

Philip Tetlock, a leading researcher on prediction, conducts forecasting tournaments to measure how well individuals – including professors, journalists, and intelligence analysts<sup>47</sup> – predict future societal events (e.g., who will win an election or whether two countries will go to war). These studies find that some individuals are better forecasters than others and that certain interventions and ways of thinking improve forecasting.<sup>48</sup> But they also find severe limitations on human predictive capacities. When making predictions five years out, forecasters' accuracy declines and approaches random chance – or, as Tetlock puts it, a dart-throwing chimp.<sup>49</sup> Tetlock and others conclude that even the best forecasters cannot make accurate predictions about society a decade from now, besides the occasional lucky guess and generalities (e.g., there will be interstate conflicts).<sup>50</sup> So if someone makes detailed claims about what society will look like in fifty years, we have little reason to take their arguments seriously.

These limitations reflect how difficult it is to predict outcomes of complex systems involving many variables in nonlinear relationships, as is the case with society. Weather forecasting offers a helpful analogy. Weather patterns are extremely complex. Small variations in initial conditions – beyond what we can accurately measure – lead to vastly different outcomes as time elapses.<sup>51</sup> So though meteorologists generally can predict the weather for the next few days, their longer-term forecasts are far less accurate. The complexity of the social world creates similar challenges. Skilled forecasters predict events in the short term with some accuracy, but their predictive powers fail them when trying to foretell events further out.

Some may accept these limits on human prediction while denying that they pose problems for ideal theory. After all, ideal theory only makes *general* predictions about the distant future (e.g., there will be moderate scarcity and

<sup>47</sup> Tetlock, *Expert Political Judgment*.

<sup>48</sup> Tetlock, *Expert Political Judgment*; Barbara Mellers et al., "The Psychology of Intelligence Analysis: Drivers of Prediction Accuracy in World Politics," *Journal of Experimental Psychology: Applied* 21, no. 1 (2015): 1–14; and Tetlock and Gardner, *Superforecasting*.

<sup>49</sup> Tetlock and Gardner, *Superforecasting*, 4.

<sup>50</sup> Tetlock and Gardner, *Superforecasting*, 243–44.

<sup>51</sup> Edward Lorenz, *The Essence of Chaos* (Seattle, WA: University of Washington Press, 1993), esp. 181–84.

pluralism), not the detailed predictions studied by Tetlock (e.g., the euro will hit this value by this date).

This defense fails to save ideal theory, however. Our inability to predict many small events can add up and lead to dramatic, unforeseen changes in society that are relevant to ideal theory. Relatedly, failures in forecasting sometimes result from the intervention of rare, difficult-to-predict, transformative events in society – what Nassim Taleb calls Black Swans.<sup>52</sup> These events, like the invention of the printing press, reshape society and alter its course in ways hard to foresee. Black Swans are especially relevant to ideal theory because their transformative effects can unsettle our considered views, causing us to rethink conceptions of justice. In other words, these societal developments are exactly the ones ideal theorists must predict to show that their principles of justice would have normative force in a future radically different from today. Yet the same feature that makes Black Swans transformative – radically departing from the status quo – also makes them difficult to predict. Predictions often extrapolate from the past, but that method fails to predict outliers like Black Swans. Such events pose a thorny dilemma for ideal theory: events with great impact on the world, which we desperately would like to predict, often are ones we are least likely to. Black Swans throw a wrench into predictive models, and give us little reason to believe that we can make reliably accurate predictions about society for the distant future.

Occasional predictions of Black Swans fail to provide much hope to ideal theorists. Pundits, academics, and others make many claims about the future – inevitably, some seemingly improbable predictions end up being right by chance. But that doesn't mean their predictions are *reliably* accurate. Forecasting studies put successful predictions, which grab our attention, into context by also tracking failed ones. And as research shows, people are poor predictors of future societal developments beyond the short term.

Such limitations severely hinder ideal theorizing. Since we cannot predict future possibilities for society, we don't know what an ideal theory's principles would look like when implemented, which prevents us from plausibly defending them. Here is a review of the argument so far and the conclusion that follows:

- (1) Defenses of navigational ideal theory are plausible only if they show that the theory's principles would have normative force in the society it envisions.

<sup>52</sup> Taleb, *The Black Swan*.



- (2) Showing that navigational ideal theory's principles would have normative force in the society it envisions requires reliably accurate predictions about science, technology, economics, and politics for the distant future.
- (3) We cannot make reliably accurate predictions about science, technology, economics, and politics for the distant future.
- (4) So, by (2) and (3), we cannot show that navigational ideal theory's principles would have normative force in the society it envisions.
- (5) So, by (1) and (4), no defense of navigational ideal theory is plausible.<sup>53</sup>

There are two points worth noting. First, this argument is not based on radical skepticism about moral truth. I assume we can identify clear examples of unjust societies and make plausible claims about what the ideal society *is not* – for instance, one with the horrors of slavery like the antebellum South. But such claims alone cannot get us to a determinate answer about what the ideal society *specifically is*, especially given the vast array of future possibilities, some of which we would have difficulty even imagining today.

Second, the argument never rejects out of hand the concept of the ideal society. It grants to ideal theorists the possibility of an ideal that represents the most just society with the potential to be realized. The problem, though, is that we lack the predictive capacities to plausibly identify and defend this ideal. That is what makes ideal theorists' plight so frustrating. They believe in the possibility of the ideal, perhaps rightly so, but prove unable to identify it with any certainty.

## IMPLICATIONS FOR POLITICAL PHILOSOPHY

One of the most famous critiques of religious belief comes from Bertrand Russell. He compares such belief to claiming that a tiny teapot, imperceptible by telescope, orbits the sun. Though no one can disprove this claim, it would be nonsense to accuse those who doubt it of being unreasonable. The onus is on those making the claim about the teapot to show its plausibility. If they fail to, we dismiss their claim as absurd. For Russell, religious belief has a similar status: it lacks plausible grounds, even if it cannot be falsified. His analogy emphasizes that those making religious claims cannot expect others to believe simply because it is impossible to disprove their claims.<sup>54</sup>

<sup>53</sup> See the Appendix for an expansion of this argument, which considers and responds to further potential objections.

<sup>54</sup> Bertrand Russell, "Is There a God?" in *The Collected Papers of Bertrand Russell, Vol. 11: Last Philosophical Testament, 1943–68*, ed. John Slater and Peter Köllner (New York: Routledge, 1997), 547–48.

Russell's analogy is relevant because it equally applies to navigational ideal theory. Though ideal theorists like Rawls distinguish ideal theory from religious belief, the former is no more immune from Russell's critique than the latter. The ideal theorist claims to give plausible grounds that a particular ideal should serve as a collective guide. Skeptics may run into difficulties disproving that claim, given deep uncertainty over the distant future. Yet for that same reason, the ideal theorist cannot make a compelling case *for* their ideal – they cannot show that it would have normative force in far-off contexts where it could be realized. Rawls's method aspires to offer an ideal that others in society have reason to accept, but ultimately fails to do so. That raises the question: Where does political philosophy go next?

Some suggest that political philosophy should abandon its focus on ideal theory. Charles Mills argues that, by placing its attention on far-off ideals, ideal theory distracts from today's most pressing injustices and has perverse effects on political philosophy's priorities.<sup>55</sup> Burke Hendrix raises the additional worry that ideal theory risks doing more harm than good by championing ideals whose full consequences cannot be known. Ideal theory often has unintended effects and can exacerbate the very injustices it seeks to remedy.<sup>56</sup> Such criticisms of ideal theory reflect underlying discontent with the dominant approach to justice in political philosophy.

Skeptics of ideal theory offer alternatives, which share the feature of rejecting a single ideal to guide efforts toward greater justice. Sen argues that we can advance justice without a perfect ideal by instead identifying the most pressing injustices and using a comparative approach to evaluate options for addressing them.<sup>57</sup> David Wiens also prefers to focus on specific injustices. He suggests institutional failure analysis, which identifies societal failures resulting in injustice and then formulates feasible measures to avoid them.<sup>58</sup> This interest in addressing injustice rather than striving after an ideal is nothing new and resembles sentiments common after the Second World War, whose horrors dashed utopian hopes.<sup>59</sup> Karl Popper reflects this mindset in a 1947 lecture on

<sup>55</sup> Charles Mills, "‘Ideal Theory’ as Ideology," *Hypatia* 20, no. 3 (2005): 165–84; and *Black Rights/White Wrongs: The Critique of Racial Liberalism* (New York: Oxford University Press, 2017), esp. 139–60, 201–16.

<sup>56</sup> Burke Hendrix, "Where Should We Expect Social Change in Non-ideal Theory?" *Political Theory* 41, no. 1 (2013): 116–43; and *Strategies of Justice: Aboriginal Peoples, Persistent Injustice, and the Ethics of Political Action* (New York: Oxford University Press, 2019).

<sup>57</sup> Sen, *The Idea of Justice*.

<sup>58</sup> David Wiens, "Prescribing Institutions without Ideal Theory," *Journal of Political Philosophy* 20, no. 1 (2012): 45–70.

<sup>59</sup> See Judith Shklar, *After Utopia: The Decline of Political Faith* (Princeton, NJ: Princeton University Press, 1957).

the ills of utopian projects, which recommends advancing justice through “the elimination of concrete evils rather than ... the realization of abstract goods.”<sup>60</sup> Popper’s influence today is most evident in Gaus’s work to develop an alternative to ideal theory. Drawing on Popper’s idea of the Open Society,<sup>61</sup> as well as John Stuart Mill’s idea of experiments in living,<sup>62</sup> Gaus argues that society should encourage its members to pursue *different* ideals and that such experimentation gives insight into what social arrangements best promote justice.<sup>63</sup>

These proposals deserve consideration as potential paths forward in theorizing about justice without navigational ideal theory. It is important, though, to recognize the aspirations that these alternatives leave behind. Without an ideal to guide action, efforts to advance justice face deep uncertainty. Even when we advance justice, there always is the worry that our efforts lead away from greater justice later. Navigational ideal theory looks to be an antidote to such uncertainty by assuring us that we’re on the right path when reforms move society closer to the ideal. Recent proposals by Sen, Wiens, and Gaus offer no such assurance. They leave open the very danger that ideal theory seeks to avoid: ending up on lesser peaks of justice because there is no end goal pointing to the highest peak. They fail to solve the problem at the heart of ideal theory – what ultimate aim should we strive for? In fact, they give up trying to solve it.<sup>64</sup>

I note this limitation not to recommend that political philosophy stubbornly defend the claims of navigational ideal theory. As we have seen, it is in no position to offer a plausible goal to guide action given future uncertainty. But recognizing this limitation helps avoid unrealistic hopes for alternatives to ideal theory. In particular, it is a mistake to claim there are reasonable grounds for believing that these alternatives lead to a society that is in any sense ideal. Despite his criticisms of ideal theory, Gaus makes this mistake when discussing his hopes for an open and diverse society. He writes: “[W]e cannot know

<sup>60</sup> Karl Popper, “Utopia and Violence,” in *Conjectures and Refutations: The Growth of Scientific Knowledge* (New York: Routledge, 2002), 485.

<sup>61</sup> Karl Popper, *The Open Society and Its Enemies* (Princeton, NJ: Princeton University Press, 2013).

<sup>62</sup> John Stuart Mill, *On Liberty*, in *On Liberty and Other Essays*, ed. John Gray (New York: Oxford University Press, 1991), 63. For more on this idea in Mill, see Elizabeth Anderson, “John Stuart Mill and Experiments in Living,” *Ethics* 102, no. 1 (1991): 4–26; and Ryan Muldoon, “Expanding the Justificatory Framework of Mill’s Experiments in Living,” *Utilitas* 27, no. 2 (2015): 179–94.

<sup>63</sup> Gaus, *The Tyranny of the Ideal*.

<sup>64</sup> Wiens explicitly concedes this point. See Wiens, “Political Ideals and the Feasibility Frontier,” 472.

what such an ideal [society] would be – unless we disagree about it. Only those in a morally heterogeneous society have a reasonable hope of actually understanding what an ideal society would be like, but in such a society we will never be collectively devoted to any single ideal.”<sup>65</sup> Gaus suggests that understanding the ideal comes through an indirect process where different individuals and groups seek their own ideal and learn from social experimentation.

Gaus’s optimism is understandable. Experimentation has been the engine behind remarkable advances, transforming fields like medicine, which cures a host of ailments it was impotent against not long ago.<sup>66</sup> Perhaps Gaus is right, and applying this approach to justice will unleash similar advances. But even if he is right, suggesting that this approach leads to the ideal overlooks the often haphazard and imperfect nature of experimentation. Chance and timing impact how knowledge from experimentation grows. Successful experiments spark interest in a hypothesis, the study of which then enjoys disproportionate attention and resources. Yet after many years, we sometimes discover that the hypothesis was wrong. Knowledge eventually grows, but along the way some experimental results direct our attention away from more promising ideas that go neglected. Far from always triumphant, the experimental approach also leads society down paths that are less than ideal.

So abandoning the aspirations of navigational ideal theory comes with real losses. Political philosophy finds itself in a tough spot without any clear ideal to light the way regarding which path best advances justice long term. We are left stumbling about in the dark, with political philosophy unable to allay doubts that actions taken to advance justice may in fact lead away from the most just possibility.

That uncertainty can create motivational hurdles to engaging in the difficult work of advancing justice. For many, a critical component of such work is hope – that the future is not condemned to the same injustices plaguing the present. Alternatives to ideal theory can foster hope in short-term progress by identifying clear injustices, outlining ways to address them, and encouraging people to take action that will bring about marginal advances in justice. But it is unclear that such hope is always enough. Steps to advance justice today can be overturned tomorrow as administrations, lawmakers, and judges change. That reality renders hope in short-term progress fragile and fleeting. Even when marginal advances endure, we cannot know if they represent steps toward the most just society. Together, these factors undermine short-term

<sup>65</sup> Gaus, *The Tyranny of the Ideal*, xix.

<sup>66</sup> See Druin Burch, *Taking the Medicine: A Short History of Medicine’s Beautiful Idea, and Our Difficulty Swallowing It* (London: Random House, 2010).

hope's potency as a source of motivation. If short-term hope is all we have, the arduous task of advancing justice risks resembling a random walk rather than a journey up a majestic peak.

That image of a random walk is far less inspiring and, for some, deeply unsettling. The difficult work of advancing justice entails sacrifices, setbacks, and frustration. Understandably, many look for reassurance that these struggles are worth it – that they lead to a goal worth striving for. Ideal theory offers that by infusing current hardships with moral significance and linking them to a far more hopeful future. So there are real worries that the loss of ideal theory leads to despair and, as a result, some resist simply leaving it behind. Whether ideal theory still has a role to play, despite its limitations, is what we explore next.

### PRESERVING UTOPIAN HOPE

Defenses of navigational ideal theory fall short. But does that failure of ideal theory force us to abandon utopian hope? Even if there are not plausible grounds to accept a proposed ideal, one could accept it on faith. That point suggests a path forward for ideal theory, albeit with tempered ambitions: concede our inability to identify the most just society possible with any confidence, yet embrace hope for an ideal on faith. In the absence of strong evidence for an ideal, faith sustains hope in it. Faith has a close relation to hope, a point that the Christian tradition has long recognized: “[F]aith is the assurance of things hoped for, the conviction of things not seen” (Hebrews 11:1).<sup>67</sup> Guided by that principle, one strategy for preserving ideal theory and utopian hope is to recognize their reliance on faith.

This approach makes explicit ideal theory's parallels with religious belief. The apocalyptic tradition, in particular, envisions a future ideal society and fosters hope that it will be realized. In a pluralistic society, it would be unreasonable to expect everyone to accept these religious beliefs based on faith. Still, for those who have such faith, it is a source of meaning and instills hope for greater justice in a world marred by injustice. Ideal theory has the potential to play a similar role. Just as individuals in pluralistic societies practice different religions, they also can embrace different ideal theories (or none at all). According to this view, no ideal serves as a collective goal for society to pursue. Rather, individuals embrace different ideals, which help assure them that their efforts to advance justice are meaningful steps toward a more perfect world.

<sup>67</sup> New Revised Standard Version.

Some will see little point in hanging on to ideal theory and reach a conclusion similar to Russell's regarding religious belief. Beyond just claiming that arguments for religious belief lack plausible grounds, Russell treats such beliefs as nonsense – like believing a tiny teapot is orbiting the sun. Admitting religious belief's reliance on faith doesn't change that fact. Belief in ideal theory is vulnerable to the same criticism. From the critic's perspective, we should treat any account of ideal theory as absurd, since so much about the future is unknown and the forms society could take are virtually endless. We have no idea what a proposed ideal would look like in practice and whether it would have normative appeal under conditions that could be radically different from today's world. Even if we avoid defending ideal theory, belief in it is nonsense.

One can arrive at this conclusion while still caring about justice. For some, progress against concrete injustices and suffering provides sufficient assurance that their efforts are worth it. To continue in this work, they do not need the further assurance that their efforts move society toward the *most just* possibility.

Others, though, yearn for this more robust hope. That is evident from its persistent expressions in religion, philosophy, literature, and popular culture. Utopian hope expresses the desire to realize an ideally just society and belief in its possibility. Such hope often looks beyond the immediate future for inspiration, an idea evident in Martin Luther King Jr.'s maxim that "the arc of the moral universe is long, but it bends toward justice."<sup>68</sup> Even if the present is not hospitable to justice, the long arc of history is. Utopian hope instills partial, imperfect steps toward justice with meaning by situating them within a longer development toward the ideal. According to this view, human efforts over time move society closer to the ideal, which once achieved will be stable and lasting – otherwise it wouldn't be a true utopia. After all, a society that quickly falls into decline after achieving its goal fails to count as utopia. Because it is meant to endure, utopia represents an end goal. For this reason, utopian hope has close links to teleological views of history, which understand history as having a purpose and moving toward a particular end. Even if there are at times setbacks, the overall course of history is moving toward utopia – or at least that is the hope – and this future ideal informs the significance of all that comes before it.

<sup>68</sup> Martin Luther King Jr., "The Current Crisis in Race Relations," in *A Testament of Hope: The Essential Writings and Speeches of Martin Luther King, Jr.*, ed. James Washington (New York: HarperCollins Publishers, 1986), 88. For a qualified defense of this view, based on the idea that injustice is inherently unstable, see Joshua Cohen, "The Arc of the Moral Universe," *Philosophy and Public Affairs* 26, no. 2 (1997): 91–134.

The goal here is not to argue that morality or reason demands utopian hope. In a world long filled with injustice, where horrific suffering seems to fall at random on the undeserving, it is both understandable and defensible to reject utopian hope. There is plenty of room in political philosophy and social movements for those who work against injustice without hoping for utopia. My modest goal is to carve out space for utopian hope and explain how someone can embrace it without falling into error.

It turns out that the same factor that undermines defenses of ideal theory – future uncertainty – ensures a place for utopian hope. The nature of hope helps explain why. Though conceptions of hope vary, most understand it as involving, at the very least, the desire for an outcome believed to be neither guaranteed nor impossible.<sup>69</sup> The desired outcome need not be likely. After all, people often hope for very unlikely things, like an experimental drug that will cure their cancer. The desired outcome just has to be *possible*. As Adrienne Martin explains in her study *How We Hope*, the mere possibility of an outcome, no matter how unlikely, provides permission to hope and act on that hope. So hope does not require inflating the odds of an outcome, and there is nothing inherently irrational about “hoping against hope” – that is, hoping for an outcome with extremely long odds. Such hope emphasizes an outcome’s possibility, which can have practical value by sustaining individuals as they pursue goals under incredibly trying circumstances, like a terminal illness.<sup>70</sup>

We can apply these insights to utopian hope. The deep uncertainty surrounding the future opens the door for such hope. Future uncertainty makes it impossible to establish what the utopian society would look like and whether it will be realized, so utopian hope must rely on faith. Yet this same uncertainty functions as a bulwark to protect faith in utopian hope. Because so much about the future is uncertain, we cannot entirely preclude the possibility of achieving the ideal society at some point. Though realizing the ideal is difficult to imagine, it still is possible, which frees people to embrace utopian hope without committing any obvious error. This hope offers needed assurance, at least for some, that the arduous work to make the world more just leads

<sup>69</sup> See R. S. Downie, “Hope,” *Philosophy and Phenomenological Research* 24, no. 2 (1963): 248–49; J. P. Day, “Hope,” *American Philosophical Quarterly* 6, no. 2 (1969): 89; Luc Bovens, “The Value of Hope,” *Philosophy and Phenomenological Research* 59, no. 3 (1999): 673; Philip Pettit, “Hope and Its Place in Mind,” *Annals of the American Academy of Political and Social Science* 592 (2004): 154; Ariel Meirav, “The Nature of Hope,” *Ratio* 22, no. 2 (2009): 218–20; and Adrienne Martin, *How We Hope: A Moral Psychology* (Princeton, NJ: Princeton University Press, 2014), 62.

<sup>70</sup> Martin, *How We Hope*, 11–71.

to more than just fleeting progress. From the perspective of utopian hope, such work – so often incomplete and imperfect – contributes to an ideal truly worth striving for.

Here some might object that the analogy between ideal theory and religious faith breaks down. This description of utopian hope, aware of its own uncertainty, seems to stand in contrast to more dogmatic forms of hope often found in religious faith. Indeed, it is common for religious believers to describe their faith as providing hope for an ideal future that is certain.<sup>71</sup>

But that view is not universal. In today's secular age, as Charles Taylor points out, religious beliefs and hopes no longer seem as self-evident as they once did. Religious belief in many contemporary societies does not enjoy the status of assumed truth, nor does the divine pervade shared perceptions of the world and the forces within it. For those who choose to embrace religious faith in this context, their faith often coexists with uncertainty and doubt.<sup>72</sup>

This variety of religious faith described by Taylor, which offers uncertain hope, serves as an apt analogy for ideal theory aware of its epistemic limitations. In both cases, beliefs grounded in faith rather than plausible evidence serve as a source of utopian hope. Ideal theory and utopian hope cannot escape this shortcoming, but can persist in spite of it. If we are to hang on to ideal theory and utopian hope with intellectual honesty, we must abandon the ambition of offering plausible grounds for others to accept our ideal and hope for it.

### RAWLS'S FAITH

The idea that ideal theory relies on faith can be jarring, since it is not usually described in this way. Ideal theory often aims to provide a common goal to strive for in a pluralistic society where religious faith fails to fulfill that role. Political philosophy thus tends to treat ideal theory as resting on more solid ground than religious faith. That perspective has its roots in Rawls, whose thought inspired much of contemporary ideal theory. A close look at Rawls's writings serves as a reminder, however, that reliance on faith is at the heart of ideal theory.

In his account of ideal theory, Rawls assigns a central role for utopian hope. No historical figure impacted his thinking on this issue more than Kant.

<sup>71</sup> David Elliot, *Hope and Christian Ethics* (New York: Cambridge University Press, 2017), 72.

<sup>72</sup> Charles Taylor, *A Secular Age* (Cambridge, MA: Harvard University Press, 2007). See also David Newheiser, *Hope in a Secular Age: Deconstruction, Negative Theology, and the Future of Faith* (New York: Cambridge University Press, 2019).



Citing Kant, Rawls stresses the need to hold on to hope for a just society if life in this world is to be worth living.<sup>73</sup> This hope strikes him as necessary, especially when individuals take on the difficult task of working to advance justice. Rawls believes ideal theory can “banish the dangers of resignation and cynicism” and meet the challenge of preserving utopian hope. “By showing how the social world may realize the features of a realistic utopia,” he writes, “political philosophy provides a long-term goal of political endeavor, and in working toward it gives meaning to what we can do today.”<sup>74</sup>

Rawls’s *Lectures on the History of Moral Philosophy* illustrate in greater detail his interest in the utopian element in Kant’s philosophy, as well as his debt to it.<sup>75</sup> His final lecture on Kant focuses on the relation among faith, reason, and hope in Kant’s thought.<sup>76</sup> As Rawls explains, Kant treats utopian hope as a necessary component of morality. Specifically, the moral law takes as its object the highest good, which for Kant is a world where happiness is proportional to and in harmony with virtue. And since the moral law only seeks ends that are possible, one of its presuppositions is that the highest good must be possible.<sup>77</sup> Without that presupposition, the moral law cannot get off the ground. “If . . . the highest good is impossible,” writes Kant, “then the moral law, which commands us to promote it, must be fantastic and directed to empty imaginary ends and must therefore in itself be false.”<sup>78</sup>

<sup>73</sup> Rawls, *Political Liberalism*, lx; and *The Law of Peoples*, 128.

<sup>74</sup> Rawls, *The Law of Peoples*, 128.

<sup>75</sup> In his *Lectures*, Rawls notes that his purpose for studying Kant and other thinkers is to “bring out what is distinctive in their approach to moral philosophy” rather than find “some philosophical argument, some analytic idea that will be directly useful for our present-day philosophical questions.” See Rawls, *Lectures on the History of Moral Philosophy*, ed. Barbara Herman (Cambridge, MA: Harvard University Press, 2000), 329. We should keep that caveat in mind, but it does not mean that the *Lectures* are irrelevant to understanding Rawls’s normative theory. In the case of Kant, Rawls explicitly and favorably cites him in his major works of political philosophy – in particular, when discussing the need to hope for a just future. The *Lectures* prove valuable in gathering a fuller picture of how Rawls understands this idea from Kant and what about it appeals to him.

<sup>76</sup> Rawls, *Lectures on the History of Moral Philosophy*, 309–25.

<sup>77</sup> Kant, *Critique of Practical Reason*, in *Practical Philosophy*, trans. and ed. Mary Gregor (New York: Cambridge University Press, 1996), 5:108–14; *Religion within the Boundaries of Mere Reason*, in *Religion and Rational Theology*, trans. and ed. Allen Wood and George di Giovanni (New York: Cambridge University Press, 1996), 6:3–6; and “On the Common Saying: That May Be Correct in Theory, But It Is of No Use in Practice,” in *Practical Philosophy*, trans. and ed. Mary Gregor (New York: Cambridge University Press, 1996), 8:279–80. For more on this idea in Kant, see Loren Goldman, “In Defense of Blinders: On Kant, Political Hope, and the Need for Practical Belief,” *Political Theory* 40, no. 4 (2012): 497–523.

<sup>78</sup> Kant, *Critique of Practical Reason*, 5:114.

Kant goes on to argue that further presuppositions are necessary for the highest good to be possible. It is here that he introduces tenets of religious faith – immortality of the soul and the existence of an omnipotent and benevolent God – along with freedom of will as necessary presuppositions of morality.<sup>79</sup> Without these presuppositions, the highest good as Kant sees it would be impossible. Only a good and all-powerful God ensures that virtue and happiness ultimately will correspond. Only immortality ensures that individuals can make continual progress in conforming their will to the moral law. And only freedom of will ensures the possibility of moral action to begin with. Together, the presuppositions of morality answer a question Kant poses to himself: “What may I hope for?”<sup>80</sup> They fill in the content of what he considers reasonable faith.<sup>81</sup> Kant avoids claiming that we can prove God’s existence or immortality. But there is also no way to disprove these religious beliefs, so reason *permits* them. Moreover, since these beliefs are necessary presuppositions of the moral law, morality *requires* them.<sup>82</sup> As Kant puts it, “Morality . . . inevitably leads to religion.”<sup>83</sup>

Rawls’s conception of utopian hope never looks as robust as Kant’s, since it lacks hope for God and immortality. But despite these departures, Rawls treats Kant’s account of utopian hope and reasonable faith with great sympathy. To distinguish between which elements are worth preserving and which to discard, Rawls uses the German term “*Vernunftglaube*” for some of Kant’s beliefs and the English term “reasonable faith” for others. *Vernunftglaube* just means reasonable faith, but Rawls specifically uses it to refer to Kant’s hope for achieving the highest good and the related presuppositions of God and immortality. Rawls then uses the term reasonable faith to refer to Kant’s hope for a “realm of ends,” a society where individuals live under conditions of justice.<sup>84</sup> This realm of ends closely resembles the realistic utopia that Rawls outlines in his ideal theory and defends as a goal to strive for. Notably, he calls the “realm of ends . . . a secular ideal.”<sup>85</sup>

Not surprisingly, Rawls sees the reasonable faith required for this ideal as more essential than Kant’s *Vernunftglaube*. He asks: “[W]hat is the content of

<sup>79</sup> Kant, *Critique of Practical Reason*, 5:122–34.

<sup>80</sup> Kant poses this question in his letter to C. F. Staudlin from May 4, 1793. See George di Giovanni, “Translator’s Introduction,” in *Religion and Rational Theology*, trans. and ed. Allen Wood and George di Giovanni (New York: Cambridge University Press, 1996), 49.

<sup>81</sup> Notably, hope for the kingdom of God is part of Kant’s conception of reasonable faith. See Kant, *Religion within the Boundaries of Mere Reason*, 6:134–36.

<sup>82</sup> Kant, *Critique of Practical Reason*, 5:122–34.

<sup>83</sup> Kant, *Religion within the Boundaries of Mere Reason*, 6:6.

<sup>84</sup> Rawls, *Lectures on the History of Moral Philosophy*, 310.

<sup>85</sup> Rawls, *Lectures on the History of Moral Philosophy*, 312.

practical faith once we take a realm of ends as the object of the moral law?" Here Rawls essentially describes the move that he makes in his political philosophy, so his response to the question proves illuminating for understanding faith's role in his ideal theory. He writes:

I suggest that while [taking a realm of ends as the object of the moral law] does not require the postulates of God and immortality, it does require certain beliefs about our nature and the social world . . . . For we can believe that a realm of ends is possible in the world only if the order of nature and social necessities are not unfriendly to that ideal. For this to be so, it must contain forces and tendencies that in the longer run tend to bring out, or at least support, such a realm and to educate mankind so as to further this end.<sup>86</sup>

For Rawls, Kant's religious beliefs are unnecessary as presuppositions for working toward an ideal society, but other presuppositions *are* necessary. Namely, we must hold on to the hope that such an ideal is possible and, relatedly, that the future is hospitable to its realization.

This hope for the future reflects aspects of a teleological view of history. Kant stresses our practical need to "hope for better times" and see history as progressing toward greater perfection.<sup>87</sup> Rawls's remarks on Kant's view suggest an affinity for it: "We must believe . . . that the course of human history is progressively improving, and not becoming worse, or that it does not fluctuate in perpetuity from bad to good and from good to bad. For in this case we will view the spectacle of human history as a farce that arouses loathing of our species."<sup>88</sup> Even as he leaves behind Kant's religious beliefs, Rawls expresses greater openness to other leaps of faith seen as necessary presuppositions of working toward the ideal society.<sup>89</sup>

The essay "On My Religion," written by Rawls near the end of his life, provides insight into why he cannot hang on to the religious elements in Kant's concept of reasonable faith, which he otherwise finds appealing. Rawls points to three events during his service in the Second World War that led him to

<sup>86</sup> Rawls, *Lectures on the History of Moral Philosophy*, 319.

<sup>87</sup> Kant, "On the Common Saying: That May Be Correct in Theory, But It Is of No Use in Practice," 8:309–10.

<sup>88</sup> Rawls, *Lectures on the History of Moral Philosophy*, 319–20.

<sup>89</sup> For clarity, it is helpful to distinguish Rawls's views on different forms of teleology. Rawls rejects teleological theories of ethics, which prioritize the good over the right, in favor of his conception of justice as fairness. See Rawls, *A Theory of Justice*, 21–30. Teleological theories of ethics are distinct from teleological views of history, which understand history as moving toward a particular end. Rawls's remarks on Kant and hope for a future hospitable to justice suggest a greater sympathy toward a teleological view of history. Specifically, he *hopes* that such a conception of history, in which society continually progresses toward ideal justice, is true.

abandon his Christian faith: (1) a sermon to him and other U.S. troops claiming that God would aid them in killing the Japanese; (2) the death of his friend Deacon on an expedition that Rawls almost went on instead; and (3) learning about the horrors of the Holocaust.<sup>90</sup> It became impossible to maintain his faith, as Rawls explains:

These incidents, and especially the third [the Holocaust] . . . , affected me in the same way. This took the form of questioning whether prayer was possible. How could I pray and ask God to help me, or my family, or my country, or any other cherished thing I cared about, when God would not save millions of Jews from Hitler? When Lincoln interprets the Civil War as God's punishment for the sin of slavery, deserved equally by North and South, God is seen as acting justly. But the Holocaust can't be interpreted in that way, and all attempts to do so that I have read of are hideous and evil. To interpret history as expressing God's will, God's will must accord with the most basic ideas of justice as we know them. For what else can the most basic justice be? Thus, I soon came to reject the idea of the supremacy of the divine will as also hideous and evil.<sup>91</sup>

Though deeply religious before the war – Rawls at one point considered seminary<sup>92</sup> – the Holocaust and his experiences in combat dashed his faith that God ensures justice now or ever. In fact, for Rawls, to imagine an omnipotent God in a world marred by such evil only crushes hope in a just future, for it suggests that the ruler of the universe is a monster.

The horrors of the Second World War pushed Rawls to give up his religious faith, yet he held on to utopian hope and the conviction that the future is hospitable to justice. Not everyone held on to such hope in the wake of these horrors. Rawls, though, adamantly rejects political despair as an option. Even after the “the manic evil of the Holocaust,” he stresses that we must start from the assumption that a realistic utopia is possible.<sup>93</sup> The thought of abandoning hope in that ideal strikes Rawls as intolerable. So a secular understanding of ideal theory steps in to be the source of utopian hope that religious faith can no longer provide.

The basis for this move lies in his distinction between forms of reasonable faith. In his lecture on Kant, Rawls makes a point to differentiate reasonable faith in an ideal society from reasonable faith (or *Vernunftglaube*) in religious

<sup>90</sup> Rawls, *A Brief Inquiry into the Meaning of Sin and Faith with “On My Religion,”* ed. Thomas Nagel (Cambridge, MA: Harvard University Press, 2009), 262–63.

<sup>91</sup> Rawls, *A Brief Inquiry into the Meaning of Sin and Faith with “On My Religion,”* 263.

<sup>92</sup> Rawls, *A Brief Inquiry into the Meaning of Sin and Faith with “On My Religion,”* 261.

<sup>93</sup> Rawls, *Political Liberalism*, lx.

beliefs like immortality and the existence of God. Rawls admits that Kant himself never makes this distinction,<sup>94</sup> but feels compelled to make it because “the plausibility of Kant’s view in these two cases are quite different.”<sup>95</sup> For Rawls, reasonable faith in a just and ideal society has greater plausibility since the object of such faith could be realized within the confines of the natural world, without any necessary reference to the supernatural.<sup>96</sup> So though hope for a future ideal society and hope in the divine both rest on faith, Rawls sees the former as the more reasonable of the two faiths.

This desire to distinguish the faith required for ideal theory as more reasonable and plausible than the faith required for religious belief is where Rawls runs into trouble. To preserve his utopian hope, Rawls must make leaps of faith no less considerable than those he eschews. His ideal theory rests on the convictions that we can identify the ideal and that the future is such that it fosters progress toward it. We already have seen earlier in this chapter that there are not plausible grounds to support the first claim. And the teleological view of history implicit in the second claim long has been criticized as wishful thinking rather than grounded in actual evidence. That seems even truer today as the world grapples with the legacies of slavery and colonialism, as well as the devastating consequences of human-induced climate change. Confronted with such overwhelming injustice from the present and recent past, some have no patience for “fairy tales that imply some irrepressible justice.”<sup>97</sup> According to that view, only a tendentious and selective reading of history suggests that it continually progresses toward greater justice.

Now it would be too strong to say that Rawls errs in hoping for utopia and a future conducive to its realization. The future *could* radically depart from the present and bring about the ideal. That possibility, regardless of its likelihood, allows one to embrace ideal theory and utopian hope without violating basic principles of rationality. But it is a mistake to think that ideal theory rests on faith so qualitatively different than religious faith that it succeeds in providing plausible grounds to accept and pursue a shared ideal. By making distinctions between ideal theory and religious faith that fail to hold up under scrutiny, Rawls set contemporary ideal theory on the wrong track from the start.

Political philosophy long has taken seriously the notion that ideal theory can identify the most just society possible, give plausible arguments for this claim, and offer an end goal to guide collective action. Such ambitions far

<sup>94</sup> Rawls, *Lectures on the History of Moral Philosophy*, 317.

<sup>95</sup> Rawls, *Lectures on the History of Moral Philosophy*, 310.

<sup>96</sup> Rawls, *Lectures on the History of Moral Philosophy*, 312.

<sup>97</sup> Ta-Nehisi Coates, *Between the World and Me* (New York: Spiegel & Grau, 2015), 70.

outstrip human capacities and fail to appreciate the deep uncertainty in the world along with its full implications. We cannot predict future developments and innovations that will shape society, and thus cannot show what proposed principles of justice would look like in far-off worlds, let alone that they would constitute the most just society. Political philosophers would be wise to admit these limitations and stop expecting the impossible from ideal theory. If theorizing about the most just society possible persists, it is important to recognize it for what it is: hope for an ideal grounded in faith.