The Case for Ideal Theory

Laura Valentini
London School of Economics and Political Science

In Oxford Handbook of International Political Theory
For citations, please refer to the published version.

1. Introduction
In recent years, political theorists have devoted increasing attention to questions of methodology. The debate on so-called “ideal vs non-ideal theory” is at the heart of this methodological turn, and prominently features controversies about international political morality. Specifically, some of the most ambitious views in international political theory are routinely criticized for being too “ideal” or “utopian.” Targets of these criticisms include: cosmopolitan theories defending global distributive equality, theories of global democracy, and so-called “revisionist” just-war theory.

My aim in this chapter is to evaluate whether and, if so, under what conditions, the critique of excessive idealism succeeds. I argue that success is rare. I proceed as follows. In Section 2, I clarify terminology and distinguish between two reasons why a theory may be labelled “ideal”: either because it contains idealizations or because it is insensitive to feasibility constraints. In Section 3, I address objections raised against international political theories that contain idealizations. I show that whether the presence of idealizations invalidates a given theory depends on the role those idealizations play within the theory as a whole. In Section 4, I address objections raised against theories that (allegedly) fail to take sufficient account of feasibility constraints. I argue that those objections only threaten normative theories that claim to deliver practically efficacious guidance. Purely evaluative theories, and normative theories that do not aspire to practical efficacy, are unaffected by this “feasibility” critique. I thus conclude, in Section 5, that there is nothing wrong with ideal theorizing as such, but only with certain misuses of it. Throughout, I illustrate these broader methodological points—which I have partly defended elsewhere¹—by reference to examples drawn from some of the most prominent areas of international political theory: global justice, global democratic theory, and just-war theory.

2. What Is an Ideal Theory?
One often hears political theorists refer to “ideal” and “nonideal” theory. Yet what the labels “ideal” and “non-ideal” stand for is far from clear (for overviews, see Stemplowska 2008; Valentini 2012; Stemplowska and Swift 2012). What is more, the very notion of “a theory” is ubiquitously, yet loosely, invoked.

First, by “a theory” I mean a set of propositions, entailed by principles, that play a given functional role in an agent’s understanding of reality (List and Valentini 2016). Depending on the relevant functional role, we can distinguish between positive and normative/evaluative theories. A theory is normative when its propositions include deontic operators (such as “ought,” “must,” “may”), and it is evaluative when its propositions include evaluative predicates (such as “good,” “bad,” “just,”

¹ I am grateful to Chris Brown, Robyn Eckersley, Seth Lazar, Pietro Maffettone and David Wiens for their helpful comments.
² This chapter offers an application of ideas developed in Valentini (2009; 2016) and in List and Valentini (2016) to the domain of international political theory in particular.
“unjust”). Otherwise, it is positive. Biological and physical theories—e.g., Darwin’s theory of evolution or Einstein’s theory of relativity—are positive. Theories of international justice and of the just war, by contrast, belong to the normative or evaluative domain.

In this contribution, I focus on normative and/or evaluative theories that may be described as “ideal.” The distinction between ideal and nonideal theory was first introduced by John Rawls (1999a, 8, 215; cf. Simmons 2010). An ideal theory, according to Rawls’s definition, is a theory designed assuming “full compliance”—i.e., all agents act as they ought to—and “favourable conditions”—i.e., circumstances are such that individuals’ basic rights and needs can be met (Rawls 1999a, 475–76). A non-ideal theory, by contrast, is a theory characterized by partial compliance, and/or the lack of favourable conditions. That said, in the recent literature, the notion of ideal theory has taken on several, and rather broader, understandings. Two of them are particularly prominent, which I discuss below.

On one understanding, normative and/or evaluative theories are “ideal” to the extent that they involve idealizations. Idealizations, in Onora O’Neill’s (1996, 41) well-known characterization, are falsities introduced for theoretical purposes. Full compliance and favourable conditions are just two examples of idealization: it is plainly false that, in our world, everyone complies with the demands of justice, and that favourable conditions always obtain. Other idealizations include statements such as: “there is no disagreement about justice across different societies”; “soldiers in war face no uncertainty”; “society is a closed system”; “individuals are always altruistically motivated,” and so forth.

A second understanding of ideal theory has developed in connection with debates about so-called feasibility constraints (e.g., Räikkä 1998; Gilabert and Lawford-Smith 2012; Lawford-Smith 2013; Wiens 2015b). From this perspective, the more insensitive to feasibility constraints a theory is, the greater its level of idealism. Feasibility constraints are those physical, biological, political, psychological, and economic facts that make the prescriptions of a normative theory hard, if not impossible, to realize. Theories of international political morality that prescribe considerable global redistribution, demand global democratic participation, or make the permissibility of killing in war dependent on individuals’ moral liability are often regarded as “ideal” in this sense—since their prescriptions appear to be either impossible or at any rate extremely hard to fulfil.

For the sake of completeness, I should mention that there is also a third understanding of ideal theory in the current literature. This is ideal theory understood as offering an end-point towards which we should aspire. This understanding of ideal theory is contrasted with transitional theories of justice, detailing how to make progress justice-wise, without necessarily achieving the ideal (Sen 2009; Wiens 2015a; 2012). Due to space constraints, I set this third understanding aside in the present chapter, and focus on the previous two.

It has become quite common to object to (international) political theories by virtue of their being ideal in either of the two senses I focus on—which involve idealization and insensitivity to feasibility constraints, respectively. Behind these critiques stands the following general concern. A key aim of (international) political theory is to help us address pressing normative questions that arise in today’s international realm: wars, refugee crises, global poverty, and much else. Ideal theories, however, are doomed to fail in this respect. At best, they are useless; at worst, their application to current (international) affairs will deliver bad outcomes.
(see, e.g., Mills 2005; Farrelly 2007; McCarthy 2004; Sen 2009; Wiens 2015a; Miller 2013; cf. Valentini 2009). In what follows, I examine the validity of this critique in relation to (i) theories of international political morality that contain idealizations and (ii) theories that disregard feasibility constraints.

3. Responding to the Charge of Idealization

Several normative/evaluative theories of international politics contain idealizations. Here I offer three prominent examples.

- **Cosmopolitan theories of global distributive justice** that recommend applying the outcome of Rawls’s original-position thought experiment to the world at large (see, e.g., Beitz 1999; Pogge 1989). To the extent that Rawls’s thought-experiment contains idealizations (not only about full compliance and favourable conditions, but also in the characterizations of the parties behind the veil of ignorance), those idealizations are carried forward in cosmopolitan theories that globalize its conclusions.

- **Rawls’s statist theory of international justice**, outlined in The Law of Peoples, contains several idealizations, including its assumption that societies are relatively independent units, distinguished into five categories: liberal, decent, burdened, outlaws states and benevolent absolutisms (Rawls 1999b). The theory, in other words, assumes away much of the complexity of contemporary international politics.

- **Revisionist theories of the just war**, which are designed often assuming away some of the most physically and psychologically debilitating circumstances of war (see, e.g., McMahan 2009; Fabre 2012; Frowe 2014).

Does the fact that these theories contain idealizations suffice to put their tenability into question? Critics of ideal theory as idealized theory would answer in the affirmative. This unconditional affirmative answer is unwarranted. Whether the presence of idealizations affects the plausibility of a normative theory, I suggest, depends on the level at which the idealizations are introduced.

As Christian List and I have argued elsewhere, one should distinguish between three levels at which idealizations may be encountered in the construction of a normative theory. First, the idealizations might be contained in the theory itself, namely in the body of propositions entailed by its principles. Second, the idealizations might be contained in the conditions of applicability of the theory, namely in the set of circumstances and agents to which the theory is meant to apply. Third, idealizations may be contained in the arguments or evidence offered in support of a given theory (List and Valentini 2016). Which kinds of idealizations are dangerous?

While the presence of idealizations within the theory itself is problematic, this is for reasons other than those cited by critics of idealization. To say that a theory contains idealizations is to say that some of its propositions are false in relation to their subject matter. If a theory of international political morality is idealized in this first sense, all this means is that its prescriptions are incorrect. For instance, if the theory prescribes global socio-economic redistribution, but this is not in fact what

---

2 I am here setting aside the so-called “realist” critique of Rawls-inspired political theory. For discussion, see Duncan Bell’s contribution to this volume.
global justice requires, then the theory is “idealized”: it contains a false statement about international justice.

What critics of idealized theorising typically have in mind are idealizations that feature at the level either of the arguments offered in support of a normative theory or of the conditions of applicability of the theory, which I discuss in turn. An example of the former is offered by cosmopolitan theories of justice: both (i) “relational” theories that globalize the outcome of Rawls’s original-position on the ground that there exists a “global basic structure of society” (e.g., Beitz 1999; Pogge 1989) and (ii) “non-relational” theories that advocate global equality independently of the existence of a global basic structure or particular global relationships (e.g., Caney 2005; for the labels “relational” and “non-relational,” see Sangiovanni 2007).

As mentioned earlier, “globalizing” Rawls’s original position involves importing all of the idealizations that characterize it. These include not only the assumptions of full compliance and favourable conditions, but also a given description of the “contractors” who are to choose the relevant principles of justice. Famously, they are prevented from knowing facts about themselves—including their talents, abilities, and social class—by being placed under a “veil of ignorance.” These assumptions are clearly counter-factual: they involve falsities about the world in which we live. But does the fact that they figure in the argument supporting Rawls’s theory—i.e., the fact that Rawls’s principles would be chosen in the original position—necessarily imply that the theory itself is somehow undermined? It is far from clear why it should (List and Valentini 2016; cf. Lawford-Smith 2010).

The original position is meant to model circumstances of fairness, such that its output may also be regarded as fair. One may of course object to the original position on the grounds that its modelling of fairness is unconvincing. But this would be a substantive objection, not a methodological one having to do with idealization. Consider the following analogy. When considering whether the grade we have given a student whose identity we know is fair, we may want to ask whether we would have given the same grade had we not known the student’s identity. To be sure, this is a counterfactual claim, yet its counterfactual nature, per se, does not appear to invalidate its heuristic value. Perhaps, we can think of a better counterfactual, say one in which not only the identity of the paper’s author is concealed, but also where the handwriting of the author is easier to read. Poor handwriting may in fact negatively influence examiners. Still, the use of idealizations in the evidence/arguments used in support of a theory is not per se problematic.

Similar conclusions may be drawn with respect to even starker instances of uses of idealizations in the “evidence” supporting a theory. Consider the counterfactuals marshalled by “non-relational cosmopolitans” in support of the view that justice demands equality between individuals independently of whether they share a Rawlsian “basic structure,” or other kinds of special relationships. For example, we may be invited to imagine individuals in completely separate plots of land, who do not know each other, and who are differently off: one has plenty, the other very little, and neither deserves his fortune or misfortune. We are then asked whether this deeply unequal distribution of resource strikes us as unjust. A positive answer to this question is evidence in support of non-relational cosmopolitanism. To be sure, the far-fetched scenario involving two individuals is far from an accurate description of the world today. Yet its point is not to offer such a description, but only to prove that our convictions about egalitarian justice do not depend on the presence of a basic structure or special relationships. And to do so, precisely those relationships have to be (counter-factually) assumed away. This is a familiar technique in causal
reasoning. To establish that our headache does not depend on chocolate consumption, we stop eating chocolate and see whether the headache persists. Similarly, to prove that (our thinking about) justice does not depend on special relationships, we assume them away, and test whether our intuitions persist in such counterfactual—relationship-free—scenarios.

The foregoing discussion has shown that the use of idealizations in the arguments or evidence offered in support of a normative theory is not per se a source of concern. Resort to idealization as such does not make an argument a bad one. Though of course many arguments can contain idealizations and also be weak or unsuccessful, for reasons that are independent of the mere fact that they feature those idealizations.

Let me finally turn to idealization at the level of the conditions of application of a normative theory. By “conditions of application,” I mean the circumstances under which its prescriptions hold. For example, consider the conditional prescription: “One ought to abide by the law if most of one’s fellow citizens comply with it.” The duty to abide by the law, here, applies in circumstances of near full compliance. But the above conditional prescription tells us nothing about what we should do when very few comply with the law.

Several theories of international political morality are characterized by the presence of idealizations at the level of their conditions of application. Rawls’s “law of peoples” and contemporary revisionist theories of the just war are two prominent examples. Rawls’s principles of international justice are designed assuming a world populated by societies that are fairly self-contained, and independent of each other. This is also a world in which societies can be easily classified in the following mutually exclusive categories: well-ordered peoples, burdened/less-developed societies, and rogue regimes. In a world populated by these types of societies, Rawls’s principles apply.

It is of course almost superfluous to say that today’s world is far from the simplified one Rawls envisages. Poor societies can also be rogue regimes. Liberal societies are often implicated in the institutional deficits of developing countries (Shue 2002). Furthermore, societies are not relatively independent, but deeply interconnected through the global economy (Pogge 2001). What this means is that the prescriptions contained Rawls’s eight principles may well not apply to, or be appropriate for, the world in which we live. Strictly speaking, they are silent about it. If we want the theory to be able to guide our action in real-world circumstances, this form of idealization is of course problematic. That is, conditional on the aim of the theory being real-world action-guidance, idealizations in its domain of application render the theory highly likely to be unsuccessful: we have very little reason to believe that principles that are appropriate for Rawls’s simplified world also work for ours. Though, if the aim of the theory is not to guide action in real-world circumstances, but simply to outline a number of conditional prescriptions taking the form “In such and such circumstances, one ought to perform such and such actions,” then the presence of idealizations in the theory’s conditions of application is not per se problematic.

In the case of Rawls’s theory of international justice, it seems interpretively plausible to suggest that The Law of Peoples is indeed meant to apply to existing circumstances—indeed, it is meant to guide the foreign policy of liberal societies. To that extent, the presence of idealizations in its conditions of application justifiably raises concerns. The charge of problematic idealization is warranted in this case (Valentini 2009).
Let me now turn to just-war theory. Key tenets of just war theory—especially of jus in bello, as famously discussed by Michael Walzer (1977)—are the moral equality of combatants and non-combatant immunity. The former corresponds to the idea that combatants enjoy the same moral status independently of whether they are fighting on the “just side”; the latter to the idea that deliberately targeting non-combatants is impermissible. Recently, the Walzerian—“orthodox”—paradigm has been challenged by so-called “revisionist” just-war theorists, who wish to tie the permissions and prohibitions applying to individuals in war to their moral liability (see, e.g., McMahan 2009; Fabre 2012; Frowe 2014). The revisionist picture of the just war—and of jus in bello specifically—seeks to make just war principles more “continuous” with ordinary peacetime morality and the ethics of self-defence. On revisionist views, the status of combatants in war should depend on whether they are responsible for unjust threats. Furthermore, for revisionists, some noncombatants may also be legitimate targets, to the extent that they share responsibility for serious enough wrongdoing.

Partly due to their focus on individual defensive killing, revisionist just-war theorists often develop their prescriptions against the background of highly stylized scenarios, some of which involve far-fetched self-defence situations—e.g., ray-guns and sizeable individuals falling down wells or bridges (e.g., Frowe 2014, 22; McMahan 2009, 172). The psychological duress, collectivist dimension, and uncertainty that characterize war are not given much attention.

Although, as discussed earlier, the use of stylized scenarios serves useful theoretical purposes, it also casts doubt on whether the prescriptions arrived at by appeal to them are appropriate for their intended context of application: war (Shue 2010). Some contemporary revisionist just-war theory, then, might also be susceptible to the charge of idealization, at least to the extent that its prescriptions are (i) meant to apply to war as we know it and yet (ii) have been developed and tested assuming circumstances that rarely obtain in real-world wars (cf. Valentini 2016; Lazar and Valentini forthcoming). I shall return to this point in the next section.

The foregoing discussion has shown that the presence of idealizations in a normative theory raises interesting concerns only when the idealizations are found in the theory’s conditions of application. In that case, there might be a mismatch between the theory’s prescriptions and the domain for which they are intended (i.e., the real world). If so, those prescriptions are ill-suited to apply to existing circumstances. Otherwise, the presence of idealizations in the arguments offered in support of a theory is not per se problematic. And the presence of idealizations in the theory itself is problematic, but uninterestingly so: it simply means that the theory includes false propositions about its subject matter (be it international distributive justice, the just war or something altogether different).

4. Responding to the Charge of Feasibility-Insensitivity
A second set of challenges raised against ideal theories of international political morality concerns their insensitivity to feasibility constraints. Once again, several such theories exist. Here are some well-known examples.

- Theories of global distributive justice, including the consequentialist view that one ought to give away all of one’s surplus resources to combat global poverty (Singer 1972), and the cosmopolitan egalitarian injunction to globalize Rawls’s difference principle (Beitz 1999; Pogge 1989). Both kinds of prescriptions appear to ignore feasibility constraints such as individuals’ lack
of selfless altruism, and the absence of powerful enough global institutions capable of realizing cosmopolitan global equality.

- **Theories of global democracy** prescribing the democratization of global politics (e.g., Held 1995; Archibugi 2008; Marchetti 2008). These have been routinely accused of ignoring feasibility constraints, such as the lack of inclusive institutions at the global level, as well as the lack of a “global demos” (Miller 2010).

- **Revisionist theories of the just war** that extend the ethics of self-defence to contexts of war. These are often criticized for failing to take into account the psychological, epistemic and institutional feasibility constraints that are typical of wars. As a result, so the critique goes, if implemented, their prescriptions would be counterproductive: leading to more wrongful killing than alternative, orthodox prescriptions (e.g., Buchanan 2006; Shue 2010; cf. McMahan 2008).

Once again, whether the charge of feasibility-insensitivity is warranted or not depends on the particular theory we are looking at, and its specific aim. First, there is an entire class of theories for which feasibility constraints are immaterial, namely *purely evaluative theories* (Gilabert 2011). These theories do not aim to set out prescriptions and permissions, “oughts” and “mays.” Instead, they aim to develop standards of evaluation, allowing us to determine whether a given state of affairs is good or bad, just or unjust. Since such theories do not issue any “oughts,” they need not concern themselves with “cans”: namely with what is possible or feasible. A state of affairs can be bad, unfair or regrettable even if nobody is responsible for bringing it about, and nobody can remedy it. For example, a natural catastrophe without remedy can be bad, and perhaps even “unfair” or “unjust” if it hits a specific group of innocent individuals. So, to the extent that theories are meant to be exclusively evaluative—hence not directly action-guiding—feasibility-insensitivity is not a concern (Gilabert 2011; Gheaus 2013; Valentini 2016). And it may very well be that some demanding cosmopolitan theories requiring global equality, or global democracy, are best understood as purely evaluative in this sense. They tell us that *there is something to be regretted* in the fact that our world is unequal, and globally undemocratic, even if bringing about a democratic and egalitarian world is currently not in anyone’s power. In doing so, these purely evaluative theories may play an important *critical* function, by highlighting the existence of problematic power relations—even if they fail to provide a clear path to remediying them.³

But the retreat to pure evaluation is unlikely to satisfy critics, nor is it one that many international political theorists will want to undertake. For many, the whole point of theorizing about international political morality is to deliver prescriptions helping us orient ourselves in the real world. And when theories are prescriptive, as opposed to purely evaluative, feasibility constraints matter: “oughts” are valid only when it is possible/feasible for their addressees to act on them—or so I shall assume for present purposes.

There are two broad classes of feasibility-insensitivity for normative theories. First, there are theories whose prescriptions are *genuinely impossible* to realize. Second, there are theories whose prescriptions are *unlikely to be acted upon*. In what follows, I consider each in turn.

³ I am grateful to the Editors of this Handbook for pointing this out to me.
First, some theories may place demands on us that we simply cannot fulfill. Theories requiring global democracy or equality now might be of this kind. No matter how well-intentioned individuals or states might be, global equality or democracy are long-term goals that cannot be brought about today. Any theory demanding their immediate realization would thus problematically fall foul of the “ought implies can” constraint (cf. Meckled-Garcia 2008). That said, the most charitable reading of international political theories advocating global democracy or cosmopolitan equality is as evaluative theories, accompanied by ancillary prescriptions requiring existing agents to do what they “reasonably can” towards the achievement of these long-term goals. A commitment to global democracy or cosmopolitan equality may thus give rise to what Pablo Gilabert calls “dynamic duties,” namely duties to gradually bring about conditions under which the demands of global justice (or global democracy) can be realized (Gilabert 2012). These duties, however, would be constrained by feasibility considerations.

Second, we can turn to theories issuing prescriptions that can be fulfilled, but whose fulfillment is unlikely. Peter Singer’s (1972) demanding principle of global beneficence is a case in point. People are simply unlikely to be motivated to donate as big a portion of their earnings as Singer recommends. In fact, faced with Singer’s extreme requests, they are likely to ignore them. Similarly, it might be argued that revisionist just-war theory asks too much of those engaged in a war. The epistemic and psychological burdens it places on soldiers are so great that they are unlikely to be willing to carry them. Consequently, the institutionalization of revisionist just war theory risks being counter-productive: the relevant prescriptions and prohibitions are likely to be either ignored or misapplied (Shue 2010; McMahan 2008; Lazar and Valentini forthcoming).

What to say about this line of critique? Several responses can be given. Each of them shows that “feasibility-insensitivity” complaints directed at theories whose prescriptions are unlikely to be realized are not really about feasibility, but ultimately trace back to deeper moral controversies. In some cases, a prescription being unlikely to be realized is just a symptom of the prescription being overdemanding (Estlund 2011, 222–23). For example, Singer’s “solution” to global poverty has come under attack precisely for this reason. Each individual, so the argument goes, should be allowed to give some priority to one’s own wellbeing and the wellbeing of one’s near and dear (Scheffler 1994). Yet Singer’s position on global justice ignores this morally salient fact. The “feasibility critique” here simply boils down to a substantive moral disagreement about the costs individuals may reasonably be asked to bear. Similar reflections can be advanced in the case of revisionist theorists of the just war: asking soldiers not to take up arms or not to defend themselves when in doubt about the justice of their cause may simply be too costly.

In other cases, a theory’s being unlikely to be realized is not a symptom of its overdemandingness, but just a by-product of people’s weakness of will or selfishness (Estlund 2014; Estlund 2011). In that case, the “feasibility” critique genuinely has no bite if the theory is meant to track “the morally right answer.” Yet the critique does have bite if the theory in question is meant to set out principles that should be efficacious when embodied in existing laws and institutions. For example, it may be

---

4 Another critique of Singer’s view relates to its lack of sensitivity to the structural causes of poverty. A policy change, oriented towards greater individual or state donations, in the absence of deeper structural reforms of the global economy, may contribute to perpetuating the status quo or possibly have detrimental effects. See, e.g., Kuper (2002). This illustrates the general trade-off between policy and structural reforms we are likely face when trying to make the world more just.
that, in a fully just society, the tax rate on the richest segment of the population should be 70% (I use a similar example in Valentini 2016). Yet, under existing circumstances, one can foresee that such a tax rate would result in considerable tax evasion, tax avoidance, and people relocating to other countries with a more favourable tax regime (cf. Estlund 2011). This being so, a theory setting out prescriptions to be embedded into institutions and policies should indeed take account of likely compliance with them. This point is well understood by revisionist theorists of the just war, who are adamant about the fact that they are only setting out principles of the “deep” morality of war, rather than principles to be embedded in the “laws of war” (McMahan 2008; Lazar and Valentini forthcoming). The latter, revisionist theorists admit, ought to instead take account of likely compliance and other feasibility considerations.

In sum, whether the feasibility-insensitivity critique applies or undermines a given theory depends on the aim of that theory. If the theory is purely evaluative, feasibility constraints are irrelevant to its validity. If the theory is prescriptive or normative, such constraints matter. Prescriptions must satisfy the “ought implies can” proviso. Yet they need not satisfy an “ought implies likely” constraint, as David Estlund has already argued at length (Estlund 2011). Whether moral demands are likely to be realized or not may matter morally only when low likelihood is a sign of overdemandingness, or when principles are meant to offer policy or institutional prescriptions. Otherwise, the fact that the addressees of certain moral demands are unlikely to act on them simply signals their moral weakness.

5. Conclusion
In this chapter, I have offered a moderate defence of ideal theory in international political theorizing. I have suggested that critiques pointing to theories’ use of idealizations and insensitivity to feasibility constraints succeed only in a limited number of cases. Global democrats, cosmopolitan egalitarians and revisionist just war theorists may be guilty of all sorts of mistakes in their theorizing, but the charge of “excessive idealism” is for the most part unlikely to suffice to invalidate their accounts. I have also, indirectly, shown that charges of excessive idealism are too indefinite to carry much significance. Instead of employing the catch-all expression “ideal theory,” it is better to level one’s charges against each position one disagrees with, pointing to the presence of idealizations where none should be, or to the prescriptions issued by the target theory being too demanding, impossible to realize or counterproductive when institutionalised. Without further explanation, a vague charge of “excessive idealism” carries very little, if any, weight.
References


