

# Ideal and Non-ideal Theory in Political Philosophy

J. Ahlin Marceta

As a pre-theoretical starting point, the term “ideal theory” in political philosophy denotes normative theories that constitute some conceptual model of society (or parts of it). For instance, John Rawls’s theory of justice is an example of an ideal theory. “Non-ideal theories” are instead simply those that are *not* constructed accordingly. However, as will be made clear in the discussion below, there is no consensus on how the two terms and their relation should be understood.

The philosophical debate on ideal and non-ideal theory is marked by various complexities. In addition to the problems concerning how to define the two terms and distinguish them from each other, political philosophers debate the feasibility of ideal theory, its methodological role or function, and whether normative guidance should rather be grounded in non-ideal theories, among other things.

In this text, I provide a brief overview of two central topics in this debate. I begin by discussing the basic conceptual ambiguities involved with ideal and non-ideal theory. Then, I account for important criticisms of ideal theory, as well as defenses thereof.

## Conceptual ambiguities

There is little agreement as to how ideal and non-ideal theory should be understood beyond the pre-theoretical starting point above. The notion of “ideal” theory, in the sense relevant to the present discussion, was first mentioned as a theoretical construct in Rawls’s *A Theory of*

*Justice* (see, e.g., p. 245). However, at least to my knowledge, the first detailed discussion of it is found in Onora O'Neill's 1987 book chapter, "Abstraction, Idealization and Ideology in Ethics." In what follows, I discuss four suggestions on how the distinction between ideal and non-ideal theory should be understood.

O'Neill takes ideal theory to be an instance of abstract reasoning. "Abstract ethical and political theories," O'Neill writes in an explication of criticism of ideal theory, "make assumptions about agency which are not satisfied by human agents" (p. 56). In the book chapter, O'Neill defends abstraction in political philosophy arguing, among other things, that it is a both necessary and an unavoidable form of thinking. She identifies two kinds of abstraction.

In a first sense, ideal theories build on idealizations such as, e.g., "moral agents are rational," and "humans have interpersonally comparable utilities." Idealizations of this kind are constructed from features that human beings generally share. It is of course known that particular individuals may lack one or more of these features—the point of ideal theory, in this sense, is to construct general ideals by abstracting away features that distinguish particular individuals from each other. These ideals can then be used as models in political philosophy for explanatory purposes similarly to how mice are used as models in science, i.e., as representations of targets.

In a second sense, O'Neill distinguishes ideal theories that are inherently normative, such as idealizations of how human beings (or other agents, distributions, institutions, etc.) *should* be. Normative ideals are constructed from generalized features that are morally desirable or praiseworthy, such as a hypothetical person who is compassionate, brave, and altruistic. These ideals can also be used as models in political philosophy, but for normative purposes rather than explanatory.

One example of ideal and non-ideal theory in O'Neill's use of the terms is Rawls's. Following John Simmons's interpretation (2010, pp. 7–18), Rawls meant for "ideal theory" to designate a model of a fully just

state-of-affairs. One element in his conceptualization is that ideal theory assumes *strict compliance*. That is, in Rawls's theory of justice (which is ideal in this sense), individual beings comply with the demands of justice. They have a basic sense of what is just and are moved by it as they lead their lives in accordance with society's institutions.

Non-ideal theories instead, on Rawls's account, are theories that are *not* marked by the assumption of strict compliance. In non-ideal theories, people may lack a sense of justice, or they comply merely to some limited extent—or even not at all—with the demands of justice. The purpose of non-ideal theories, in this sense, is to provide guidance with regard to how ideals should be met; they specify the road to justice rather than justice itself.

Laura Valentini (2012) has proposed that the terms ideal and non-ideal theory carry (at least) three distinct meanings in this context. The first sense corresponds to Simmons's analysis, namely that "ideal theory" means "full-compliance theory" (p. 654). In a second sense, it may instead be taken to mean "utopian or idealistic theory" (*ibid*). On this understanding, disagreements regarding ideal and non-ideal theory concern feasibility and considerations of what constraints there are or should be to normative political theory. In the third sense, "ideal theory" means "end-state theory" while "non-ideal theory" means "transitional theory" (*ibid*). Here, the debate concerns "whether a normative political theory should aim at identifying an ideal of societal perfection, or whether it should focus on transitional improvements without necessarily determining what the 'optimum' is" (*ibid*).

Finally, to conclude this section on conceptual ambiguities, Alan Hamlin and Zofia Stemplowska identify four broad approaches to the distinction between ideal and non-ideal theory (2012, pp. 48–9): (1) Full compliance and non-full compliance, (2) idealization and abstraction, (3) fact-sensitivity and fact-insensitivity, (4) perfect justice (or another value) and local improvement in justice (or another value). Any theory in political philosophy can be analyzed in either of these dimensions,

although Hamlin and Stemplowska find implausible categorical distinctions between ideal and non-ideal theories focusing on only single dimensions (p. 49). Taken together, they call this “a theory of ideals.”

In addition to their theory of ideals, Hamlin and Stemplowska suggest that there is a continuum of ideal and non-ideal theory “concerned with the identification of social arrangements that will promote, instantiate, honour or otherwise deliver on the relevant ideals” (p. 53). The continuum marks the degree of feasibility of a theory, where the ideal end is “Panglossian” (p. 56), i.e., very optimistic with regard to the theory’s practical viability, and the non-ideal end is “realist.” Thus, on Hamlin and Stemplowska’s understanding, the debate on ideal and non-ideal theory is multi-dimensional concerning what a particular normative political theory is actually about (as per their theory of ideals) and sensitive to the degree of feasibility of it.

### Criticisms and defenses of ideal theory

Amartya Sen (2006) argued famously against ideal theory starting from the question, “what do we want from a theory of justice?” Sen does not phrase his arguments in the terminology described above. Instead, he uses the term “transcendental” to denote Rawls’s theoretical approach, which is to identify “perfectly just societal arrangements” (p. 216). The upshot of Sen’s arguments is that ideal (or transcendental) theory cannot “address questions about advancing justice and compare alternative proposals for having a more just society” (p. 218).

More specifically, Sen argues first that ideal theory does not suffice to ground judgments between alternative outcomes in a non-ideal world. An ideal theory such as Rawls’s, for instance, does not account for “(1) different fields of departure, (2) varying dimensionalities of transgressions within the same general field, and (3) diverse ways of weighing separate infractions” (p. 219). One answer to Sen’s original question is thus that we want a theory of justice to be able to take into

account and provide normative guidance with regard to the contextual circumstances in any particular justice-related case.

Secondly, Sen argues that ideal theory is not necessary to ground such judgments (pp. 221–2). Two alternative outcomes in a non-ideal world can be internally ranked in moral terms without reference to a moral ideal. In an often-cited illustration, Sen writes that we do not need to know that Everest is the tallest mountain in the world to compare the heights of Kanchenjunga and Mont Blanc (p. 222). Likewise, we do not need to know what ideal justice is to tell which one of two (or more) alternative states is more just. Therefore, in Sen’s view, ideal theory is neither sufficient nor necessary in political philosophy.

Other criticisms of ideal theory include, among other things, the views that ideal theory is ideological (Mills 2005), ineffective (Farrelly 2007), and uninformative or misleading (Wiens 2015). Jacob Levy (2016) has formulated the perhaps most categorical denial of ideal theory in an article suggestively titled, “There Is No Such Thing as Ideal Theory.”

In a defense of ideal theory, Erman and Möller (2013) reject three charges against ideal theory, namely that ideal theory is impossible, distorting, and not action-guiding. The charge that ideal theory is impossible rests on the basic assumption in moral philosophy that “ought implies can” and the observation that political ideals cannot be reached. Therefore, the argument goes, ideal theory cannot ground ought-claims. Erman and Möller responds to it by noting that there is a difference between something being on the one hand *unlikely* or *practically* impossible, and on the other hand *actually* impossible. Only political ideals that are actually impossible are vulnerable to the charge.

The charge that ideal theory is distorting builds on the argument that ideals fail to account for how the world actually is. Theoretical ideals do not include real power structures, and “depicts human agency and social institutions in unrealistic ways” (Erman and Möller 2013, p. 40). Therefore, the argument goes, ideal theory lacks the conceptual tools needed to ground ought-claims. Erman and Möller argue that

the charge is unsubstantiated. It does not follow from the fact that a theoretical ideal is not characterized by all the properties of the real world that the ideal *obscures* those properties or fails to provide guidance with regard to them.

Finally, the charge that ideal theory is not action-guiding comes in two types. The first is epistemological. According to it, ideal theory provides less knowledge than non-ideal theory about what should be done in practice. Erman and Möller discuss various versions of the epistemological charge. However, their response to the charges as a group can be summarized in two sentences: Political ideals do not provide full knowledge with regard to action-guidance, but this is not a problem. Ideal theory acknowledges that practical judgments must be made in particular cases, it is a mistake to expect complete *a priori* guidance from political ideals (pp. 27–32).

The second charge is psychological. Ideals are not immediately directed at the complexities and difficulties of the real world, but only concerns hypothetical states that are distant from ordinary people's everyday experiences (p. 32). Therefore, the argument goes, ideal theory is less motivating for real agents than non-ideal theory. Erman and Möller responds that if the charge is true, which it is not obvious that it is (pp. 33–4), ideal theorists “are not ignorant about the need, in our society, for extra-moral motivators” (p. 36). That is, it may not be the case that ideal theory *should* provide ordinary people with moral motivation—that might be someone else's job.

Other defenses of ideal theory include Valentini (2017). In her book chapter, titled “The Case for Ideal Theory,” Valentini argues that criticism of ideal theory's use of idealization and its insensitivity to feasibility constraints only succeed in a limited number of cases, and that the charge of “excessive idealism” carries little weight, if any. There have also been attempts to re-formulate ideal theories in non-ideal terms. For instance, Marcus Arvan (2014) has re-constructed Rawls's

“original position,” i.e., the thought example of a veil of ignorance and its theoretical role in justice as fairness, as a non-ideal position.

### Concluding remarks

To summarize, there are various conceptual ambiguities involved with the terms ideal and non-ideal theory. Most broadly construed, the term “ideal theory” in political philosophy denotes normative theories that constitute some conceptual model of society (or parts of it), whereas “non-ideal theories” are instead simply those that are *not* constructed accordingly. However, there is no consensus on how to understand and distinguish the two terms.

Furthermore, ideal theory has been criticized on various accounts. Among other things, it has been argued that it is both insufficient and unnecessary to ground normative judgments. However, there are also noteworthy defenses of ideal theory. As political philosophy has taken a “methodological turn” in recent years (Valentini 2017), the debate between proponents and opponents of ideal theory is likely to continue.

### References

- Arvan, M. (2014). First Steps Toward a Nonideal Theory of Justice. *Ethics & Global Politics* 7(3), 95–117.
- Erman, E. & Möller, N. (2013). Three Failed Charges Against Ideal Theory. *Social Theory & Practice* 39(1), 19–44.
- Farrelly, C. (2007). Justice in Ideal Theory: A Refutation. *Political Studies* 55, 844–864.
- Hamlin, A. & Stemplowska, Z. (2012). Theory, Ideal Theory and the Theory of Ideals. *Political Studies Review* 10, 48–62.
- Levy, J. T. (2016). There Is No Such Thing as Ideal Theory. *Social Philosophy and Policy* 33(1-2), 312–333.
- Mills, C. (2005). ‘Ideal Theory’ as Ideology. *Hypatia* 20(3), 165–184.

- O'Neill, O. (1987). Abstraction, Idealization and Ideology in Ethics. In J. D. G. Evans (Ed.), *Moral Philosophy and Contemporary Problems* (pp. 55–69). Cambridge University Press.
- Rawls, J. (1971). *A Theory of Justice*. Harvard University Press.
- Sen, A. (2006). What Do We Want from a Theory of Justice? *The Journal of Philosophy* 103, 215–238.
- Simmons, J. (2010). Ideal and Nonideal Theory. *Philosophy & Public Affairs* 38, 5–36.
- Valentini, L. (2012). Ideal vs. Non-Ideal Theory: A Conceptual Map. *Philosophy Compass* 7, 654–664.
- . (2017). The Case for Ideal Theory. In C. Brown and R. Eckersley (Eds.), *The Oxford Handbook of International Political Theory* (pp. 664–676). Oxford University Press.
- Wiens, D. (2015). Against Ideal Guidance. *Journal of Politics* 77(2), 433–446.