THREE ROLES OF IDEAL THEORY

Lars J. K. Moen
Department of Philosophy, University of Vienna
lars.moen@univie.ac.at

ORCID: 0000 0002 6052 9948

Abstract: Rawlsian ideal theory is meant to perform various roles in non-ideal theory. In this paper, I distinguish between three roles, and I consider the extent to which we can expect ideal theory to perform them. It is meant to serve as a target to guide non-ideal theorising in the long-term. It is also supposed to provide a way of comparing different injustices to tell us which is worst and therefore in most urgent need of a remedy. Finally, ideal theory is the basis for a model of fairness that restricts the set of morally permissible measures in non-ideal theory. I show how the first two roles—the target and urgency roles—are less plausible than the fairness role.

Keywords: Fairness; ideal theory; justice; non-ideal theory; Rawls.

Resumo: A teoria ideal rawlsiana pretende desempenhar diversos papéis na teoria não-ideal. Neste artigo, distingo entre três desses papéis e indago até que ponto podemos esperar que a teoria ideal seja capaz de realizá-los. Ela serve de baliza que deve guiar a produção de teoria não-ideal no longo prazo. É também suposto que ofereça uma forma de comparar diferentes injustiças e assim dizer-nos quando delas é a pior e, portanto, a que precisa mais urgentemente de uma solução. Por fim, a teoria ideal é a base para um modelo de equidade/justiça que limita o conjunto de medidas moralmente aceitáveis no âmbito da teoria não-ideal. Mostra-se aqui como os dois primeiros papéis – o de servir de baliza e o de [determinação] da urgência [relativa] – são menos plausíveis do que o papel em termos da equidade.

Palavras-chave: Equidade; teoria ideal; justiça; teoria não-ideal; Rawls.

Introduction

Rawlsian ideal theory formulates a model of an institutional arrangement that fully satisfies the principles of justice. These principles require that everyone can effectively exercise the basic liberties; there is fair equality of opportunity, so that no one is denied opportunities to pursue advantaged positions in society because of morally

arbitrary features like race or gender; and social or economic inequalities are justified only if they benefit the worst off. In ideal theory, we assume full compliance with these institutions and no shortage of the means necessary for operating them (Rawls, 2001, pp. 13, 47).

This ideal is then meant to serve different roles in non-ideal theory, where these assumptions, and particularly the full compliance assumption, do not hold. Non-ideal theory account for these realistic, non-ideal conditions and considers how to improve them. Ideal theory is meant to provide a target to guide this non-ideal theorising in the long-term. Stemplowska and Swift (2012, p. 376) therefore call this ideal theory's "target role". Ideal theory is also meant to provide a way of comparing the graveness of different unjust practices in the non-ideal world. The injustice that deviates the most from the ideal is in most urgent need of a remedy and should therefore be prioritised. This is therefore ideal theory's "urgency role".

Critics have challenged both of these practical purposes of ideal theory extensively, especially over the last two decades. Most of the focus has been on whether institutions actually can be guided towards an ideal. Critics have rejected this idea of long-term ideal guidance as unrealistic (Farrelly, 2007; Gaus, 2016; Rosenberg, 2016; Schmidtz 2011; Sen, 2006, 2009; Wiens, 2012, 2015). It has even been considered potentially harmful because it distracts us from pursuing closer and achievable improvements (Brennan & Pettit, 2005). Some have also challenged the urgency role by questioning whether we can derive from ideal theory a way of comparing injustices to work out which to prioritise (Gaus, 2016; Sen, 2009).

In this paper, I outline the target and urgency roles and show why critics are right to challenge them. Section 2 focuses on the target role, while Section 3 is about the urgency role. But I also identify a third and largely unexplored role ideal theory serves in non-ideal theory. I call this the *fairness* role, as ideal theory is here a basis for a constraint that rules out certain measures in non-ideal theory as unfair. As I explain in Section 4, ideal theory identifies the conditions under which everyone enjoys basic rights and liberties and fair equality of opportunity. In non-ideal theory, where institutions do not provide these goods to everyone in the society, we cannot deny anyone these goods even if that might be beneficial to others who do not possess them. That would be unfair treatment of those who have no more than they have under ideal conditions and are therefore not responsible for others having less. The fairness role is less ambitious than the other roles, as it merely identifies one consideration to be taken into account in non-ideal theorising. It is therefore unsurprising that it is also the most plausible of the three roles.

1.   Target

In ideal theory, we make certain assumptions, some more realistic than others. An apparently realistic

assumption, at least in modern well-developed societies, is that there is no lack of "economic means, or education, or the many skills needed to run a democratic regime" (Rawls, 2001, p. 47). A less realistic assumption is that "(nearly) everyone strictly complies with … the principles of justice" (Rawls, 2001, p. 13). Everyone, or nearly everyone, is assumed to comply with institutions that satisfy these principles.

Such ideal conditions with institutions operating perfectly according to the principles of justice are what we should work towards in non-ideal theory. Rawls thus takes ideal theory to provide a target for non-ideal theory. Stemplowska and Swift (2012, p. 376) understandably call this ideal theory's "target role". The model of the ideally just institutional basic structure is, in Rawls (1999a, p. 193) view, "the end of political action to which practical decisions are to be made". Without this ideal, he thinks non-ideal theory will be significantly impaired. Without it, "nonideal theory lacks an objective, an aim", he says (Rawls, 1999b, p. 90). Non-ideal theory, for Rawls (1999b, p. 89), considers "how [the] long-term goal might be achieved, or worked toward, usually in gradual steps". Simmons (2010) also endorses this connection between ideal and non-ideal theory. Ideal theory, he says, "dictates the objective, nonideal theory dictates the route to that objective" (Simmons, 2010, p. 12). This ideal target is crucial, he argues, as we would otherwise myopically focus only on short-term improvements and lose out on larger and more distant improvements achievable only by first defining an ideal that we then work out how to steer towards.

But for the ideal to serve this role in non-ideal theory, it cannot be too ambitious, or too remote from the actual world. It clearly cannot "exceed the capacity of human nature", Rawls (1999a, p. 154) says. While it could be clearer what this restriction entails, it at least involves some sensitivity to human motivation. To satisfy the principles of justice, institutions should not have to demand so much of individuals that compliance would involve a high cost (Rawls, 1999a, p. 126). People in a modern society have different conceptions of the good, and no one should have to compromise one's development and pursuit of a conception of the good, at least insofar as it is reasonable—that is, it is compatible with treating everyone as a free and equal member of society.[1]

Principles that go beyond this restriction and require people to compromise their own conceptions of the good are utopian in a bad sense. Rawls's (2001, p. 13) theory, by contrast, is "realistically utopian", he says, as "it probes the limits of the realistically practicable". Just and legitimate institutions should enable everyone to pursue and develop their own conceptions of the good. Everyone, or every reasonable person, can then see the personal benefit of such institutions and will therefore voluntarily comply with them. Such an institutional arrangement, Rawls (1999a, p. 119) says, will "generate its own support".

The target role is the focus of much of the literature on ideal theory. We have seen that Rawls considers it very

---

[1] A person's conception of the good, Rawls (2005, p. 104) explains, consists of the ends and purposes the person considers worthy of her or his pursuit over a complete life.

important by understanding non-ideal theory to lack a clear objective without it. It also receives support from others, such as Simmons (2010), as we have seen, and from Valentini (2009). But the target role has also attracted much criticism. Many critics make the objection that ideal theory cannot specify a target we can expect to motivate people to work towards. Ideal theory abstracts away from real-world problems of how to attract compliance, and we therefore cannot realistically hope to move institutions towards the ideal (Brennan & Pettit, 2005; Farrelly, 2007; Galston, 2010; Levy, 2016; Schmidtz, 2011). Brennan and Pettit (2005) even argue that the ideal target might be harmful, as we might miss out on close and achievable improvements by focusing on large and more distant ones. We are sometimes better off pursuing the "second-best", they argue (see also Goodin, 1995, pp. 52–55; Wiens, 2012, pp. 55–56).

But Rawls (1999a, p. 398) thinks principles are "seriously defective" if people are not motivated to act as they require. The ideal must therefore be sensitive to what we can expect people to comply with it. While this might appear to be a response to the critics' concerns, we should consider that it must involve an analysis of human motivation, and such an analysis is in no way straightforward. It is a complicated analysis that we cannot be sure to clearly determine whether an ideal is realistic enough to guide long-term non-ideal theorising. It is difficult to say how people might be motivated to behave, and this will vary across time. We can design an incentive structure, but it is uncertain whether, or to what extent, people will respond to it as intended. It will therefore be difficult to determine whether a principle is defective or not. As we learn more about human motivation and how people respond to incentives, we might find that principles must be continuously revised to ensure that they can realistically expect compliance. This gives us a moving target whose requirements keeps shifting, and not a fixed point to direct non-ideal theorising towards in the long-term (Gaus, 2016, pp. 59–61; Rosenberg, 2016, pp. 63–65).

In response, Valentini (2009, p. 340) argues that these objections are simply misguided. Ideal theory gives us a framework for criticising individuals' behaviour, she argues, but it need not track actual motivation. But this is no defence of the target role, since steering institutions towards satisfying ideal principles does depend on motivating a certain kind of behaviour. Ideal theory probably can, as Valentini suggests, serve as a basis for evaluating human behaviour. We can stipulate what it means to fulfil the principles and then criticise people for failing to bring about that state of affairs. But to actually motivate a shift towards the ideal state is a different and far more ambitious matter, and it is unclear whether it is possible. Steering institutions towards a particular ideal over time requires an understanding of how to motivate certain behaviour within a complex web of individual interaction, and social scientists still lack such understanding (Moen, 2022a, pp. 294–297; Wiens, 2012, pp. 63–64).

But even if we ignore the problem of how to make individuals behave in the right manner, a further challenge is to specify what this behaviour actually is (Moen, 2022a, pp. 297–298). It is not obvious what individuals should

do in everyday situations to make their institutions conform to the ideal. Problems for the target role therefore need not arise from individuals' unwillingness to comply; they could arise from difficulties of specifying exactly what compliance involves (Goodin, 1995, pp. 39–40).

Rawls (1999a, p. 267) also accepts that ideal theory cannot always specify right actions. Non-ideal theory must deal with complicated real-world situations, and ideal theory's guidance will sometimes "no doubt fail", he says. It might provide no satisfactory answer to the question of how to behave in certain situations. Simmons (2010, p. 21), in his defence of Rawlsian ideal theory, also concedes that ideal theory will sometimes provide no determinate guidance under non-ideal conditions, and he sees a need to consult social science to work out what is politically possible and effective in these situations. Neither Rawls nor Simmons are worried by this observation, however (see also Valentini, 2009, p. 342). Ideal theory need not always specify a right course of action in non-ideal theory. We can still, Simmons (2010, p. 24) claims, "simply muddle through, keeping the target ideal of justice firmly in mind". It can still "point to the overall direction of political action", as Rawls (2005, p. 285) intends it to do.

But it is unclear why these defenders of ideal theory remain so confident in the target role despite these indeterminacies. It is not obvious why this would be no more than an insignificant limitation, and these defenders provide no empirical evidence to support that the ideal can in most cases provide the guidance it is meant to deliver. Motivating the right kind of behaviour is a major challenge. On this point, then, a stronger defence of the target role is needed.

2.   Urgency

The model of a perfectly just society developed in ideal theory is also meant to enable us to compare different injustices in the real world. The idea is that by comparing unjust practices and arrangements to the perfectly just society, we can say which is most unjust (Rawls, 1999a, p. 216). This problem is the one in most urgent need for a remedy, and therefore the one to be prioritised when we try to make improvements on actual, non-ideal conditions. Stemplowska and Swift (2012, p. 376) therefore call this ideal theory's "urgency role". We should strive to remove any injustice, but we should begin with "the most grievous as identified by the extent of the deviation from perfect justice" (Rawls, 1999a, p. 216). Ideal theory, he says, provides "the only basis for the systemic grasp" of "pressing problems" under non-ideal conditions (Rawls, 1999a, p. 8). In light of the preceding discussion of the target role, we might not feel confident about how to remedy an injustice, but the urgency role is at least meant to direct our attention when we under non-ideal circumstances are faced with several injustices.

Rawls's lexical ordering of the principles of justice is important for making justice, or injustice, comparisons. A

violation of a lexically prior principle is more unjust than a violation of a posterior principle (Rawls, 1999a, p. 216). A violation of the basic rights and liberties principle is therefore more important to attend to than a violation of the fair equality of opportunity principle, which in turn is more pressing to deal with than a violation of the difference principle. And while we can remedy a violation of a prior principle by violating a posterior principle, we cannot remedy a violation of a posterior principle by violating a prior principle. That is, no gain in fair equality of opportunity or improvement for the worst off is justifiable if it compromises someone's basic rights and liberties, and no benefit to the worst off can compromise someone's basic rights and liberties or fair equality of opportunity.

The lexical ordering is, of course, very strict. It is too strict, according to some critics, who consider the strict priority to liberty especially implausible (Hart, 1973; Sen, 2009, p. 65). Even Rawls himself does not fully commit to this strict priority rule. It can only provide "an approximate solution to the priority problem", he says (Rawls, 1999a, p. 39). As we have just noted, the priority rule implies that no permissible attempt to remedy an injustice can compromise someone's basic liberties. Rawls (2005, p. 365) adds, however, that this might not be so if the consequences of this strict priority to liberty is intolerable. Simmons (2010, pp. 21–22) thinks Rawls would give priority to overall justice when a remedy to one injustice itself involves violating a principle, thus causing a new injustice. This view gives no clear purpose to the ordering unless we somehow assign appropriate weights to the principles in accordance with the ordering. Remedying a violation of the first principle will then be a greater contribution to overall justice than would remedying a violation of the second principle. But if the benefit from remedying the latter violation is sufficiently great, it could be prioritised.

But without a strict priority rule, we need a sophisticated measure of overall justice to work out, case by case, which injustice to prioritise. But as Gaus (2016, pp. 42–43) points out, this is a considerable challenge. With different dimensions of justice, we need a clear way of weighting each dimension, as we might find that when we compare two arrangements, one is closer to the ideal on one dimension but further away on another dimension. One dimension might be a concern with fairness, which I consider in the next section. Another is Rawls's (1999a, p. 154) requirement that it be publicly known that institutions satisfy the principles of justice. This publicity requirement might rule out the most effective way of enhancing overall justice, or if publicity is itself a dimension of justice, it must be given some weight in the measurement of overall justice. Either way, it is not clear how much weight to give to publicity or any other consideration when we work out how to deal with injustices in non-ideal theory. With adequate knowledge of the context, we might work out what we ought to do, but then we do not rely on a general priority rule formulated in ideal theory.

Valentini (2009) is one defender of ideal theory who acknowledges that committing to a strict priority rule might have unexpected and morally counter-intuitive implications. In her example, strict adherence to the principle

"you ought to be honest" might never have bad consequences under ideal conditions, but it might be self-destructive under non-ideal conditions, where many people will be dishonest (Valentini, 2009, p. 341). But we nonetheless need an ideal model for responding to injustices in the real world, Valentini (2009, p. 346) argues. The model does not specify how to deal with these unjust practices, but it tells us what to aim for. Since the ideally just society is free from racism and sexism, eliminating all real-world racism and sexism should be our target.

But this is no response to the problem of working out a priority rule that can tell us how to work towards the ideal. It is obvious to most of us that racist and sexist practices are unjust and that they have no place in the ideal society. But when we work towards that ideal, we need to make priorities, and we will perhaps consider taking measures that cause new injustices for the sake of remedying existing greater existing ones. We need to make trade-offs, but defenders of ideal theory have not provided clear rules for how to do so. These are no doubt great challenges that defenders of ideal theory have not dealt with, and it is not easy to see how they can do so.

3.   Fairness

The third and final role of ideal theory I discuss here has received less attention than the other two. And it is not explicitly discussed by Rawls, though we shall see how it fits well into his contractualist framework. We shall also see that it is the most plausible of ideal theory's three roles. I call this the fairness role, as ideal theory here provides the basis for a model of fairness that is one normative consideration to be accounted for in non-ideal theorising.

To play this role, ideal theory's full compliance assumption is particularly important (Moen 2022a). Everyone is assumed to comply with institutions satisfying the principles of justice. Under these ideal conditions, everyone holds the basic rights and liberties, there is fair equality of opportunity, and the distribution of income and wealth makes the worst off as well off as possible. Everyone can then be said to possess a fair share of these primary goods—that is, goods all rational persons want regardless of their plans in life. It is not clear what it means that income and wealth is distributed so as to maximise the position of the worst off, but we can see what it means to have conditions under which no one's basic rights and liberties need be compromised to ensure others' basic rights and liberties. And with fair equality of opportunity, no one is denied an opportunity to pursue an advantaged position in the society because of a morally arbitrary feature like race or gender.

When we then turn to non-ideal theory, where we try to make an actual society conform to the ideal model, we rule out institutional requirements that compromise the share of basic rights and liberties and opportunities to hold advantaged positions as impermissible. A compliant individual's share of these goods is no greater than it would be under ideal conditions, and we therefore cannot compromise this individual fair share to improve the position of

others who are less well off. This compliant individual does not contribute to the unfair distribution and compromising her share would therefore itself be unfair.

We can thus see how ideal theory, by performing its fairness role, imposes a constraint on non-ideal theory (Moen, 2022b). The most effective measures in the pursuit of societal improvements might involve compromising compliant individuals' share of basic rights and liberties and opportunities to occupy advantaged positions, but these measures are ruled out as unfair and impermissible. In non-ideal theory, Rawls (1999b, p. 89) says, we look for "morally permissible" ways of pursuing the ideal target, and these unfair measures are not morally permissible however effective they might be.

The fairness role shows how ideal theory plays a role similar to that of the Kingdom of Ends in Kant's moral philosophy. In the Kingdom of Ends, no one is treated as a mere means but always as an end in oneself (Kant, 1998, pp. 41–46). Everyone acts on the categorical imperative by acting according to a maxim one would want everyone to act on. This morally perfect behaviour applies also under non-ideal circumstance, where some do not act on such a maxim. Lying is therefore always wrong even if it produces bad consequences due to others' dishonesty. The Kingdom of Ends thus provides moral rules for the non-ideal world. Rawls differs from Kant in that he uses ideal theory to determine guidelines for social institutions, not for individuals' moral behaviour (Korsgaard, 1996, p. 150). But they both use it to identify constraints on the pursuit of the good.

The fairness role has not been much discussed by political philosophers, at least explicitly. Taylor (2009) offers perhaps the most explicit consideration of this role when he considers the permissibility of affirmative action on the basis of Rawlsian ideal theory. Under ideal conditions, everyone enjoys fair equality of opportunity, so that no one is denied opportunities to pursue advantaged positions in society because of morally arbitrary characteristics. In non-ideal theory, this ideal is used as a constraint as no one can be denied opportunities everyone enjoys under ideal conditions. Doing so might be an effective way of providing opportunities to others, but this measure is impermissible because it involves denying individuals opportunities they enjoy under ideal conditions. A quota system would therefore not be a permissible way of promoting fair equality of opportunity, Taylor (2009, pp. 491–494) argues, but training, mentoring, and funding members of disadvantaged groups would be permissible. Rawls (1999a, p. xv) also notes that fair equality of opportunity should be "secured by provisions for education and training".

The fairness constraint plays an important role in Rawlsian contractualism, as it ensures that every reasonable person benefits from the just institutional arrangement. The key consideration of this contractualist approach is what is acceptable to all reasonable persons. That is, we consider what everyone could accept given that they want an institutional arrangement that enables everyone to pursue their conceptions of the good insofar as these conceptions

are compatible with treating everyone as a free and equal member of society. If all reasonable persons enjoy these benefits from the arrangement, we can expect them to accept it. Under such a mutually beneficial scheme, we can also expect adequate compliance, which makes it stable. For Rawls, just institutions are stable institutions, and they therefore cannot demand more than people can reasonably expect from each other. Reasonable persons should have a persistent desire to comply with these institutions (Rawls, 1999a, p. 296).

The fairness constraint fits into this contractualist framework by ensuring that the institutional arrangement makes no reasonable person worse off. Reasonable persons want no unfair advantages—no advantages they do not enjoy under ideal conditions—but they might not want to concede advantages they would hold under ideal conditions of full compliance. They might therefore reasonably reject institutions that demand such sacrifices. They might make such sacrifices voluntarily, but just institutions cannot require them to do so.

Critics of ideal theory primarily focus on its target and urgency roles—for good reasons, as we have seen. Their criticism does not bear on the fairness role. We have seen that Rawls takes fairness, as well as other considerations discussed in the previous section, to constrain the pursuit of the ideal target. However, the fairness role is distinct from the target role, and we can apply the former while rejecting the latter, as we can use it to rule out measures for short-term as well as long-term improvements.

Pettit (2017, p. 335) is one critic who challenges what he calls "a collective deontological constraint – a constraint that all citizens are required to satisfy together" derived from ideal theory. This constraint, he says, tells us "nothing about what should happen if there is less than universal compliance". Following Sen (2009), Pettit (2017, p. 336) claims Rawls gives no answer to the question of how institutions should function when the ideally just society is not obtainable. It is not clear what constraint Pettit has in mind, but we can now see that there is an ideal-theory-based constraint that can inform non-ideal theorising under conditions of partial compliance. Under such non-ideal conditions, institutional requirements cannot deny anyone basic rights and liberties or opportunities to pursue advantaged positions that everyone enjoys under ideal conditions of full compliance. The fairness constraint can therefore inform "what should happen if there is less than universal compliance".

But while ideal theory can thus serve a purpose in non-ideal theory, we can argue about the extent to which it ought to play its fairness role. Rawls (1999a, p. 26) admits that no concern for consequences would be "irrational, crazy". There might be cases where applying the fairness constraint would have very undesirable consequences. With a low level of compliance, we might see that institutions operating under the fairness constraint are either too demanding or not demanding enough. These institutions are too demanding on compliant individuals when overall compliance is so low that one individual's compliance will have no beneficial consequences. Waldron (1993) applies a "test of effectiveness" to determine whether overall compliance is high enough to give individuals a natural duty to

comply. Waldron (1993, p. 21) notes, however, that institutions in modern liberal societies generally attract adequate compliance to pass this test and that people in these societies therefore have a natural duty to comply.

We might also find that compliant individuals do less than can be reasonably expected of them due to others' non-compliance. This is the type of case Ashford (2000) describes in her example where each of two boats has enough life jackets to rescue a whole group of struggling swimmers. Under ideal conditions, each boat rescues half of the swimmers. But under the non-ideal conditions Ashford describes, one boat leaves without rescuing a single swimmer. Fairness then demands no more from the remaining boat than that saves only one half of the swimmers. It can leave the other half to drown even though it can rescue all of them. Surely, the fairness constraint cannot dictate what should be demanded in this situation. Rawls also acknowledges this. We have a natural duty to do more than our fair share, he argues, when doing so has no significant cost to ourselves and great benefits to others (Rawls, 1999a, 98).

These cases nonetheless illustrate a limit to ideal theory's fairness role in non-ideal theory. And without an overarching principle for weighting different moral considerations, the extent to which the fairness constraint should inform non-ideal theorising will remain unclear. But these cases do not undermine the fairness role insofar as we take the ideal-theory-based fairness model to provide just one concern to be taken into consideration in non-ideal theory. The model does not tell us what to do all things considered; it only illuminates one criterion to consider when we work out how institutions should operate under non-ideal conditions. Its significance depends on the weight we give it. The fairness role is thus a moderate and plausible role ideal theory can play in non-ideal theorising.

4.   Conclusion

Ideal theory is meant to perform different roles in non-ideal theory. We have seen that Rawls intends it to provide a target to aim towards when try to achieve societal improvements under real, non-ideal conditions. The model of a perfectly just society formulated in ideal theory is also meant to help us figure out which existing injustice is worst by enabling us to compare different injustices to the ideal. The arrangement or practice that deviates the most from this model is then the worst and the one we should prioritise as we try to achieve a more just society.

We have seen that there are good reasons to doubt whether ideal theory can perform these target and urgency roles. Given the difficulty of knowing what people can be motivated to do, it will be hard to figure out what a realistic ideal target would actually look like. Motivational constraints will also vary across time, and as we learn more about how people respond to incentives, we will have to adjust our views of what would be a realistic ideal. The target would therefore likely be a moving target and not a fixed point to aim towards in the long-term, as Rawls intends it to be.

The problem with ideal theory's urgency role is that we need a metric for comparing the justice, or injustice, of different arrangements and practices, and it is unclear how ideal theory can provide such a metric. We have seen that there are various considerations that should be taken into account, and it is unclear how we can give these appropriate weights so as to produce a general measure of justice.

The fairness role is the most plausible and the least ambitious of the three roles. Here we use the ideal model of a perfectly just society to see what basic rights and liberties and opportunities to pursue advantaged positions people have when they deny no one the same share of these goods. Under non-ideal conditions of partial compliance, it would then be unfair to deny compliant people rights, liberties, and opportunities they have under ideal conditions, because they do not contribute to the existing unfairness. We have seen that there are limitations to the extent to which ideal theory ought to play this fairness role in non-ideal theory, but by providing a model of fairness that should be taken into account, ideal theory will at least serve some purpose in non-ideal theory.

Acknowledgment

References

Ashford, E. (2000). Utilitarianism, integrity, and partiality. *Journal of Philosophy*, 97(8), 421–439. https://doi.org/10.2307/2678423

Brennan, G., & Pettit, P. (2005). The feasibility issue. In F. Jackson & M. Smith (Eds.), *The Oxford Handbook of Contemporary Philosophy* (pp. 258–279). Oxford University Press.

Farrelly, C. (2007). Justice in ideal theory: A refutation. *Political Studies*, 55(4), 844–864. https://doi.org/10.1111/j.1467-9248.2007.00656.x

Galston, W. A. (2010). Realism in political theory. *European Journal of Political Theory*, 9(4), 385–411. https://doi.org/10.1177/1474885110374001

Gaus, G. (2016). *The tyranny of the ideal: Justice in a diverse society*. Princeton University Press.

Goodin, R.E. (1995). Political ideals and political practice. *British Journal of Political Science*, 25(1), 37–56.

https://doi.org/10.1017/S0007123400007055

Hart, H.L.A. (1973). *Rawls on liberty and its priority*. University of Chicago Law Review, 40(3), 534–555.

Kant, I. (1998). *Groundwork on the metaphysics of morals* (M. Gregor, Ed.). Cambridge University Press.

Korsgaard, C.M. (1996). *Creating the kingdom of ends*. Cambridge University Press.

Levy, J. T. (2016). There is no such thing as ideal theory. *Social Philosophy and Policy*, 33(1–2), 312–333. https://doi.org/10.1017/S026505251600025X

Moen, L. J. K. (2022a). Making sense of full compliance. *Social Theory and Practice*, 48(2), 285–308. https://doi.org/10.5840/soctheorpract202231156

Moen, L. J. K. (2022b). Ideal theory and its fairness role. *Journal of Value Inquiry*. Published online. https://doi.org/10.1007/s10790-022-09905-6

Pettit, P. (2017). Political realism meets civic republicanism. *Critical Review of International Social and Political Philosophy*, 20(3), 331–347. https://doi.org/10.1080/13698230.2017.1293912

Rawls, J. (1999a). *A theory of justice* (Revised ed.). Belknap Press.

Rawls, J. (1999b). *The law of Peoples*. Harvard University Press.

Rawls, J. (2001). *Justice as fairness: A restatement* (E. Kelly, Ed.). Belknap Press.

Rawls, J. (2005). *Political liberalism* (Expanded ed.). Columbia University Press.

Rosenberg, A. (2016). On the very idea of ideal theory in political philosophy. *Social Philosophy and Policy*, 33(1–2), 55–75. https://doi.org/10.1017/S0265052516000376

Schmidtz, D. (2011). Nonideal theory: What it is and what it needs to be. *Ethics*, 121(4), 772–796. https://doi.org/10.1086/660816

Sen, A. (2006). What do we want from a theory of justice? *Journal of Philosophy*, 103(5), 215–238. https://doi.org/10.5840/jphil2006103517

Sen, A. (2009). *The idea of justice*. Belknap Press.

Simmons, A. J. (2010). Ideal and nonideal theory. *Philosophy and Public Affairs*, 38(1), 5–36. https://doi.org/10.1111/j.1088-4963.2009.01172.x

Stemplowska, Z., & Swift, A. (2012). Ideal and nonideal theory. In D. Estlund (Ed.), *The Oxford Handbook of Political Philosophy* (pp. 373–389). Oxford University Press.

Taylor, R. S. (2009). Rawlsian affirmative action. *Ethics*, 119(3), 476–506. https://doi.org/10.1086/598170

Valentini, L. (2009). On the apparent paradox of ideal theory. *Journal of Political Philosophy*, 17(3), 332–355. https://doi.org/10.1111/j.1467-9760.2008.00317.x

Waldron, J. (1993). Special ties and natural duties. *Philosophy & Public Affairs*, 22(1), 3–30.

Wiens, D. (2012). Prescribing institutions without ideal theory. *Journal of Political Philosophy*, 20(1), 45–70. https://doi.org/10.1111/j.1467-9760.2010.00387.x

Wiens, D. (2015). Against ideal guidance. *Journal of Politics*, 77(2), 433–446. https://doi.org/10.1086/679495