In Defense of Strict Compliance as a Modeling Assumption

**Abstract:** Rawlsian ideal theory has as its foundational assumption strict compliance with the principles of justice. Whereas Rawls employed strict compliance for his particular positive purpose, I defend the more general methodological point that strict compliance can be a permissible modeling assumption. Strict compliance can be assumed in a model that determines the most just set of principles, but such a model, while informative, is not straightforwardly action-guiding. I construct such a model and defend it against influential contemporary criticisms of models that assume strict compliance.

**Keywords:** ideal theory, justice, methodology, modeling, Rawls, strict compliance

**I. Introduction**

Ideal theorists who follow in the tradition of John Rawls (1999a: 7–8, 215–16, 308–9; 1999b: 89–90; 2003: 13, 66) hold that the optimally just social world – strictly speaking, the design principles according to which the basic structure of such a world is to be organized – must be identified at the outset to serve as the destination that the transition from the nonideal present ultimately aims to reach. Ideal theory is said to have *priority* over nonideal theory in that the former is needed to serve as a guide for the latter.[[1]](#footnote-1) Transitional justice is not to be styled on a desultory college road-trip, but on a carefully planned out family vacation.

Rawlsian ideal theory both tells us where we should be going and collaborates with nonideal theory to determine how to get there. The Rawlsian model, as described by Gerald Gaus and Keith Hankins (2017: 177), is like “using Google Maps to get directions” as there are “two relatively clear, fixed points: where we are now, and the ideal” and the “practical problem of political philosophy is to find a route between them.” Proponents of this broadly Rawlsian model of ideal theory seem committed to two modeling tasks which are of different types. The first task is to construct a model of where we should be going in the form of design principles for the optimally just social world. Call that model *Ideal Destination*. The second task is to provide a model of the ideal transitional steps from the present social world to the one identified in the first modeling task. Call this model *Ideal Transition*.

This paper concerns the justifiability of the modeling assumption that Rawls (1999a: 8) called *strict*, or what some in the literature call *full*, *compliance* – which is the assumption that supposes that everyone will “act justly” and “do his part in upholding just institutions” – in determining the content of *Ideal Destination*. I argue that using strict compliance as a modeling assumption can have a valuable and non-distortionary role to play. To do so, I introduce a two-stage model which I call *Ideal Justice* in Section II. The first stage – what I call the *ordering stage* – constructs a priority ordering, *irrespective of compliance considerations*, of possible social worlds organized according to different sets of principles of justice. The second stage – what I call the *selection stage* – reintroduces compliance considerations in the determination of what is appropriate for nonideal conditions.

However, Alexander Rosenberg (2016), David Schmidtz (2011; 2016; 2017), Jacob Levy (2016), and Christopher Freiman (2017) have all offered forceful criticisms of assuming strict compliance in a model of ideal justice. At the heart of this general line of criticism is the concern that assuming strict compliance will distort our theorizing. Though in the same vein as Charles Mill’s well-known assessment of ideal theory, as such, being distortionary, the claim of distortion presently under consideration is restricted to a subset of ideal theory – namely, those models that assume strict compliance – and, thus, has a narrower scope.[[2]](#footnote-2) Given the prominence of these criticisms, I devote a substantial portion of the paper to defending the model from its various critics. In Section III, *pace* Rosenberg, I follow many philosophers of science and social science in thinking that an idealization, such as strict compliance, can be justifiably made in order to make a problem tractable. Section IV and Section V address criticisms raised by Schmidtz of the model that results when strict compliance is assumed. I make the case that assuming strict compliance does not necessarily produce an incoherent model in Section IV and that it does not necessarily yield an irrelevant model in Section V. In Section VI, I rebut arguments made by Levy and Freiman that assuming strict compliance assumes away the very problem under consideration. In Section VII, I conclude that strict compliance can be a justifiable modeling assumption.

**II. *Ideal Justice*: A Dynamic Model**

Rawls thought that strict compliance must be assumed in order to accurately identify the ideally just destination.[[3]](#footnote-3) But how does this assumption actually help determine the destination to be modeled? Rawls is less than clear. A. John Simmons suggests that the purpose of assuming strict compliance is to control for “noise” when comparing societies arranged in accordance with different principles of justice.[[4]](#footnote-4) By this, Simmons seems to mean that strict compliance allows for societies ordered according to different principles of justice to be compared in terms of a *single* parameter – justness. This single-parameter comparison allows for a determination of what organizing principles lead to the most just social world. The society that ranks highest is to be understood as the ideal destination.

Supposing Simmons’s filling of the Rawlsian justificatory-lacuna is correct, Rawls leaves no such gap in how the principles, once chosen, get applied in a given context. Through the procedure that Rawls (1999a: 171–76) calls “the four-stage sequence,” the abstract principles of justice are given increasingly determinate institutional-specification for the particular context. While the principles alone “are not sufficiently determinate to yield recommendations for a specific society”, the four-stage sequence tailors the “general content” of the principles “to the particular conditions of a given society” (Ronzoni 2014: 290). The first stage determines the principles of justice themselves via the original position with a full veil of ignorance – which excludes arbitrary facts, including those about compliance, from impacting the selection of principles – intact.[[5]](#footnote-5) At the second stage, a “constitutional framework for the basic structure” of society is chosen (Ronzoni 2014: 291). To aid the chooser(s) in making this determination, Rawls lifts the veil to admit knowledge of the “history their own society has undergone so far; what level of economic development the country has reached, as well as which natural resources it is endowed with; and what its general political culture and tendencies are” (Ronzoni 2014: 291). In the third or legislative stage, the veil is further lifted to allow the chooser(s) knowledge of “the full range of general economic and social facts” so that they can design legislation for realizing the two principles under the constitutional framework outlined in stage two (Rawls 1999a: 175). Importantly, the second and third stages are not necessarily to be understood as a strict sequence. Only by “moving back and forth between the stages of the constitutional convention and the legislature” is “the best constitution is found” (Rawls 1999a: 174). The fourth and final stage – what Bill Edmundson (2017: 65) calls “the administrative/judicial stage” – has the veil fully lifted and includes a “judicial review of legislation to assure that it is constitutionally sound, as well as execution of legislation and the behavior of persons in their role as citizens.” At this point, “everyone has complete access to all the facts” including “particular facts about individuals such as their social position, natural attributes, and peculiar interests ” (Rawls 1999a: 175).

Despite the various merits of Rawls’s approach, my interest is not in defending either Rawls’s four-stage sequence or his own particular use of strict compliance in ideal theory. Rather I will be defending a more general position: the use of strict compliance in what I call the *ordering stage* – in which a priority ordering of alternatives *irrespective of considerations about compliance* is constructed – of an ideal theory of justice. The ordering stage creates a ranking of alternative sets of organizing principles. The highest-ranking set of principles, as Rawls (1999a: 215) puts it, “define then a perfectly just scheme; they belong to ideal theory and set up an aim to guide the course of social reform.” Then what I am calling the *selection stage* reintroduces considerations about compliance into the decision-making calculus. When deciding what set of organizing principles are appropriate for actually guiding reform in the nonideal world, compliance considerations cannot be overlooked. The chooser in the selection stage determines how to trade-off considerations about compliance with considerations about justice. The higher the alternative is on the priority ordering constructed in the ordering stage, the more just the selection will be. But it will not necessarily receive as much compliance. And it is important not to forget that what is opted for at selection stage is not necessarily – indeed, likely is not – the highest-ranking option in terms of justice.

An illustration of this framework in the form of a model – which I call *Ideal Justice* – may prove instructive.

*Ideal Justice*: A set of individuals, S, is responsible for determining the optimal principles of justice from the set {A, B, C, D, E} where each refers to a different set of principles.[[6]](#footnote-6) Irrespective of concerns about *compliance*, S strictly prefers A > B, B > C, C > D, and D > E. Thus, the ordering stage produces the priority ordering of {A > B > C > D > E}. In the selection stage, compliance concerns are reintroduced. Suppose compliance with A requires that agents be unrealistically compliant in a way that would be socially destructive. Given this constraint, S rules out A and goes on to select B. Still A would be more justice than B and serves as the ideal if people were (or become) more inclined to comply than they are currently.

*Ideal Justice* employs a two-stage process. In the ordering stage, strict compliance with the principles is assumed and S ranks the principles solely according to the justness of the principles. Then, in the selection stage, S opts for the most just set of principles that is consistent with real levels of compliance.[[7]](#footnote-7) Indeed, S may have reason at the selection stage to opt for a lower ranked alternative on the priority ordering if it easily generates compliance and other higher ranked alternatives have difficulty generating compliance. Of course, different choosers will employ different tradeoff costs in the selection stage. What is important is that the priority ordering constructed for a given criterion (or set of criteria) in the ordering stage is generated on the basis of that criterion (or set of criteria) and is not corrupted by looming tradeoff costs which enter the picture at the selection stage.

Having introduced a model that assumes strict compliance, it is worth making a few methodological remarks about the relationship between *Ideal Justice* and Rawls’s four-stage sequence. Whereas Rawls is defending a *particular* use of strict compliance in a *particular* ideal theory of justice, I am advancing the more general methodological point that strict compliance can be assumed when theorizing about justice. If it were definitively shown that Rawlsian ideal theory was wrong, it could still be the case that strict compliance is a permissible modeling assumption in a different, non-Rawlsian ideal theory. But if I am wrong and strict compliance is not a permissible modeling assumption, then Rawls must also be wrong because his particular version of ideal theory assumes strict compliance.

This is not to say that Rawls would embrace *Ideal Justice* full-stop. The first stage of Rawls’s sequence could be viewed as a particular version of the ordering stage. The original position offers a – though certainly not only – way of modeling the choice situation that parties confront when constructing a priority ordering of alternative sets of principles of justice. So, while Rawls offers a version of the ordering stage, it is unclear whether he would be comfortable with the latitude that *Ideal Justice* allows in how to model the ordering stage.

On this point, it is worth ruminating on the different ways of employing strict compliance in the ordering stage. Rawls does so in the context of the original position, but one could defend assuming strict compliance with alternative sets of principles of justice.[[8]](#footnote-8) To arrive at a set of principles that are not identical to Rawls’s two principles while retaining the assumption of strict compliance, one could modify the choice problem being modeled along, at least, two dimensions.[[9]](#footnote-9)

The first way would be to “reveal” or make known aspects of the *circumstances* and the *environment*. In Rawls's (1999a: 118) theory, parties behind the veil of ignorance do not know their

place in society, [their] class position or social status…fortune in the distribution of natural assets and abilities, [their] intelligence and strength…[their] conception of the good, the particulars of [their] rational plan of life, or even the special features of [their] psychology…the particular circumstances of their own society…[and] have no information of which generation they belong.

One may deem some or all of this information relevant to deciding what principles of justice are the ones with which individuals are to strictly comply. And admitting this information could change the principles selected. Rawls could be mistaken in his determination about what circumstantial and environmental information it is important that parties faced with the choice problem have, but right that whatever the parties decide should be strictly complied with.

The second – what we might call *feasibility* – dimension that could lead to different outcomes at the ordering stage concerns how to “delineate the feasible set” – which is to “decide how utopian we are willing to be” in setting the bounds of what is politically possible (Cowen 2007: 1). While an extensive literature has developed on questions concerning feasibility, what is relevant here is that different conceptions of feasibility will determine what sets of principles are to be considered in the ordering stage (e.g. Southwood 2018). Those with more realist inclinations will set the bounds more tightly. Perhaps the principles necessary to realize Cohen's (2009) ideal socialist camping trip will be deemed outside the feasible set as a result of being too utopian, and, therefore, will not be among the alternatives considered in the ordering stage. Those with more utopian leanings will consider additional sets of principles at the ordering stage. Perhaps the principles needed to realize Cohen’s ideal socialist camping trip will be considered and ranked higher than Rawls’s two principles. I suspect that there will be non-trivial disagreement about whether Rawls’s two principles are counted as being amongst the live alternatives to be considered in the ordering stage. The function of the ordering stage is ultimately to create a ranking of sets of organizing principles in terms of how just they are; yet, there is not consensus on what is under consideration. Again, Rawls (2003: 4) could be mistaken about his conception of justice being, as he says, within “the limits of practicable political possibility” but right that full compliance with whatever is the optimal conception of justice would be ideal.

Having demonstrated the pliability of *Ideal Justice*, it is also worth stressing its *dynamism*. Perhaps the best way of demonstrating this aspect of *Ideal Justice* is by considering an extremely important concern raised by Adam Smith (1982: 234) in *The Theory of Moral Sentiments* in which he writes: “to insist upon establishing, and upon establishing all at once, and in spite of all opposition, every thing which that idea may seem to require, must often be the highest degree of arrogance. It is to erect his own judgment into the supreme standard of right and wrong.” Smith is worrying about models of justice that are committed to a *static*, once-and-for-all destination of justice. Herein lies the trouble with the extensively discussed topographical metaphors used by Amartya Sen (2011), A. John Simmons (2010), and David Schmidtz (2011). Whether we are talking about the mountain peaks of Sen and Simmons or about Schmidtz’s featureless desert plane, these are *fixed* geological structures. Smith’s point is made in the language of the mountain metaphor by Alexander Rosenberg (2016: 62) who argues that the appropriate geological metaphor is “a desert landscape with sinkholes that appear, disappear, change shape and location, over time, and influence one another’s sizes and shapes as they do so.” Smith and Rosenberg maintain that an ideal theory of justice must be *dynamic*. And I agree. The relevant question becomes whether an ideal theoretic model that assumes strict compliance – like *Ideal Justice* – can be dynamic?

Answering this question necessitates first clarifying what is meant in claiming that “an ideal theory of justice must dynamic.” On one interpretation, the claim could be asserting that the content of *Ideal Justice*, specifically the priority ordering constructed in the ordering stage, must be dynamic. The content of the ideal changes as the conditions on the ground change. This seems to be the view defended by Alex Guerrero (2017) and Shmuel Nili (2018: 99–104), respectively, in recent papers. A commitment to dynamic ideal theory, so understood, entails the rejection of Platonic ideal theory as well as other forms of universal realist utopian theorizing.[[10]](#footnote-10) This seems to be Rosenberg’s criticism. Yet, if this is what is meant by dynamic ideal theory, then we have moved from a difficult methodological debate to a seemingly intractable metaphysical and meta-ethical debate. If forms of universalistic realism are false, then theorizing will likely need to be dynamic as a result of the fact-sensitivity of justice. If, however, it is true, then static theorizing remains a live alternative. Reconciliation on this front seems unlikely.

Fortunately, there is another justification for dynamic theorizing that does not follow from a commitment to a particular moral ontology. Irrespective of metaphysical convictions, the need for dynamic ideal theorizing can be understood as following from the *epistemic* limitations of theorists in constructing models of ideal justice. The way to excavate insight about ideal justice is by adopting “the view of scientists about the use of idealized models” which is “pragmatic and pluralistic” (Ismael 2016: 14).[[11]](#footnote-11) To be engaged in dynamic ideal theory, then, is, in part, to not assert the truth of one’s preferred model of ideal justice with too high a credence. Indeed, as Simmons (2010: 36) acknowledges, we are susceptible to becoming “preoccupied with particular targeted injustices that seem to us especially grievous.” Such preoccupations lead theorists to bias their models of ideal justice in ways that overly reflect their priors. Given the complexity of ideal justice and the fact that it “is an integrated goal and that activism in one domain has the potential to affect adversely the achievement of justice in another,” going “all-in” on one’s preferred model of ideal justice exhibits at least some degree of methodological hubris.[[12]](#footnote-12) Dynamic ideal theory, so understood, is the enterprise of constructing a plurality of models of ideal justice all aimed at furthering our understanding of a complex social phenomenon.[[13]](#footnote-13) As social learning occurs, models are updated. Some old models prove untenable, such as those that defended inherent superiority along racial and gendered lines, and are progressively replaced with new models that express equality along those dimensions. Thus, the “optimal political structure cannot be fixed once and for all” as “which political structure is best” will be one that is “able to adapt to…changes and remodel itself in a way that constantly approximates the optimal structure” (Aligica, Boettke, and Tarko 2019: 148).

What this importantly shows is that dynamic models and models that assume strict compliance need not be at odds. A model of “justice” that assumes strict compliance with racist and sexist principles is problematic not because it assumes strict compliance, but because it does so on unjust principles. Such a model is simply incorrect. But we need not throw out the strict compliance baby with the static model bathwater. By decoupling the two, it becomes clear that we can identify as ideal a world in which strict compliance with the optimal principles of justice obtains while remaining cognizant that we may be wrong about what the optimal principles are. That is, we can – and likely will need to – update on with what design principles there should be strict compliance. A dynamic model allows for feedback mechanisms to, over time, improve the veracity of the priority ordering.

Put differently, the static approach to ideal theory only works if one has constructed *Ideal Destination* correctly. However, if one constructs a less than completely correct *Ideal Destination* model, then the static approach seems in need of a justification in that its very motivation was to prevent us from arriving at the wrong place. As Gaus (2016: 88–89) puts the epistemic worry:

A century witnessing Stalin, Mao, and Pol Pot disastrously confirmed Mill’s judgment; Robespierre is insignificant compared to this utopian trinity. The worry that certain judgments of the unchanging ideal will give rise to recommendations for immediate implementation is by no means a “utopophobia” of the liberal fallibilist.

History shows that some models of *Ideal Destination* that have purported to be the final destination of justice, when implemented,moved us vastly further away from the normative target. Gaus is certainly correct to emphasize this cautionary note. The worry indicts static theorizing and those who advocate the “immediate implementation” of the product of said theorizing. But the worry need not implicate all theorizing about justice that assumes strict compliance, for assuming strict compliance is consistent with dynamic theorizing. Static theorizing is a product of a broader modeling strategy, not an entailment from assuming strict compliance. Specifically, static theorizing is not a fallibilistic modeling strategy, while dynamic theorizing can be.

Social progress is not linear, and this is a problem for all methodologies – those static and dynamic – with epistemically limited parties. So, while the proper analogy for social progress is not a desultory college road trip, it also is not a carefully planned family vacation. It is more like a, to borrow a term from Elisabeth Ellis (2005), *provisionally* planned vacation that is prone to revision. There may be deep disagreement about where to go and the destination chosen may ultimately not be the best one amongst the alternatives, but this does not mean that the project of trying to identify and arrive at the optimal vacation spot is a Sisyphean endeavor.

There are real challenges both in determining *what* model of ideal justiceis the correct one and in figuring out what to do given deep *disagreement* about this determination. But from this it neither follows that particular specifications of *Ideal Justice* are not candidate models, nor does it follow that strict compliance is an inappropriate modeling assumption in constructing such models. Moreover, employing strict compliance as a modeling assumption does not render the resulting model static. Those who defend static ideal theory do so as a result of a broader modeling strategy, not because strict compliance is assumed in the model.

Having presented the preliminary case in defense of using strict compliance as a modeling assumption, the following sections defend it from its critics.

**III. Is Assuming Strict Compliance Merely an Unjustified Simplification?**

To model is to simplify. By definition, a model of x leaves out actual features of x. The modeler is responsible for determining which features get represented and which get idealized or assumed away. Some determinations of what is to be idealized will be justifiable in that they will not undermine the purpose of the respective models. Others will not be justifiable in that they will do precisely that.

Ideal theorists in the Rawlsian tradition believe that using strict compliance as a modeling assumption is an instance of the former. However, one may argue that it is actually an instance of the latter by appealing to the modeler’s motivation in making the determination. Perhaps if a modeler *only* makes an assumption to simplify the problem or make it epistemically tractable, then the modeler is not justified in making this assumption. A simplification must be theoretically defensible and appealing to the modeler’s epistemic limitations is not. In our case, the worry is that if the modeler only assumes strict compliance in order to make the task of identifying the ideal principles of justice more determinate, then the modeler is not justified in making the assumption. Admitting the realism of partial compliance induces indeterminacy into identification of the ideal principles, but this indeterminacy is part and parcel of the task. A good modeler must face the indeterminacy both about precisely what degree of compliance is required for properly evaluating the justness of the principles and about how judgments can be made about the justness of the principles themselves if compliance considerations are simultaneously being evaluated head-on.

This seems to be Alexander Rosenberg’s critique of Rawls’s project. Rosenberg (2016: 59) writes, “Rawls needs a more compelling reason than a desire to simplify his problem by excluding noncompliance. Abundance would simplify his problem even more, but Rawls does not make that assumption.” He interprets Rawls as assuming, without justification, strict compliance *merely* because partial or noncompliance makes the problem too hard to be solved without doing so.

First, we have reason to be skeptical that making an idealizing assumption in order to make the problem solvable necessarily counts as a bad reason to make it. Such idealizing assumptions are routinely made in scientific and social scientific practice. Michael Weisberg (2007: 640), discussing types of idealizing assumptions in science, calls this a “Galilean idealization” which “is the practice of introducing distortions into theories with the goal of simplifying theories in order to make them computationally tractable.” The justification for a Galilean idealization is pragmatic: “We simplify to more computationally tractable theories in order to be able to learn from them” (Weisberg 2015: 99). In *Ideal Justice*, strict compliance *qua* Galilean idealization allows us to learn about the justness of the principles themselves by fixing the compliance parameter. Concerns about indeterminacy and complexity, though real, are set aside, in part, to make the task of constructing a priority ordering tractable.

Importantly, this cannot be the end of the story for the Galilean idealizer. Keeping the learned insights in mind, the Galilean idealizer must begin to reintroduce the constraints that were idealized away. As Weisberg (2015: 99) puts it: “Since the justification is pragmatic and tied to tractability, advances in computational power and mathematical techniques should lead the Galilean idealizer *to de-idealize, removing distortion and adding back detail to her theories*.”[[14]](#footnote-14) In *Ideal Justice*, this seems to be precisely what the selection stage does. The compliance considerations that were idealized away in order to construct a priority ordering based wholly on the justness of the principles get reintroduced. Compliance factors into the decision made at the selection stage, where tradeoff costs (i.e. how much compliance with the principles versus how just are the principles) are to be weighed.

Suppose one is not convinced that assuming strict compliance to make the problem more tractable justifies it. Rosenberg’s argument might be that an adequate justification would provide *independent* reason, which presumably tractability is not, to make the assumption. Is there such a reason to assume strict compliance? Simmons suggests that Rawls introduces the strict compliance assumption to (1) help overcome the indeterminacy problem resulting from admitting degrees of partial compliance *and* (2) ensure that the specification of the optimal set of principles of justice is determined *solely* by the justness of the principles.[[15]](#footnote-15) Rosenberg holds that (1) is a bad reason to assume strict compliance. He expresses this by saying that assuming abundance (more accurately, superabundance) could simplify the problem by making it so that the circumstances of justice do not apply, but this would be to assume away the problem.[[16]](#footnote-16) However, Rosenberg says nothing about (2). On Simmons’s interpretation of Rawls, (1) enables (2) to be successfully performed. The problem is simplified not just in order to be able to be solved, but in order to be able to be solved *correctly*. Rawls wants to make sure the priority ordering is constructed according to the single relevant parameter – the justness of the principles – by controlling for variables – namely, the motivational shortcomings of agents – that are not directly relevant to that parameter.

**IV. Is Assuming Strict Compliance Incoherent?**

David Schmidtz has provided arguably the most sustained criticism of using strict compliance as a modeling assumption. One dimension of his criticism concerns the *coherence* of assuming strict compliance in a model of ideal justice. Schmidtz suggests that there is a sort of theoretical incoherence when one models the level of compliance with the principles of justice as one exogenous variable and the type and quality of the principles as another exogenous variable.[[17]](#footnote-17) Contrary to this modeling strategy, compliance must be modeled as “an endogenous variable; the extent of compliance is not externally determined but is instead a function of the principles chosen. When we choose a principle, and any particular way of trying to put it into practice, we choose a compliance problem at the same time” (Schmidtz 2011: 778). The idea is that “good” principles of justice will be ones that incentivize agents to comply with them; “bad” principles will not. The content of the principles should do the motivating.[[18]](#footnote-18) We cannot assume away compliance because the question we are trying to answer is precisely ‘what are the most just principles with which agents will comply?’ To do so would be incoherent.

This question is certainly an important one, but I fail to see why coherence requires that we think that it is the *only* one. A second, though I believe lexically prior, question seems to be: ‘what principles are the most just?’ To illustrate the difference in the questions, recall the structure of *Ideal Justice*. The ordering stage provides the answer to the question ‘what principles are the most just?’ while the selection stage identifies the answer to the question ‘what are the most just principles with which agents will comply, here and now?’

Schmidtz does not find reason to differentiate between the questions. It is, Schmidtz (2017: 137) says, “a mistake to think we are imagining what *is* ideal when we imagine what *would* be ideal if compliance were something we got for free, rather than being the precarious achievement that it is.” Neither Rawls nor Simmons think that compliance is free in deciding what is the ideal thing *to do in the circumstances*. Simmons (2010: 9) is explicit that “the assumption of strict compliance should not be understood to amount to an assumption that all possible principles for ordering society’s basic structure are equally good at motivating compliance.” Both Rawls and Simmons do think that it would be a good thing if people complied with whatever principles were deemed most just at the completion of the ordering stage. Then, the selection stage would just opt for the highest option on the priority ordering.

Far from incoherent, distinguishing between the two questions seems imperative. Absent an antecedently specified answer to the question posed at the ordering stage, it is unclear how the second question can even be answered. In order to answer what are the most just principles with which agents will comply, we need to have a *justice-ranking* of the principles in the first place. Otherwise, though we may be able to tell whether agents will comply with some set of principles, *j*, we neither seem able to tell whether *j* is the *most* just set of principles in the superset, J, nor are even able to determine how *j* stands in relation to other candidate sets of principles in J.

Putting the point differently, Schmidtz does not distinguish between what would be ideal *simpliciter* – that is, if people were better than they actually are – and what would be ideal (a better word may be ‘best’) *in the circumstances* – taking people as they actually are. Returning to *Ideal Justice*, A is ideal *simpliciter* and B is ideal *in the circumstances*. In deciding what to do in our present social world, Schmidtz is undoubtedly correct that compliance ought to be considered alongside justice as well as many other normative and empirical considerations. Yet, this does not mean that figuring out what is ideal *simpliciter* is theoretically incoherent.

Assuming strict compliance is only incoherent if one thinks that the *only* relevant question about justice is the one answered in the selection stage. Indeed, it may be the case that Schmidtz believes that the question of what would be ideal *simpliciter* is logically incoherent. The argument would go that if the question is logically incoherent, then theories that attempt to be answers will also be incoherent. But this does not follow from Schmidtz’s incoherence argument. A more ambitious argumentative strategy would be needed. More likely, Schmidtz would acknowledge the difference in the questions but argue that the counterfactuals involved with what would be ideal *simpliciter* are irrelevant. The question to be answered at the ordering stage is not theoretically flawed, but practically. This response is explored in the next section.

**V. Does Assuming Strict Compliance Make the Resulting Model Irrelevant?**

Schmidtz (2016: 6) raises a distinct, but related objection to using strict compliance as a modeling assumption. He worries that claiming that we would not have non-compliance in the ideal world is akin to saying that we would not need to drive defensively in the ideal world, and such remarks are “about a world whose problems, and therefore whose solutions – whose ideals – are not like ours.” Whereas the objection in Section IV challenges the *coherence* of the strict compliance assumption in a model, the present objection concerns the *relevance* of the resulting model. Assuming strict compliance yields a model world that is so distant from the actual world that it lacks application. The distance between the worlds is so vast that solutions in the strict compliance world are irrelevant to the actual one.

I embrace Schmidtz’s analogy. In the ideal world, we would neither have compliance problems nor have to drive defensively. Schmidtz takes this to be a sort of analogical *reductio*, whereas I embrace the truth of the statement. He is certainly correct that a world where we do not need to drive defensively is a world whose problems are not like ours. Indeed, it seems like a vastly *better* world. What we certainly do not want to do is act in the actual world *as if* we did not need to drive defensively. Doing so would lead to catastrophically bad consequences. This point was famously made by Adam Smith (1982: 234):

[I]n the great chess-board of human society, every single piece has a principle of motion of its own altogether different from that which the legislature might chuse to impress upon it. If those two principles coincide and act in the same direction, the game of human society will go on easily and harmoniously, and is very likely to be happy and successful. If they are opposite or different, the game will go on miserably, and the society must be at all times in the highest degree of disorder.

Fortunately, ideal theorists can heed Smith’s advice. That is precisely the purpose of distinguishing between the ordering stage and the selection stage. The ordering stage figures out the best possible world, while the selection stage figures out what is best thing in this world. Notice that the ordering stage is not inconsistent with Smith’s position. Indeed, Smith (1982: 234) seems to advocate for something akin to the ordering stage when he says: “Some general, and even systematical idea of the perfection of policy and law, may no doubt be necessary for directing the views of the statesman.” Following Smith, there is no *a priori* reason to think that assuming strict compliance at the ordering stage makes its results irrelevant.

Perhaps Schmidtz thinks we *cannot* reach a world where there is not a need to drive defensively or where everyone willing complies with the principles of justice. Such a modal claim requires investigation. It is unclear to me what notion of possibility one could have in mind to think that it is *impossible* to reach one of those worlds. More likely, the claim is that we *will not* reach a world where there is not a need to drive defensively or a world where everyone willingly complies with what justice demands. It is a question of *probability*, not of *ability* (e.g. Estlund 2011; 2014).

How far away are we from a world where we do not need to drive defensively? I actually do not think that far. A world with self-driving cars is on the horizon. It is certainly true that a world with self-driving cars is “a world whose problems are not like ours.” But that world is close and those problems will be *our* problems soon. Laying the groundwork in the present such that those problems can be more easily addressed in the future seems to be the appropriate course of action.

It is certainly the case, though, that we are closer to not needing to drive defensively than we are to not needing to structure institutions in a way that individuals do not view others as ‘opponents.’ The problem may just be with the analogy. What Schmidtz seems to have in mind is that the probability that individuals will develop the requisite moral motivations to realize a world in which strict compliance obtains is simply too low to justify preparing for that world. If this is the case, then the issue might seem to turn empirical. We would need to investigate whether there is reason to think that the social world is moving (and, will continue moving) in a direction where more (and, eventually all) people are complying with what justice demands.

Fortunately, the issue does not ultimately turn on empirical considerations. Even if the empirics end up pointing decisively in Schmidtz’s favor – though I am not sure that they would, they well might – he still would need to show why knowledge of a strict compliance world has *no* relevance in the actual world. Granting that the actual world will never realize strict compliance on the principles of justice, such an admission would not entail that consideration of such a world could not be *informative* in the actual world. This is precisely the point made by Weisberg about Galilean idealizations. This point gets seconded by Kogelmann and Gaus (2017: 232) when they say that in “social science, no model or theory explains most of the variance of interestingly complex phenomena, but this does not mean the model is without predictive and explanatory power.”

In responding to Schmidtz’s forceful objections, we are led to the question: is *Ideal Justice* informative in the sense of offering predictive and explanatory power? The answer, I believe, is yes. A model of a world in which there is strict compliance with the ideal principles of justice can be informative in the sense of offering predictive and explanatory power in the actual world. As Ismael notes, there are many ways idealized models can be informative. The use of ideal models, such as those in which strict compliance is used as a modeling assumption, allows “us to isolate certain relationships, suppress complicating factors that we are not interested in, and explore in isolation features of the world that always come together in practice” (Ismael 2016: 16). In *Ideal Justice*, the relationship isolated is between justice and compliance, as there are many other potentially relevant parameters that are not under inspection. The complicating factor being suppressed is non-compliance by self-directed, rather than justice-directed agents.[[19]](#footnote-19) The result is the parameter of justice is isolated in the model – though, of course, and no one denies, justice and compliance always come together in practice.

An example may help illustrate the way in which an ideal model can be relevantly informative. Suppose that the highest ranking principles of justice in the priority ordering constructed in the ordering stage require the institutional implementation of a Carens Market which holds that “everyone should be taxed in such a way that everyone’s after-tax income is equal; however, all should nevertheless voluntarily strive to produce as much pretax income as possible, out of a sense of social duty.”[[20]](#footnote-20) Let us first consider a way in which the ideal model would *not* be informative in the actual world. The strict compliance world would not be informative as a prescription of what ought to be immediately implemented. Schmidtz would certainly be correct that at present it would be a mistaken to implement a Carens Market because individuals would likely respond as we expect them to – namely, work less hard. This would be an undesirable result and precisely why we need the selection stage. Ideal and nonideal theorists agree on as much.

Where is the disagreement, then? I believe that it concerns whether a strict compliance world that will not obtain can serve as the standard-bearer for measuring what qualifies as a transitional step to an increasingly just world. To illustrate this point, consider an analogy with the standard in golf. The ideal score on any hole in golf is one. Assume that no golfer will ever score a hole-in-one on a particular hole; it is simply too demanding.[[21]](#footnote-21) The fact that golfers will fail to score a hole-in-one are like the agents who will fail, *ex hypothesi*, to develop the motivational motivations needed to make a Carens Market work.[[22]](#footnote-22) Given this reality, a Carens Market should not be what is opted for in the selection stage. As noted above, Schmidtz and the ideal theorist will agree about this. The disagreement will concern whether knowing that the standard is one is informative.

The ideal theorist will hold that the knowledge of the standard *orients* one’s playing of the hole: the aim is to get as close to one as possible. This does not change even for holes where it is accepted that the ideal score will not be achieved. Notice this does not mean that knowing the standard is directly *action-guiding* in the sense that the golfer should intentionally attempt to hit a hole-in-one. Given a golfer’s abilities, a less ambitious strategy will often be what is appropriate. Very good golfers should strategize to play in ways in which their scores will closer approximate one, while worse golfers ought adopt strategies that reduce the probability of producing a bad score even if it means settling for a good but not great score. Likewise, societies compromised of agents that are more just should strive to closer approximate – leaving precisely how close this approximation should be intentionally ambiguous and to be determined by the particular features of the society – a Carens Market, while societies compromised of agents that are less just are better off opting for a more realistic alternative in the selection stage. Still, this latter selection should be made with an eye, in part, towards institutions that will cultivate the moral motivations of the citizenry required to eventually make a Carens Market a realizable alternative.[[23]](#footnote-23) This point holds irrespective of what is the content of the ideal. I think a Carens Market is not a desirable – let alone the ideal – institutional implication of the principles of justice. One can insert different institutional implications ranging from G.A. Cohen’s (2009) camping trip to Jason Brennan’s (2014) utopian Mickey Mouse Clubhouse as one likes.

By contrast, Schmidtz would seem to think that knowing the standard is one is irrelevant as it is, by assumption, unattainable.[[24]](#footnote-24) Since it is unattainable, it cannot be the relevant standard for measuring performance. What we need to know and evaluate performance with respect to is *par*. In golf, par measures what a good player would score on a given hole. This is importantly different from what a perfect player – whatever that might mean – would score. Perfect players, Schmidtz might say, do not exist. Extrapolating from the analogy, Schmidtz includes concerns about compliance in the determination of what is the standard for justice. An informative standard of justice is one that real agents can achieve. Hence, the ordering stage, in which strict compliance is assumed, of *Ideal Justice* is irrelevant because it does not create a priority ordering of principles of justice that would be informative *for real agents*.

However, Schmidtz’s motivation for this view remains unclear to me. It seems odd, say, to suggest that the standard in golf should be lowered from one to par simply because golfers were unable to achieve it. Knowing that the standard is one and not par seems of paramount importance for strategizing how to play the hole.[[25]](#footnote-25) Indeed, lowering the standard changes the *incentive structure* of the game. If a plus-handicap golfer (one who routinely shoots *better* than par, such as your favorite professional player) has the standard set at par and not one, what reason is there to attempt to shoot better than par if it is not rewarded? I cannot think of any that would withstand reflection. This restructuring of the standards of the game would undermine it.

The same seems to be true of justice. Schmidtz’s account creates the wrong incentive structure for the most just agents. By tying compliance so deeply with justice, the resulting lowered normative bar has as an externality the disincentivizing of optimally just behavior by society’s most just agents. This sort of incentive problem should be avoided in golf, in the market, and in theorizing about justice. Inability to reach a standard does not mean that knowledge of the standard is not informative in pursuit of it.[[26]](#footnote-26) Of course, how to optimally pursue the optimal outcome is a separate and considerably more epistemically difficult task – what I called at the outset the *Ideal Transition* model. There is a reason that though I know the standard in golf is one, I have never hit a hole-in-one.

I conclude this section on relevance with a brief discussion of Ismael’s remarks on the different purposes the ideal principles of justice can have and which purpose is the relevant one. Ismael (2016: 23–24) writes:

Rawls should have been firmer about the distinction between an ideal conception as (i) a gauge for how just a society is, (ii) the model or template on which an actual society is to be built, and (iii) something to “aim at” in taking steps to make an unjust society more just. He should have endorsed (i), but not (ii) or (iii).

Put differently, Ismael’s identifies three purposes for the ideal principles of justice: (i) as a standard of normative measurement; (ii) as design principles for building the optimal society for actual agents; and (iii) as a normative target. Ismael is undoubtedly correct that Rawls should have been clearer about the purpose of ideal justice. However, Ismael too quickly dismisses (ii) and (iii). Though I do not want to defend either as stated, both are defensible upon clarification.

Purpose (i) was discussed in the preceding analysis of Schmidtz’s criticism. I will not comment further. The present focus will be on purposes (ii) and (iii). Admittedly, I am unclear on Ismael’s argument against (ii). The general worry seems to be that using the ideal principles of justice as design principles will lead to poorly fitted institutions for the agents that must actually operate within them. If this is correct, then a further distinction needs to be drawn about the model in question. Consider the following two claims. The first is that we can construct a model of the optimal society for real agents. The second is that we can construct a model of the optimal society for real agents *and* we ought to immediately begin implementing that model. It requires this second claim to generate Ismael’s skepticism about purpose (ii), but, as noted earlier, this is not the claim ideal theorists like Rawls and Simmons are making. Ideal theorists are only committed to the first claim.[[27]](#footnote-27) Thus, using the ideal principles of justice to construct a model on the basis of which the optimal society is to be built should not seem problematic. A problem only emerges if one begins building prematurely.

A similar concern applies to (iii). Ismael’s worry seems to be that ideal justice should not be used as a target to guide the transition of society to increasingly just social worlds.[[28]](#footnote-28) To assess whether Ismael’s worry is well-founded requires that we make more precise what it would be for a target to guide. If what is meant by this is that the target gives a conclusive and perfectly accurate set of steps that should be taken in the present to reach it, then I sympathize whole-heartedly with the worry. However, a target seems able to guide, at least in principle, in a less demanding sense by telling us, as Simmons (2016: 88) says, “what is necessary in order to perfect…the “rightfulness” of our political condition.” That is, an ideal target could guide in the sense of giving us something to, as Ismael says, “aim at” in that ideally we would end up there *without* guiding in the sense of giving us a set of directions for how to get there. The highest-ranking alternative in the ordering stage could orient our aim, but knowing how to bring it about could be out of our epistemic purview. In this respect, I am more skeptical about the ability of an ideal target to guide-action than ideal theorists like Rawls and Simmons, but more optimistic than Ismael in believing that an ideal conception of justice can give us something at which to aim.

**VI. Does Assuming Strict Compliance Beg the Question?**

Jacob Levy and Christopher Freiman each argue that strict compliance assumes away the very problem that the institutionally just state is supposed to solve.[[29]](#footnote-29) Levy (2016: 313–14) puts the point metaphorically: “political life is *about* friction: no friction, no politics or justice.” Assuming away the friction assumes away the need for political justice. Freiman (2017: 11) calls the assumption of strict compliance a “*self-obviating idealization*“ which is “when a model of x is idealized such that it assumes away the problem x is intended to solve.”

Both Levy and Freiman interpret ideal theory – by which they mean any model social world that assumes strict compliance – as being logically inconsistent.[[30]](#footnote-30) Levy (2016: 314) writes:

The ideal theories that appear relevant to political life necessarily smuggle in nonideal premises in order to justify the need for politics and justice altogether. Those that fail to do so also fail to be relevant, at best collapsing into a moral theory that lies across an unbridged gap from an articulation of political ideals of justice.

Freiman (2017: 7) identifies the same tension:

[I]deal theory houses an internal inconsistency. On the one hand, we must assume that people are *not* fully just to generate a need for the state in the first place. On the other hand, if people are not fully just, the state itself won’t be fully just either (it’s an institution run by people, after all.) In short, the assumption that generates a need for the state – that people aren’t fully just – at the same time undermines the assumption that the state is fully just.

The argumentative upshot of these remarks is that in order for a just state – which is essentially a dispute-resolver or conflict-resolver – to be necessary, there needs to be conflict.[[31]](#footnote-31) If people strictly comply with justice, then there will be no such conflict. Strict compliance assumes that people always do what is just. So, there are no conflicts for the state to solve, thereby, making justice or a just state superfluous. Elsewhere, I have called this the Levy-Freiman argument (Carroll Forthcoming).

The part of this argument that requires investigation is the conditional: if people strictly comply with justice, then will be no disputes or conflicts in need of resolution. The conditional is false when strict compliance with justice is assumed to obtain, but there remain conflicts for the state to resolve. Given that the antecedent is true by assumption, the truth of the conditional turns on the truth of consequent. Will there be conflicts for the state to resolve when strict compliance is assumed to obtain? There are, at least, two reasons for thinking that the answer is yes.[[32]](#footnote-32)

The first reason concerns the lack of specificity about the object of strict compliance. Assuming that there is “strict compliance on justice” typically means on the principles of justice. If this is the case, then there remains room for disagreement. Suppose Anita and Brian agree that the set *E* represents the correct set of principles of justice with which to comply. That is, *E* is the highest-ranking option in the priority ordering constructed in the ordering stage. However, Anita believes realizing *E* requires implementing redistributive socialist institutions, while Brian believes realizing *E* requires implementing free-market libertarian institutions.[[33]](#footnote-33) Anita and Brian’s disagreement about the proper institutional form of *E* is an instance of the type of dispute a state could have a role in resolving. One possible resolution would be putting the issue to a vote. An election will most likely introduce the need for a state to administer the election.

Defenders of the Levy-Freiman argument might respond that it is not just the principles of justice that are being strictly complied with, but all of justice including the principles and the institutional implementation. This maximally precise specification of the object of strict compliance moves us, though further from Rawls, closer to vindicating the consequent of the conditional, at least when paired with significant epistemic idealizations. Is this enough to vindicate the consequent? I think not. For why, we turn to the second reason.

The second reason concerns the scope of conflicts that fall under the aegis of the state. While we obviously do not want the state settling disputes, say, between friends about what television show to watch, it is also clear that justice does not subsume all of socio-political morality. Though Rawls (1999a: 3) says that “justice is the first virtue of social institutions,” he does not say it is the only virtue. In fact, it seems to imply that there are others. Where to draw the boundaries on justice is unclear. Nonetheless, important values like care (Held 2007), compromise (Wendt 2016), peace (Gray 2002; Kukathas 2003), social trust (Vallier 2019) and various others do not seem to be wholly subsumed by considerations about justice and relate to issues in which the state could have a role.[[34]](#footnote-34) Even when everyone knows and does what justice demands, there is nothing that precludes individuals from disagreeing about the rest of socio-political morality as long as it does not result in one acting unjustly. Such disagreements may concern what are the other values as well as the relative weights of these values. If non-justice, political and social disagreements can occur, there can be conflicts that the state can, in principle, play a role in resolving. Whether the state is better suited to resolve such conflicts than individuals on their own would depend on a more complete account of the situation. Importantly, this is a concern that emerges *after* the consequent is shown to be false.

On the basis of these two reasons, the consequent appears false. Insofar as the consequent of is false, then the conditional is false. If this is correct, the Levy-Freiman argument is unsound. Using strict compliance as a modeling assumption does not entail that the resulting model is question-begging by stipulatively rendering the state otiose because strict compliance only assumes away injustice, not all of political disagreement.

**VII. Conclusion**

A theme in critiques of employing strict compliance as a modeling assumption is that it is somehow distortionary. Some, like Rosenberg, suggest that the distortion comes as a result of attempting to make determinate a problem that is simply wrought with indeterminacy. Others, like Schmidtz, maintain that the distortion follows either from attempting to parse two things that cannot be coherently parsed – compliance and justice – or from the assumption being so idealistic that the resulting solution becomes irrelevant to the actual problem. While those like Levy and Freiman hold that assuming strict compliance distorts the very conceptualization of what the problem is by assuming the problem itself away.

I have argued that *Ideal Justice* does not succumb to any of these distortions and offers a way of providing content to *Ideal Destination*. This does not mean that *Ideal Justice* offers either the only way or the best way to employ strict compliance as a modeling assumption. But it does offer *a* way.[[35]](#footnote-35)

Jeffrey Carroll

Department of Philosophy

University of Virginia

jmcarroll422@gmail.com

REFERENCES

Aligica, Paul Dragos, Peter J. Boettke, and Vlad Tarko. 2019. *Public Governance and the Classical-Liberal Perspective: Political Economy Foundations*. Oxford University Press.

Anderson, Elizabeth. 2009. “Toward a Non-Ideal, Relational Methodology for Political Philosophy: Comments on Schwartzman’s Challenging Liberalism.” *Hypatia* 24 (4): 130–45.

Brennan, Jason. 2014. *Why Not Capitalism?* New York: Routledge.

Carens, Joseph. 1981. *Equality, Moral Incentives, and the Market*. Chicago: University of Chicago Press.

Carroll, Jeffrey. Forthcoming. “Is Ideal Theory Anarchic?” *The Journal of Politics*.

Cohen, G. A. 2009. *Why Not Socialism?* Princeton: Princeton University Press.

Cowen, Tyler. 2007. “The Importance of Defining the Feasible Set.” *Economics and Philosophy* 23 (1): 1–14.

Edmundson, William A. 2017. *John Rawls: Reticent Socialist*. New York: Cambridge University Press.

Ellis, Elisabeth. 2005. *Kant’s Politics: Provisional Theory for an Uncertain World*. New Haven, Conn: Yale University Press.

Erman, Eva, and Niklas Möller. 2013. “Three Failed Charges Against Ideal Theory.” *Social Theory and Practice* 39 (1): 19–44.

Estlund, David. 2007. *Democratic Authority: A Philosophical Framework*. Princeton: Princeton University Press.

———. 2011. “Human Nature and the Limits (If Any) of Political Philosophy.” *Philosophy and Public Affairs* 39 (3): 207–237.

———. 2014. “Utopophobia.” *Philosophy & Public Affairs* 42 (2): 113–34.

———. 2016. “What Is Circumstantial About Justice?” *Social Philosophy and Policy* 33 (1–2): 292–311.

———. 2017a. “Prime Justice.” In *Political Utopias: Contemporary Debates*, edited by Michael Weber and Kevin Vallier, 35–56. New York, NY: Oxford University Press.

———. 2017b. “The Ideal, the Neighborhood, and the Status Quo: Gaus on the Uses of Justice.” *Ethics* 127 (4): 912–28.

Freiman, Christopher. 2017. *Unequivocal Justice*. New York: Routledge.

Gaus, Gerald. 2016. *The Tyranny of the Ideal: Justice in a Diverse Society*. Princeton: Princeton University Press.

Gaus, Gerald, and Keith Hankins. 2017. “Searching for the Ideal: The Fundamental Diversity Dilemma.” In *Political Utopias: Contemporary Debates*, edited by Kevin Vallier and Michael Weber. New York: Oxford University Press.

Gaus, Gerald, and John Thrasher. 2016. “Rational Choice in the Original Position.” In *The Original Position*, edited by Timothy Hinton. New York: Cambridge University Press.

Gray, John. 2002. *Two Faces of Liberalism*. Oxford: The New Press.

Guerrero, Alex. 2017. “Political Functionalism and the Importance of Social Facts.” In *Political Utopias*, edited by Michael Weber and Kevin Vallier, 127–50. New York: Oxford University Press.

Hamlin, Alan, and Zofia Stemplowska. 2012. “Theory, Ideal Theory and the Theory of Ideals.” *Political Studies Review* 10 (1): 48–62.

Held, Virginia. 2007. *The Ethics of Care: Personal, Political, and Global*. Oxford: Oxford University Press.

Hinton, Timothy, ed. 2016. *The Original Position*. New York: Cambridge University Press.

Huemer, Michael. 2013. *The Problem of Political Authority: An Examination of the Right to Coerce and the Duty to Obey*. New York: Palgrave Macmillan.

———. 2016. “Confessions of a Utopophobe.” *Social Philosophy and Policy* 33 (1–2): 214–34.

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

Kogelmann, Brian. 2017. “Justice, Diversity, and the Well-Ordered Society.” *The Philosophical Quarterly* 67 (269): 663–84.

Kogelmann, Brian, and Gerald Gaus. 2017. “Rational Choice Theory.” In *Research Methods in Analytic Political Theory*, edited by Adrian Blau, 217–42. Cambridge: Cambridge University Press.

Kukathas, Chandran. 2003. *The Liberal Archipelago*. Oxford: Oxford University Press.

Levy, Jacob T. 2016. “There Is No Such Thing as Ideal Theory.” *Social Philosophy and Policy* 33 (1–2): 312–33.

Lipsey, R. G., and Kelvin Lancaster. 1956. “The General Theory of Second Best.” *The Review of Economic Studies* 24 (1): 11–32.

Mason, Andrew. 2010. “Rawlsian Theory and the Circumstances of Politics.” *Political Theory*, 2010.

———. 2016. “Justice, Feasibility, and Ideal Theory: A Pluralist Approach.” *Social Philosophy and Policy* 33 (1–2): 32–54.

Mills, Charles W. 2005. “‘Ideal Theory’ as Ideology.” *Hypatia* 20 (3): 165–184.

Neufeld, Blain. 2011. “Review of ‘Why Political Liberalism?’ On John Rawls’s Political Turn.” Electronic Journal. Notre Dame Philosophical Reviews. 2011.

Nili, Shmuel. 2018. “The Moving Global Everest: A New Challenge to Global Ideal Theory as a Necessary Compass.” *European Journal of Political Theory* 17 (1): 87–108.

Rawls, John. 1999a. *A Theory of Justice: Revised Edition*. Cambridge, Mass: Belknap Press.

———. 1999b. *The Law of Peoples: With “The Idea of Public Reason Revisited.”* Cambridge, Mass: Harvard University Press.

———. 2003. *Justice as Fairness: A Restatement*. Edited by Erin Kelly. Cambridge, Mass: Belknap Press.

Ronzoni, Miriam. 2014. “The Four-Stage Sequence.” In *The Cambridge Rawls Lexicon*, edited by Jon Mandle and David A. Reidy, 290–92. Cambridge: Cambridge University Press.

Rosenberg, Alexander. 2016. “On the Very Idea of Ideal Theory in Political Philosophy.” *Social Philosophy and Policy* 33 (1–2): 55–75.

Schmidtz, David. 2011. “Nonideal Theory: What It Is and What It Needs to Be.” *Ethics* 121 (4): 772–796.

———. 2016. “A Realistic Political Ideal.” *Social Philosophy and Policy* 33 (1–2): 1–10.

———. 2017. “Realistic Idealism.” In *Methods in Analytical Political Theory*, edited by Adrian Blau, 131–51. New York, NY: Cambridge University Press.

Sen, Amartya. 2011. *The Idea of Justice*. Cambridge, Mass: Belknap Press.

Simmons, A. John. 2010. “Ideal and Nonideal Theory.” *Philosophy & Public Affairs* 38 (1): 5–36.

———. 2016. *Boundaries of Authority*. New York: Oxford University Press.

Smith, Adam. 1982. *The Theory of Moral Sentiments*. Edited by D. D Raphael and A. L Macfie. Indianapolis: Liberty Fund, Inc.

Southwood, Nicholas. 2018. “The Feasibility Issue.” *Philosophy Compass* 13 (8): 1–13.

Stemplowska, Zofia. 2008. “What’s Ideal About Ideal Theory?” *Social Theory and Practice* 34 (3): 319–340.

Stemplowska, Zofia, and Adam Swift. 2012. “Ideal and Nonideal Theory.” In *Oxford Handbook of Political Philosophy*, edited by David Estlund, 373–89. Oxford: Oxford University Press.

———. 2013. “Rawls on Ideal and Nonideal Theory.” In *A Companion to Rawls*, edited by John Mandle and David A. Reidy, 112–27. Oxford: Wiley Blackwell.

Thrasher, John, and Kevin Vallier. 2013. “The Fragility of Consensus: Public Reason, Diversity and Stability.” *European Journal of Philosophy*, 1–22.

Tomasi, John. 2013. *Free Market Fairness*. Princeton: Princeton University Press.

Vallier, Kevin. 2019. *Must Politics Be War? Restoring Our Trust in the Free Society*. Oxford: Oxford University Press.

Waldron, Jeremy. 1999. *Law and Disagreement*. Oxford: Oxford University Press.

Watson, Lori, and Christie Hartley. 2018. *Equal Citizenship and Public Reason: A Feminist Political Liberalism*. New York: Oxford University Press.

Weisberg, Michael. 2007. “Three Kinds of Idealization.” *Journal of Philosophy* 104 (12): 639–659.

———. 2015. *Simulation and Similarity: Using Models to Understand the World*. Oxford: Oxford University Press.

Weithman, Paul. 2013. *Why Political Liberalism?: On John Rawls’s Political Turn*. Reprint edition. New York; Oxford: Oxford University Press.

Wendt, Fabian. 2016. *Compromise, Peace and Public Justification: Political Morality Beyond Justice*. Palgrave Macmillan.

Wiens, David. 2012. “Prescribing Institutions Without Ideal Theory.” *Journal of Political Philosophy* 20 (1): 45–70.

———. 2015. “Against Ideal Guidance.” *Journal of Politics* 77 (2): 433–446.

1. For defenses of the priority of ideal theory, see Simmons (2010); Stemplowska (2008); Stemplowska and Swift (2012); Hamlin and Stemplowska (2012); and Stemplowska and Swift (2013). David Wiens (2012; 2015), Gerald Gaus (2016) along with Keith Hankins (2017), Jacob Levy (2016) and Christopher Freiman (2017) have, in different ways, criticized the priority of ideal theory. [↑](#footnote-ref-1)
2. Mills (2005). Eva Erman and Niklas Möller (2013) offer, what I take to be a highly compelling response to Mills. See also Watson and Hartley (2018: chap. 1). [↑](#footnote-ref-2)
3. Ideal theory, according to Rawls (1999a: 216) specifies “a conception of a just society that we are to achieve if we can.” [↑](#footnote-ref-3)
4. Simmons (2010: 8) writes: “if we compare the operation of societies ordered by competing principles of justice while assuming strict compliance with these principles, the different effects we observe can reasonably be taken to be wholly the responsibility of the different ordering principles themselves.” [↑](#footnote-ref-4)
5. For more on the original position, see Hinton (2016). [↑](#footnote-ref-5)
6. Nothing hangs on what principles of justice are under consideration. Not only can different principles of justice be used as inputs, but this model works if one deems it necessary to evaluate whole social worlds rather than the principles of justice that are used to order them. Thanks to Wes Siscoe for a helpful discussion on the scope and potential of the model. [↑](#footnote-ref-6)
7. Indeed, it is possible that the ordering stage proves largely unhelpful for the selection stage. As Rawls (1999a: 215–16) says: “even granting the soundness of these principles…we must still ask how well they apply to institutions under less than favorable conditions, and whether they provide any guidance for instances of injustice. The principles and their lexical order were not acknowledged with these situations in mind and so it possible that they no longer hold.” [↑](#footnote-ref-7)
8. There may be disagreement about what are the correct principles of justice. This has come to be called the problem of justice pluralism. See Kogelmann (2017). [↑](#footnote-ref-8)
9. For an illuminating discussion on the different ways Rawls models the choice problem throughout his work, see Gaus and Thrasher (2016). [↑](#footnote-ref-9)
10. The rejection of this type of ideal theorizing amounts to the rejection of the approaches defended by Simmons and Estlund. [↑](#footnote-ref-10)
11. Elizabeth Anderson also advocates a scientific methodology as the model for political theorizing. Interestingly, Ismael thinks a scientific perspective leads us to ideal theory, whereas Anderson sees it as being the model for non-ideal theory. See Anderson (2009: 135). [↑](#footnote-ref-11)
12. Or at least the potential for it. Simmons (2010, 36). [↑](#footnote-ref-12)
13. For an interesting paper on methodological pluralism, albeit of a different variety, see Mason (2016). [↑](#footnote-ref-13)
14. Emphasis added. [↑](#footnote-ref-14)
15. Simmons (2010: 8–9) recognizes that if we assume “normal” levels of noncompliance, “we will likely find both that our evaluations yield quite indeterminate results and that the results depend on more than simply the different ordering effects of the principles being compared.” [↑](#footnote-ref-15)
16. Let us grant that an assumption of superabundance would assume away the problem being solved by justice. Admittedly, I am skeptical that Rosenberg is correct about this. See Estlund (2016; 2017a). [↑](#footnote-ref-16)
17. One way of conceptualizing this challenge is in terms of a x-y plane with justice on the x-axis and compliance on the y-axis. One option is to stipulate that the y-axis is maximized and the relevant optimization problem is just how to maximize the x-axis. A different approach holds that the x-axis and the y-axis are interdependent. The relevant task is to determine how to optimize over both dimensions. Schmidtz defends *exclusively* the latter option, while the ideal theorist engages in the former in the ordering stage and the latter in the selection stage. Thanks to Marcus Hunt for a helpful discussion. [↑](#footnote-ref-17)
18. As Schmidtz (2016: 7) puts it, “To choose an incentive structure is to choose a compliance problem. To set aside our chosen compliance problem, as a detail best ignored, is to set aside the nature of what we are choosing as a detail best ignored.” [↑](#footnote-ref-18)
19. Ismael (2016: 16) makes a similar point: “The full development of a complete theory of justice should have the resources to deal with the effects of noncompliance, but the ideal part of the theory is the part of the theory that most clearly displays the *content* of, and *justification* for, [Rawls's] notion of justice.” [↑](#footnote-ref-19)
20. This description comes from Huemer (2016). For the full account in full, see Carens (1981). [↑](#footnote-ref-20)
21. Of course, this claim of demandingness is context-sensitive. Likewise with justice, a different social world may make it easier to generate compliance on principles of justice. [↑](#footnote-ref-21)
22. Schmidtz might maintain that agents will never be sufficiently morally motivated for a Carens Market to work, while the ideal theorist might leave it as an open question. It is being granted that Schmidtz is right on this point. [↑](#footnote-ref-22)
23. One might worry that concerns informed by the theory of second best loom large here. This is a real worry. Determining how to weigh institutions that will bring about the ideal outcome against institutions that will merely bring about the local best outcome must be done on a case-by-case basis. Often, we lack the epistemic resources to be able to make anything close to a fully informed decision. Any *a priori* declaration would be too far-reaching. For the seminal essay on this topic, see Lipsey and Lancaster (1956). [↑](#footnote-ref-23)
24. A stronger claim that Schmidtz could make is that there is no such justice-analog of one in golf. However, this would be a much stronger claim about the metaphysics, as opposed to merely the methodology, of justice and would require an argument in defense of it. [↑](#footnote-ref-24)
25. Though I agree with Schmidtz that it is also of paramount importance that one tailors one’s strategy for playing the hole to one’s abilities. This is the role of the selection stage in the model. [↑](#footnote-ref-25)
26. A parallel point has been made in many different guises by David Estlund (2007: chap. 14; 2011; 2014; 2016; 2017a; 2017b). [↑](#footnote-ref-26)
27. Strict speaking, ideal theorists need not even be committed to the first claim. Only ideal theorists that follow in the Rawlsian legacy and give theoretical primacy to the *Ideal Destination* modeling task are so committed. [↑](#footnote-ref-27)
28. This point has been forcefully made by Wiens (2015) and Gaus and Hankins (2017). [↑](#footnote-ref-28)
29. Drawing on the work of Jeremy Waldron (1999), Andrew Mason (2010) offered an earlier version of the problem Levy and Freiman investigate in detail. [↑](#footnote-ref-29)
30. I believe Levy and Freiman underappreciate the variety of ideal theory by tying it so tightly to the assumption of strict compliance. Two of the most prominent ideal theorists – G.A. Cohen and David Estlund – are not defined in terms of the strict compliance assumption. [↑](#footnote-ref-30)
31. Perhaps one would argue that the role the state plays is not as a conflict-resolver, but as an assurance-provider. For extensive discussions of assurance, see Weithman (2013), Neufeld (2011), and Thrasher and Vallier (2013). [↑](#footnote-ref-31)
32. This is not to say that the state will be a better resolver of those conflicts than private institutions. See Huemer (2013). [↑](#footnote-ref-32)
33. Though one may think the example is outlandish, for a defense of libertarian institutions and Rawlsian principles, see Tomasi (2013). [↑](#footnote-ref-33)
34. Unless we define ‘justice’ as referring to ‘the resolution of conflicts under the aegis of the state’ and stipulate that the set of conflicts that fall under the aegis of the state extremely capaciously. However, this strategy seems to beg the question. [↑](#footnote-ref-34)
35. Thanks to Matthew Adams, Jerry Gaus, Brian Kogelmann, Loren Lomasky, John Simmons, Kevin Vallier and anonymous reviewers with the journal for either helpful comments on or fruitful discussion of earlier versions of this paper. One such version was given at the PPE Society Conference in March 2018 in which Amy Berg, Harrison Frye, David McCabe and David Wiens all provided constructive feedback. A special thanks is owed to Greg Robson for providing detailed comments on the entire manuscript. [↑](#footnote-ref-35)