May 15th, 9:00 AM - May 17th, 5:00 PM

The Normative Impotence of Ideal Models

John Woods
THE NORMATIVE IMPOTENCE OF IDEAL MODELS

John Woods
Department of Philosophy
University of Lethbridge
©1998, John Woods

Abstract:
In the methodology of theory construction, the concept of "intuitions" is commonly assigned a central role. This is especially true of philosophical and social scientific theories or rational human agency. An equally important trait of such accounts is the theorist's employment of "ideal models" or rational agency. It is frequently supposed that the concept of intuitions and the concept of ideal models link in such a way as to give rise to a coherent and load-bearing notion of "objective normativity." This paper shows, with reference to a wide range of contemporary theories, (a) that the employment of ideal models is otiose, and (b) that the supposedly related concept of objective normativity is groundless.

***

What are the procedures and methods by which an agent performs rationally? Under what conditions is it rational to re-enter a burning building, or to buy a used Pinto (or a used Hansen & Pinto), or to remind your opponent that he too smokes, or to detach a proposition on the warrant of modus ponens?

What would it be like to have a theory that answered such questions persuasively and authoritatively? A common suggestion is that the theorist of rational performance proceeds in two main stages. At stage one he marshalls his "intuitions" about what it would be rational to do in certain cases or rough ranges of cases. Intuitions are "untutored"; they are pre-theoretical beliefs which the theorist holds with conviction and which, as he assumes, are widely shared by others. They are what the theorist supposes everyone with the relevant curiosity and attentiveness already knows about the issues at hand. They are the data for his theory.

At the second stage, the theorist ventures beyond. He tries to generalize on his data, to link them up in nomically attractive connections. However this is done, and whatever else the theorist is doing at stage two, it is essential to the enterprise that cases not reported in the data be dealt with so that the theorist's generalization can be taken as covering unobserved instances and predicting future instances. Stage two also provides the wherewithal to reconsider the data in a principled way. If, as they emerge, the theory's provisions appear to require the revision or abandonment of a pre-theoretical belief, the theorist is at liberty to do so. This is a third function performed by the theory's laws (or, to be careful, "laws"). They enable us to correct pre-theoretical beliefs, if not retire them wholesale.

It is rather striking that the theoretical "laws" of accounts of rational performance fail the prediction test rather substantially. People don't actually do what the "laws" can be taken as saying they will do. This presents the enquirer with a genuinely interesting problem in the methodology of theory construction. There are situations in which the theorist feels himself justified in refusing the status of counterexample to actual instances which contradict his "laws". The problem is one of showing that he is justified in this refusal. To that end, the theorist will commonly invoke the notion of approximations. In mechanics, the laws of frictionless surfaces fail even on the pre-game ice of Maple Leaf Gardens. The physicist says that such laws are true at a mathematically describable
but physically unrealized limit of the slipperiness of real life, and to which the slippery world approximates only more or less well. In probabilistic inference theory, laws relating to the conjoining of probabilities are also routinely dis honoured in practice. Rather than abandon the calculus of probability, the theorist invokes the device of an ideal agent, and describes his behaviour in an ideal model in which, he deems himself free to say, the conjunction rule is universally obeyed. The theorist then ventures that an actual reasoner is rational to the extent that his behaviour approximates to that of the ideal agent. Here is Keynes on this point:

Progress in economics counts almost entirely in a progressive improvement in the choice of models...But it is of the essence of a model that one does not fill in real values for the variable function. To do so would make it useless as a model... The object of statistical study is not so much to fill in missing variables with a view to prediction, as to test the relevance and validity of the model. (Letter to Harrod, 4 July 1938, in Keynes [1973], vol. 14).

There are, of course, more ideal models than one can shake a stick at. How does the theorist choose among rival models? How does he test for "relevance and validity"? Consideration of two cases might throw some light on the matter. In the first case, as with probability theory or first order logic, the theorist already knows (or thinks he does) that his axioms are true, that his derivation strategies truth-preserving, and so on. This creates an abiding temptation to specify as ideal any least model preserving those axioms and strategies. They are the "laws" of his theory, obeyed without exception by the ideal agent (the ideal probabilizer; the ideal logician, etc.) The second case is trickier. The theorist's judgement, in economics, for example, as to which theoretical "laws" are true is inseparable from his choice of the model in which they are true. It is interesting that in neither case need the theorist be worried about the underdetermination of theories by observation, or by the fact that no (first order) theory determines its objects up to isomorphism. The problems raised by our cases are not primarily ontological. In part they are semantic. They are the problem of finding models in which favoured "laws" come out true. But the more central case-two problem is that of validating the model independently of having to demonstrate the truth of the "laws" that describe it (and which actual agents routinely shirk). Suppose that the theorist is trying to construct a theory of argument. Waiving the implausibility of it, what if he is drawn to the following "law":

(L): If X and Y are opponents in an argument, and if X asserts that P, Y will beg the question with respect to P.

L is sometimes honoured in actual practice, but mostly it is otherwise. L fails empirically. What is the theorist to say for himself? What if he says that, while false in practice, it is true in his ideal model? Critics won't like his choice of ideal model, of course; and their dislike of it will turn entirely on their conviction that L is a bad law—a faulty norm—conformity to which, or approximate conformity to which, would ensure an agent's lack of rationality. So, again, how is the case—two theorist to pick his "laws" from sets of empirically false generalizations, short of already having at hand a validated ideal model in which to rescue him from this very embarrassment? How, in turn, is he to pick a model as validated, short of stocking it with empirically false generalizations which, even so, have the property of "deserving", there, to be "laws" against which rational agency is measured? This is a difficulty serious enough to have a name. Let us dub it the bootstrapping problem.

Stipulation offers promise of a way out of the bootstrapping problem. The theorist now stipulates that his ideal model is constituted by the perfect conformity of its objects with those "laws" to which the theorist is drawn. They are now analytically true in the model (waiving for now philosophical reservations about analyticity). But with careless disregard for the secundum quid fallacy, the theorist forgets the qualification "in the model". He
puts it that his favoured "laws" are analytically true, and he concludes that an actual agent is rational to the extent that his behaviour complies with them. For how can an agent be rational if he dishonors analytic truths? 5

Such seems to be the position which ideal model theorists find themselves in, in certain branches of discourse analysis and the theory of argument. Consider:

In order to clarify what is involved in viewing argumentative discourse as aimed at resolving a difference of opinion, the theoretical notion of a critical discussion is given shape in an ideal model specifying the various stages in the resolution process and the verbal moves that are instrumental in each of these stages. The principles authorizing these moves are accounted for in a set of rules for the performance of speech acts. Then together, these rules constitute a theoretical definition of a critical discussion. 6

Moreover,

[a]nalyzing argumentative discourse amounts to interpreting it from a theoretical perspective. This means that the interpretation is guided by a theoretically motivated model that provides a point of reference for the analysis. . . As a consequence of the adoption of such a theoretical perspective, specific aspects of the discourse are highlighted in the analysis (ibid, 191)... [And so], [e]ven a discourse which is clearly argumentative will in many respects not correspond to the ideal model of a critical discussion.... (ibid, 299, n. 49)... In this model, the rules and regularities of actual discourse are brought together with normative principles of goal-directed discourse. The model of critical discussion is an abstraction, a theoretically motivated system for ideal resolution-oriented discourse (ibid, 311; emphases added).

Case one stands out in apparent relief from these difficulties. We do indeed know (or so it is said) that the laws of logic and probability theory are analytic; and we know these things without the nuisance of having to specify ideal models in which they are true by stipulation. Here all talk of ideal agents in ideal models is idle. An actual agent is rational to the extent that he obeys, e.g., the analytic truths of logic and probability theory. Full stop.

It won't work. The proposal is afflicted by well-rehearsed problems, and they need not detain us long. 7 The present story forces upon us rational performance regulae which, owing mainly to the licence they give to complexity, provide a conception of rationality which defeats any hope of a credible approximation relation connecting actual performance to it. For the regulae count as rational the ideal agent's perseverance with problems that are intractable, i.e., whose solutions involve a "computational explosion". The presumed conception of rationality is one full of fidelity to which it would be irrational for an actual agent even to aspire. Most case-one theorists see the difficulty sufficiently well to reconsider the device of ideal models. An ideal logician is no longer one whose behaviour wholly conforms to the laws of logic. His ideality consists in perfect fidelity to softer norms, which are restrictions of the laws of logic. In many writings, this is something to welcome; it is occasion to make something of the difference between implication and inference. The theorist will grant that the laws of logic correctly describe the (or some designated) implication relation, but, he will now say, few if any "laws" of deductive inference will be mere simulacra of the laws of implication. So he will propose restrictions and add qualifications. So placed, the case-one theorist finds that he is in the same difficulty as his case-two colleague. As with his colleague, his judgement as to which are universally correct "laws" of inference is indistinguishable from his judgement as to which norms define and validate his ideal model. Finding himself at risk for the bootstrapping problem, the case-one theorist is beset by similar temptations regarding a way out. If he
yields, he will stipulate an ideal model as one in which his favourite "laws" are made analytically true. Again, "analytically true" here means "true in the model", i.e., "stipulated in the ideal model". If the theorist concludes that his favourite "laws" are analytically true (full stop), he too has committed the howler of secundum quid. He need not, of course, be sprung in this trap. He may say simply that an actual agent is rational to the extent that his deductive inferences conform to the laws of the model. He could say this, right enough, but without independent reason to like the model, he won't have provided us with the slightest reason to believe him.

By and large, social scientists are not much drawn to epistemological questions, even to those that bear upon their own theoretical practice. When I say that an economist's idealizations or simplifications have the effect of making his laws analytic in the model, I don't mean that this is a consequence which the theorist always or even typically recognizes expressly. Indeed he will often, in effect, deny the analyticity of his laws by insisting on their descriptive character. This is explicable. Such laws are analytic in the model. They are not analytic.

It would be wrong to leave the impression that simulators of ideal models are tendentious cynics, that they are simply making it up as they go along. Rarely is it so. The theorist of ideal models is no trifler; his exactions are too heartfelt for that. He is not even taking cover in Quine's amusing dicta that theories are "free for the thinking up" and that they are "put up jobs". The "laws", norms, and performance standards with which our theorist stocks his model are precisely those that he thinks are sound. He puts them in because he believes that they are objectively correct. If he is philosophically-minded, he might explain himself this way: These "laws", norms, and standards arise from, and are validated by an analysis of the concept, of the very idea, of rational performance. The analytic truths of his model are now conceptual truths about rationality.

Perhaps this is so. Perhaps the canons of the model are indeed the correct ones; maybe they do express the conceptual content of the very idea of rational performance. What remains is to show that they do. Short of this, the theorist has surrendered his whole stake in the theory to the category of "data for theory", i.e., statements confidently believed pre-theoretically. This matters. If the theorist's "laws" are no more than what he takes to be so, we have no theory, and are met instead with the phoney imperiousness of what the theorist already believes. Having no theory, he is left with his intuitions.

It is a fair and deeply important question whether there could be anything deserving of the name of normative theory. I say "normative", since this is what we have been discussing all along. We imagined the theorist in quest of regulae that determine what human agents out and ought not do. The general answer was that an actual agent ought to approximate to conformity with the conceptual norms of ideal models, and to avoid disconformity. "Norms" is a heavy word. People are impressed by it. But, as we see, a theorist's norms are nothing but what upon reflection he thinks people should do.

In speaking of the phoney imperiousness of what the theorist believes, there is an understandable inclination to resist the imperiousness. But it is not intended that we make light of the beliefs. Often enough our theorist is sharper than the rest of us and sees what we have missed. Sometimes a hitherto unnoticed distinction will clear away some damaging confusion. There are those who think that so far as concern the conditions on human flourishing, the best that can done is be done dialectically, preferably at the elbow of a master. So: I may think that I should always try to close my beliefs under consequence. You might point ought how profligate and unpointful that would be. You will have pressed me with consequences of my own view and, not liking them, I will have started to see things your way. If this were the canonical way of proceeding across the whole spectrum of normativity, then we could say with certain of the ancients that theories aren't possible for such matters. It could then be proposed that we end the pretense of supposing otherwise—that we abandon all talk of
norms and ideal models as empty.  

This is too much for some people to bear, anyhow for some philosophers to bear. For what is a philosophical theory about rational performance if not a theory that tells us in a principled way what people should and should not do? On the present suggestion, there are no such theories; hence there are no philosophical accounts that tell us in a principled way what people should and should not do. Philosophy, i.e., normative philosophy, is out of business. There are, of course, philosophers galore for whom the very idea of the impossibility of philosophy is absurd. Perhaps this is right. Perhaps the very idea of the impossibility of philosophy is an idea to be scorned. Even if true, the disreputability of that idea is not self-guaranteeing. So the theorist will be obliged to construct a case for the disreputability of the very idea of the impossibility of normative philosophy. In work underway, for which there is no time here, I try to say why the theorist's case cannot successfully be constructed.

So here I must content myself with something more modest. The ideal model theorist is met with the bootstrapping problem. It is a real problem that does real damage. My conclusion is indeed a modest one, for it is a truism: The ideal model theorist must solve the bootstrapping problem. I do not think that he can succeed in this endeavour. I could be wrong, of course. I am ready to wait and eager to see what the ideal model theorist has to say.

That the bootstrapping problem is a problem for ideal model theorists may have little to do with how serious a problem it is, and nearly everything to do with their unawareness of it. I want now to consider an exception to this rather entrenched indifference, namely ideal model theories, such as neoclassical economics or decision theory, which are heavily mathematical in character.

So let us consider Subjective Expected Utility theories of rational choice in economics and decision theory. In these theories, rational deciders maximize their expected payoffs. A decider's utility function is something derived from patterns of choices under risk, or "gambles", and subjective probability is defined in turn in terms of these utilities. The *locus classicus* of Subjective Expected Utility theories is the axiomatization of Savage. Of the Savage axioms perhaps the most contentious are those proclaiming transitivity and independence as norms of decisional rationality or economic agency. The Subjective Expected Utility theorist shares in the difficulties already discussed. In particular, he tends to be drawn, however implicitly, to the following argument:

1. Certain Axioms on ideal agency constitute a definition of rationality.
2. An actual agent is irrational if, or to the extent that, his behaviour disconforms to the axioms (by the definition of rationality).
3. Consequently, axioms on ideal agency are norms for rational agency.

As before, the mistake occurs in (1). It may be that the axioms on subjective expected utility define rationality in the model. It doesn't follow that this is rationality for actual agents, that is, rationality outside the model against which actual agency should be appraised. Of course, it is true that something gets to *be* an axiom in the model because it encodes the theorist's prior belief that the content of the axiom is indeed normative for actual deciders. But making it an axiom in the model is not a way of making the prior belief true. It is just a way of re-expressing that belief sententiously. So here, too, the bootstrapping problem recurs.

There is an old joke to the effect that economists can't get the success of physics out of their minds, and that they won't rest until they are able to do for socio-economic nature what the physicist has done for nature herself. In reconciling his theories to the paradigm of physics, the economist assigns a load-bearing role to mathematics. Like the physicist, he is struck by the fact that "... beyond a minimal level, we do not know how to do natural
science without mathematics".  

Neoclassical economics is an instructive case in point. As is widely known, the neoclassical theory replaced the law of diminishing marginal utilities with the law of diminishing marginal rates of substitution. With the additional "simplification" that goods are infinitely divisible, the theory had direct access to the firepower of calculus and could be formulated mathematically.

Consider, too, as a second example from economics, the so-called semantic view of theory construction. On this view, theories are set theoretic predicates. It sees the physicist's theory of motion and gravitation not as statements about the world, but rather as an analysis of the predicate, "is a system". Such predicates, may of course, apply to the world. Let $C$ be a consequence of the of the system predicate. Then $C$ is assertible about the world under the empirical assumption that the universe is. It may seem to some that the theorist's attention to the analysis of predicates rather than to the description of nature is ultimately a diversion, since nature ends up getting described after all.

Why, then, bother with theories constrained by the semantic view? A good part of the answer is that it makes it "easier to reconstruct the claims of science in a rigorous or mathematical way if one employs the semantic view".

Indeed,

This kind of endeavor is particularly prominent in economics, where theorists devote a great deal of effort to exploring the implications of perfect rationality, perfect information, and perfect competition. These explorations are...what economists (but not econometricians) call "models". One can thus make good use of the semantic view to understand theoretical models in economics. (idem).

It is apparent that the mathematical economist sees thing in one or other of two ways:

**Ontologically.** Nature, or the natural world, is to a considerable extent the physical realization of mathematical systems. That is, a mature scientific theory could only be "formalized as an extension of some part of mathematics". Socio-economic nature is part of the natural world. So it too, to a considerable extent is a socio-physical realization of mathematical systems. In both cases, the more the theorist is able to mathematize his accounts, the closer they will bring him to the true nature of things. Therefore, if shaping the theory's idealizations in a certain way—even with apparently ad hoc latitude—enlarges the theory's mathematical power, then, other things being equal, the idealizations are to that extent justified.

**Pragmatically.** Let $T$ be a mathematizable theory. Suppose that a particular idealization enables $T$ to engage some mathematics in ways that make possible simpler and deeper laws, and more transparent, encompassing and unified nomic connections. Other things being equal, the idealization is justified because, or to the extent that, it facilitates the evolving of a better theory, a theory that handles its subject matter economically, deeply, and with a high degree of comprehensiveness.

There is much that can be said against these ways of thinking. Critics will bring charges ranging from neopythagoreanism to the fallacy of division, and thence to complaints of reductionist simple mindedness. I shall press none of these objections here. I am interested only in the point that when idealizations are made in order to facilitate mathematical power, then, other things being equal, the idealizations are justified; they take on, so to
speak, the objective validity of the mathematics that they were contrived for.

Whether he thinks of his deployment of mathematics ontologically or pragmatically the theorist chooses and adjusts his idealization to facilitate the mathematics. But this is not done in isolation from the requirements of empirical adequacy, as van Fraassen calls it. On the score of empirical adequacy, even quantum physics does marvelously well (as it would have to, given its vexing conceptual arcana), and economics does rather abysmally. Of course, there is an important difference between physics and economics. Physics is a descriptive enterprise, and economics of the sort under review is normative. This all but guarantees some degree of differentiality attaching to the empirical failure of theoretical pronouncement. A theory under empirical attack is not, just so, put out of business. Various compensatory adjustments are available to the theory's loyalists. But there is one plea that the physicist cannot even consider entering, which for the economist is a metatheoretical commonplace. What the physicist cannot say in the face of recalcitrant experience is that nature is not operating as she should. This is precisely what the economist will sometimes say and might be right in saying. Bill and Sue and Fred are not acting as they should. When this is so, it will also sometimes be true that their behaviour contradicts a theorem in the theorist's ideal model. The contradiction will not matter unless those theorems also confer norms upon actual agency. We might even suppose that the theory brims with rich mathematical content, and that, for this to have been so, the behaviour of the ideal model had to be shaped to that end. The two facts, however, are not linked as the theorist may have wished. If the theory's laws are indeed normative for actual economic agency, then the norms of actual agency will enjoy mathematically rich assays. In this there may be a metaphysical connection worth remarking upon. It might be true that the norms for actual agency have lots of mathematical content. But it cannot be true that having lots of mathematical content suffices for normicity. Neither is it true that idealizations are justified because they facilitate engagement of the mathematics. So the economist and decision theorist, no less than the pragma-dialectician, is re-met with the bootstrapping problem.

Notes


3. An interesting problem emerges in so supposing, though I shall not deal with it here. It is the problem of fitting together actual performance and ideal performance by way of describable approximation relations.

4. The example is implausible, and is for illustrative purposes only. In actual theoretical practice, L would fail because of its incompatibility with pre-theoretical data. Any theorist of argument "already knows what everyone else already knows", that question-begging is not a good thing.

5. The extent to which the social sciences are normative is still a matter of lively controversy. Keynes in a letter to Harrod want[s] to emphasize strongly the point about economies being a moral science. I mentioned before that it deals with introspection and with values. I might have added that it deals with

As against this, many economists see their enterprise as "value free." As against this, is the classic riposte of Myrdal. There is little point in allowing these contentions to engage us. Some social scientists who see their work as value-free evidently incorporate into their theories normative concepts, such as rationality, and often they let their ideal models authorize prescriptions for rational conduct. Even so, they see their work as non-normative in the further sense that, e.g., one can describe a mechanism for reducing unemployment without in any sense endorsing or recommending it. (But see again Myrdal). For our purposes, it is enough that the discussion admit instances of social scientific theories that are normative in first sense but not the second. In the case of economics, I myself am pretty much at ease with a classification which I borrow from Patrick Suppes, 1961, "The Philosophical Relevance of Decision Theory," The Journal of Philosophy 58, 606.

<table>
<thead>
<tr>
<th>Individual Decisions</th>
<th>Group Decisions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classical economics</td>
<td>Game theory</td>
</tr>
<tr>
<td>Normative theory</td>
<td>Statistical decision theory</td>
</tr>
<tr>
<td></td>
<td>Moral philosophy</td>
</tr>
<tr>
<td>Descriptive theory</td>
<td>Learning theory</td>
</tr>
<tr>
<td></td>
<td>Survey studies of voting behavior</td>
</tr>
</tbody>
</table>


8. Exceptions are theorists such as Keynes, for whom economics is a branch of logic. Another exception is the Austrian School of economics, for which the fundamental postulates of economics are synthetic à priori. See Ludwig von Mises, 1949, *Epistemological Problems of Economics*, (trans. G. Riseman), New York: New

10. With Socrates, for example, who repeatedly made the following *autoepistemic* argument: If there were knowledge of such things, I more than anyone would have it. But I have no such knowledge. Therefore, there is no knowledge of them. Further, since theories are a bridge from belief to knowledge, there are no theories about these matters. Indeed, we just have our beliefs and the dialectical mechanisms that drive and re-shape them. Wisdom is knowing this.

11. It is worth emphasizing that the otioseness of the ideal models methodology is compatible with the normative soundness of theories, such as pragma-dialectics, it seeks to serve. My complaint here is not against the pragma-dialectical norms, but rather against the pragma-dialectical penchant for ideal models.


17. In econometric usage, a model is a set of assumptions which the theory has not targeted for testing. Models thus contain what we have been calling "data for theory" or "pre-theoretical beliefs". Perhaps it should be remarked in passing that although Keynes was something of an econometrician in his *Treatise on Probability* and although he has opinions about models which he passed on in his letters to Harrod of July 1938, Keynes is not there employing the word "model" in the econometrician's sense.


20. Quantum theory accurately describes the colour of gases, the heat potential of solids, and the nature of the chemical bond. It explains the periods in the periodic table of the elements, it can fathom the electrical properties of conductors, semiconductors, and insulators; it gives a theoretical foundation for lasers and masers, superconductors and superfluids; it adumbrates neutron diffraction and electron imaging; it fixes numbers of properties of the elementary particles and of radioactivity and so on.
The difficulty had by economists in hitting empirical targets were noticed by Mill (John Stuart Mill, 1836/1967), "On the Definition of Political Economy and the Method of Investigation Proper to It", in Mill (John Robson ed., Collected Works of John Stuart Mill, vol iv, Toronto: University of Toronto Press) in what is still one of the best essays on the methodology of economics yet to be written. A hundred and two years later Keynes was noticing the same thing, but not apologizing for it. Mill and Keynes had the same diagnosis. To get smooth and powerful economic laws, the economist must engage in hefty abstraction from the raw world. Beyond a certain point abstraction is distortion. One of the more interesting empirical failures of neoclassical economics was the discovery (R. Lester, 1946. "Shortcomings of marginal Analysis for Wage-Employment Problems", American Economic Review, 36, 62-82, and R. Lester, 1947. "Marginalism, Minimum wages and Labour Markets", American Economic Review, 37, 135-148) that for a profit companies do not seek to maximize profits.