

DESIRE-BASED REASONS, NATURALISM AND TOLERABLE REVISIONISM: LESSONS FROM MOORE AND PARFIT

Attila Tanyi
Universidad de Estocolmo

1. Naturalism and the Model

In the theory of normative reasons one particular approach enjoys widespread support: the Desire-based Reasons Model (from now on: the Model). The Model, moreover, typically comes with an ethical naturalist (from now on: naturalist) background. Ethical naturalists hold that normative judgments owe their normativity to the facts they refer to in the natural world. Not everyone is happy with this claim, however. The most influential argument launched against naturalism is G. E. Moore's Open Question Argument (OQA). But recently Derek Parfit has joined Moore in rejecting naturalism by employing a knock-down argument — the Triviality Objection (TO) - against naturalism *per se*. My aim in this paper is to investigate to what extent these objections can be used to reject the Model. To this end, we first need a suitable account of what naturalism is and in this context we then have to locate the Model; then we can turn to the objections mentioned.

In this paper I am concerned with what is often called substantive naturalism. (Railton 1990, p. 155; 1993b, p. 315) On this view naturalism is understood as proposing an account of normative properties in terms of *natural* properties or relations.¹ Substantive naturalism involves three approaches: analytical naturalism, non-analytical naturalism and one-term naturalism. (Dancy 2005, 126) Analytical naturalism holds that normative properties are natural properties and the two, normative and descriptive ways of capturing them are synonymous. (Lewis 1989; Jackson 1998, Chapters 5-6; Smith 1994, Chapter 2) In contrast with analytical naturalism, on non-analytical naturalism concepts and properties come apart: though to each normative

¹ Here I set aside the difficulties surrounding the notion of natural property. Instead, I will act on the supposition that such an account can be given. For a good overview of different definitions and the difficulties they face see Copp (2003) and Ridge (2003).

term there is a corresponding descriptive term and these terms refer to identical properties, the two terms are not synonymous. Although normative properties are reducible to natural properties, the identity statements employed are synthetic not analytical. (Railton 1997, 2003; Miller 1979, 1985; Brink 1989, Chapter 6)

Finally, there is the position of one-term naturalism. (Sturgeon 1985ab, 1986ab, 2005; Boyd 1988) Like non-analytical naturalism, this view only makes claims about property identity, but unlike non-analytical naturalism, it does not claim that there is a descriptive way of capturing normative properties. It is in this sense non-reductive: normative terms may not have corresponding descriptive terms, even though the properties they refer to are identical. That is, even though we may know that normative properties are natural properties, we may not be able to tell *which* properties they are. Recently, in explaining the motivation behind the position Sturgeon has argued that it is extravagant to assume that we have descriptive terms for all natural properties. (2005, 98-9) The paradigm examples of natural properties are those science deals with and even in their case, due to the terminological innovation that is a standard feature of scientific progress, new terms are introduced for properties not previously recognized, and there is no reason to think that this process is ever going to end. Hence there is no assurance that if some normative term stands for a natural property, we can (in fact: will ever be able to) find a suitable descriptive term to represent that property. Nevertheless, since there is good reason to suppose —contrary to certain critics— that normative properties play, or at least seem to play a causal role, there is good reason to think that they are natural properties. In fact, the role these properties play in the causal network may help us say something about them without giving an explicit reduction in descriptive terms.

With these distinctions in mind, let us now see how naturalists interpret the Model. On a naturalist reading, the Model designates two properties, one normative (the property of being a reason), the other natural (the property of desire satisfaction), and connects them in some way. The connection is that of identity (although there is a metaphysical alternative, which I will mention later) and the terms used to capture the properties may or may not be synonymous. That is, the sentence ‘You have a reason to φ ’ and the sentence ‘ φ -ing would satisfy one or more of your desires’ either means the same and/or refers to the same property. It is a question to what extent one-term naturalism can support the Model, soon I will say more about this. Any other reading of the Model —and there are alternatives— would not, however, do

justice to the idea that on the Model desires *ground* or *provide* reasons for action. (cf. Dancy 2000, 17-19, 26-9) Further variations of the naturalist reading of the Model, moreover, need not concern us here. The objections to be examined target naturalism, not any particular version of the Model.

2. Two objections, one response

There are two relevant objections: one based on G. E. Moore's classical argument, the other, more recent coming from Derek Parfit. Both aim to refute naturalism as such and both, I believe, fail. Yet, their failure is instructive spelling at least potential trouble for the naturalist Model. I take up each objection in turn.

G. E. Moore's Open Question Argument (OQA) has a long history. (Moore 1903, 10-21; Ayer 1936, 103-106; Brink 1989, 152-3, 162-3; Sturgeon 2005, 95-6) Famously, Moore has claimed—or, more precisely, has been interpreted to have claimed—that since for any natural property F it is possible to think that x has F while at the same time thinking that x is not right (good, rational etc.), it follows that, for no natural property F is it self-contradictory to claim that x has F but x is not right (good, rational etc.). But it would have to be self-contradictory, if it were possible to give a naturalistic definition of rightness (goodness, rationality, reason etc.). Therefore no synonymy can exist between normative and natural terms. The OQA employs several controversial thesis concerning the nature of analysis, the right form of naturalism and so on, and each of these theses has been questioned and rejected. Although not everyone accepts that the OQA is a failed argument, here we can accept that this is so.²

Even so, there may be a way to resurrect the OQA. Perhaps naturalist critics of Moore are right and a *direct* defeat of naturalism is impossible. But this doesn't rule out that an *indirect* defeat would also be impossible. Peter Railton makes the following point. (Railton 1990, 158) Every naturalistic account of normative terms leaves some questions open. An account of normative concepts is necessarily revisionist to some extent. It gives us a propo-

2 See Firth (1952), pp. 330; Lewis (1989), pp. 129-30; Smith (1994), pp. 38; Jackson (1998), pp. 151 for the analytical naturalist answer, Sturgeon (1985a), pp. 25-6; Railton (1990), pp. 156; Brink (2001); Copp (2000); Brandt (1979), pp. 14-6, 126-8 for the non-analytical naturalist answer and Sturgeon (2005) for his defence of one-term naturalism *contra* the OQA. For a recent criticism of the analytical answer see Hatzimoysis (2002), pp. 11-13 and Gampel (1997). Horgan and Timmons (1990-1), (1992ab) argues against the idea of rejecting the theory of meaning employed by Moore.

sal that aims to capture everything about the given term, but there may be some functions of the pre-reductive term, which the proposal cannot account for. This is why open questions occur. From this perspective it does not matter whether the proposal in question offers a conceptual analysis or only a synthetic identity statement. They both want to fulfill the task of giving us an account as perfect as possible and they both fail if they leave us with questions concerning important function(s) of the given term.³ Although the OQA may be a failure, in this indirect way it has a lasting heritage.

Railton calls the corresponding condition *tolerable revisionism*. (Railton 1990, 159; 1993a, 282; 1993b, 316; cf. Stevenson 1944, 36-39; 1963, 11; Brandt 1979, 14-6, 126-8; Gibbard 1990, 32) As said above, tolerable revisionism is a condition on almost all forms of naturalism. The exception is Sturgeon's one-term naturalism. Since his idea is that although normative properties are natural properties, they cannot be captured in descriptive terms, there is no way we can evaluate his reduction. For strictly speaking Sturgeon does not propose a reduction; he only argues for the possibility and plausibility of naturalist reductions, but deliberately says nothing about *particular* reductions: he does not tell us which natural property is identical with or constitutes the given normative property. In lack of such candidates, however, there is nothing we can say, it seems. And there is something disconcerting about a meta-ethical position that does not say anything concrete about its own substance. For instance, we might try to say that Sturgeon's naturalism offers a good background for the Model. But does it? We don't know. We don't know if this would be a reduction Sturgeon accepts since by his own lights he cannot say anything about the reduction basis of the property of being a reason. If he does, his naturalism immediately transforms into one of the other forms of naturalism and becomes liable to the requirement of tolerable revisionism.

For this reason I don't think that Sturgeon's version of naturalism would be of much help for the Model. The same feature that makes it a more defensible form of naturalism, namely the non-reductive element it incorporates, makes it unable to serve as a background for the Model. And as to the other

3 This claim needs to be qualified. Since revisionism is a comparative enterprise, we may conclude that we should not discard a particular theory even if it cannot account for an important function. This can happen, for instance, if, after we have examined all the available attempts, we find none that can account for the particular function while being no better, or in fact worse than the attempt under examination. This is where 'real' revisionism may come into play: we can declare that if this is so, we should perhaps do without that bit of meaning (however important it is), and reform our present use of the term. This is what some naturalists explicitly propose. Brandt (1979), Chapter 1 and Rawls (1971), 60-63§ are the stock examples.

forms of naturalism, the charge of intolerable revisionism always hangs in the air. Next I want to show that another knock-down argument against naturalism leads to the same requirement and, *ipso facto*, to the same diagnosis. The Triviality Objection (TO), as we may call it, was most recently advocated by Derek Parfit, but it too has a history. Parfit himself attributes it to Henry Sidgwick and it seems to play a role in Allan Gibbard's rejection of naturalism. (Parfit 1997, 123-4; Sidgwick 1907, 26n; Gibbard 1990, 33) And the most recent, and to my mind best formulation of the objection comes from Jonathan Dancy. (Dancy 2005, 131-2)

I suggest that the right interpretation of the TO is the following. On both forms of naturalism, i.e. on both analytical and non-analytical naturalisms we seem to lose the distinctness of the normative. Let me put it this way. We have two vocabularies, one normative and one non-normative. If either of the two naturalist accounts goes through, however, we will lose the normative vocabulary. This is clearest in the analytical case since here by construction we reduce the normative vocabulary to the non-normative. But, in a more roundabout way, the same happens in the non-analytical case as well. Here we will have two, one normative and one non-normative way of capturing the same fact; but because the fact captured is the same it seems that we will be able to say everything with the non-normative vocabulary. We won't any more have the normative vocabulary; normative terms will be *eliminated*. The result is that instead of explaining normativity, we explain it *away*. On analytical naturalism, to take Dancy's example, explaining the claim that something ought to be done with reference to the idea that doing it minimizes suffering would be like explaining why a man is married with reference to the fact that he is a bachelor. And on non-analytical naturalism, it will be like explaining why something is gold with reference to the atomic structure exhibited by it.⁴ These aren't real explanations since the explanandum simply restates the explanans; or, to put it in the language of reasoning, the agent's conclusion simply restates his premises without adding anything to them. Consequently, we don't really explain here anything; we don't really conclude about anything. But we do think that there is something there to be explained, something to be concluded about; something ethical and normative. Naturalism, however, does not live up to these expectations.

There are several ways to respond to the TO. One, of course, is to adopt Sturgeon's one-term naturalism. I, however, have already offered my misgivings about this theory as a background for the Model. Another option would

4 Both analogies are taken from Tännsjö (2006), pp. 219-220.

be to turn to the idea that though normative facts are not reducible to natural facts, they are nothing over and above such facts. Many advocate this idea—often put as the distinction between identity and constitution—on the ground that normative properties can be realized in multiple ways. (e.g. Brink 1989, Chapter 6) And it seems that if the idea is defensible, naturalists can find a way to get around the TO without adopting Sturgeon’s one-term naturalism. For this form of naturalism also has its non-reductive element, this time on the level of properties; hence the claim on which the TO rests, that we can do without the normative entirely, is no longer valid. Of course, as with every meta-ethical theory, here too there are doubts. It is questionable, for instance, whether this non-reductive kind of naturalism does not eventually collapse into one of its reductionist rivals. And there are other serious problems regarding its account of properties. (McNaughton and Rawling 2003) But I do not want to investigate this view here; doing so would take us too deep into metaphysics, something that goes beyond the scope of this paper. It is moreover an issue I have investigated elsewhere (Tanyi 2006).⁵

There is, however, a third response that is open to the naturalist, even if he advocates reduction and identity. He has to hold that normative properties, though reduced to the natural and capable to be captured in descriptive terms, are nevertheless genuine properties with an independent role to play in human practice and discourse. Peter Railton calls this project *vindictory reductionism*. (cf. Schroeder 2005, 2007, Chapter 4) He puts the idea in the following way: “[...] the naturalist who would vindicate the cognitive status of value judgments is not required to deny the possibility of reduction, for some reductions are vindications—they provide us with reason to think the reduced phenomena are genuine”. (Railton 1990, 161) And at another place he says: “Some reductions explain away the reduced phenomenon, but others simply explain it—and thereby show it to be well-founded”. (Railton 1993b, 317) To support his case he brings examples. The reduction of water to H₂O or the reduction of salt to NaCl, he says, reinforces rather than impugns our sense that there really is water or salt. In short, it is vindictory not eliminative. Similarly, the reduction of seaworthiness to a set of physically realized dispositional properties of vessels, does not eliminate talk of seaworthi-

5 Dancy claims that Frank Jackson’s view—that ethical properties, more precisely, the property of being right is a vast descriptive property—also escapes the TO. See Dancy (2005), pp. 132. Although Dancy’s reasoning at this point is not entirely clear to me, even if he is right, this is still not a problem. For the Model does not use a Jackson-type vast descriptive property in its account of the property of being a reason, hence its defenders cannot appeal to Jackson’s view at this point.

ness; instead, it vindicates our use of the term. (Railton 1990, 166) We can find counterexamples too, cases when the reduction did (or would, if proposed) eliminate the reduced term: phlogiston, caloric fluid, vital force, polywater, the non-divine reduction of the sacred are all like that. Hence, Railton points out, whether a reduction is vindicative or eliminative will depend on the specific character of what is being reduced and what the reduction basis looks like” (Ibid. 161).

This is still vague. What makes a reduction in one case eliminative and vindicative in the other? The key term is ‘vindicative information’. The idea is that vindicative reductions provide us with crucial information about the notion reduced by placing it in the world in an unproblematic way. (Railton 1993b, 318) Take the reduction of water to H_2O . In knowing that water is water, what we knew was that water is the colorless liquid that flows in rivers, falls from clouds as rain, etc. But in coming to know that water is H_2O , we were told that water is a substance whose molecules consist of two hydrogen atoms bonded to one oxygen atom. This is an important piece of information that explains why water, i.e. the colorless liquid exists and takes the form as it does. Hence, though the facts reported are the same, it is important that they can be reported in some other way: for such possibility conveys vindicative information. As opposed to this, eliminative reductions either do not produce any information or, if they do, it is such that it proves the reduced notion to be redundant or non-existent.

I have said that vindication is a matter of information produced by the proposed reduction. The examples used, the reduction of water to H_2O or the reduction of salt to $NaCl$ seem to suggest that for a reduction to be vindicative it needs to produce important *new* information. This is probably so, but I don’t now know how to conceptualize this requirement. The problem is that the information must not merely be new but must also be vindicative—and it is this bit that I cannot just now give an account of. In any case, there is one *minimal* condition that every reduction must fulfill in order to be vindicative: it must reproduce all the vital but ‘old’ information we associate with the given term. This also seems to be a good methodological way to proceed. After all, if a given reduction cannot even reproduce the already existing information, it can certainly not be vindicative. Whether it produces information beyond this point is another question, which may or may not be crucial; we need not deal with it now. This is certainly how Railton understands vindication and since I believe that the Model will fail even this minimal requirement, I see no reason to demand more at this point. But if this is so, our

response to the indirect OQA and our response to the TO join paths. For this requirement, i.e. that vindicative reduction reproduces all the vital but old information we associate with the given term, is exactly the condition of tolerable revisionism.

3. Conclusion

The possibility of informative reduction provides metaphysical grounds for vindication. It shows that it can matter whether some fact can be reported in a different way or that the term some fact refers to has a synonym. This brings with it an important reduction in the scope of the TO, invoking the condition of tolerable revisionism and tying the fate of the TO to that of the indirect OQA. As a result, while both the TO and the OQA were presented as knockdown objections against naturalism *per se*, the possibility of tolerable revisionism makes their success depend on a meticulous case by case analysis. For to show that some reduction is intolerably revisionist, we must dwell deep into its content. This task, however, must wait for another occasion; until then the indirect OQA and the TO remain serious but not yet fulfilled promises to refute the Model.⁶

⁶ This paper has greatly benefited from the detailed and helpful comments of Krister Bykvist, János Kis, Ferenc Huoranszki, Torbjörn Tännsjö, Greg Bogner, Andras Miklos, Eric Brown and many others. I also thank audiences at the ECAP 5 Conference in Lisbon, the 'Ethical Naturalism' conference in the University of Durham, the XLV Reuniones Filosóficas at the University of Navarra, and in research seminars at the University of Oslo, Stockholm University and the Central European University. Work on the paper has been made possible by a Guest Fellowship from the Swedish Institute.