

Value Naturalism Unleashed

David Brax, Department of Philosophy, Lund University
david@brax.nu

Introduction

To what are you committed in posing as a value naturalist? Quite a few things, it seems. You are committed to making *sense* of how a property as queer as value can comfortably (or awkwardly, for that matter) fit into a picture of the natural world as such. As a value naturalist you are committed to the claim that value is natural, of course, but also to some subset of optional features, the commitment to which is fixed by the kind of naturalist you take yourself to be. If you are an *analytical* naturalist, for instance, you take value to be whatever natural property it is as fixated by the *meaning* of the concept “value”. If you are a *synthetic* naturalist, you’d be characteristically hesitant to make that claim. A sort of analytical, although of an a posteriori variety, naturalist would say that if we knew what value *meant* we would also know what natural property it *was*. You might also turn it around and say that if you know what value *is*, you would presumably know what ‘value’ *means*. The distinction between these two versions would be a mere matter of method.

Now, why would one want to say such a thing? Well, this might be why: we don’t know that much about what we mean with ‘value’. We know some things about it, we can certainly use it intelligibly enough, and we do know about a few things that seem to *have* it. But we don’t know for sure what it actually is. If we did, we would presumably have the means with which to make analytically true statements about value: we would just stipulate that “value” means that which we have found out that value was. This does sound a bit naïve, and so it is: concepts do not evolve as easily as that. There is no obvious way of analysing “value” that would satisfy all competent users. And this holds not just for naturalist proposals, but for value theoretical proposals of *any* kind, it seems. If we are to make any progress in this field, we need to consider other approaches. Presumably, how we go about our business doing *this* will not satisfy all competent users either.

Any (naturalistic) conception of value is bound to sit awkwardly with some of our evaluative notions. Either it will sit awkwardly with value-*concepts* and with linguistic practice, or it will sit awkwardly *ontologically*. In most cases, it will sit rather awkwardly with both, and a stupendous amount of philosophical sophistry has gone into *cushioning* it, as it were. Naturalism in general is driven by the general idea that value must somehow fit into the natural world, and that value language should bend to fit with it. The alternative is to find a close (irrealist) conceptual fit and give up on the ontological claim (realist or naturalist). This alternative would be all very fine if there actually were a promising conceptual fit to be had. If none can be found, naturalism has at least that going for it that it makes a serious attempt to close the annoying fact-value gap. What else speaks in favour of naturalism depends on what *kind* of naturalism is defended. Naturalism *as such* does not disclose much about value, justification must be done on the basis off its instances, not of its schema.

In “Naturalism and Prescriptivity”, Peter Railton distinguished between what he called *methodological* and *substantive* naturalism. Methodological naturalism says that philosophy does not possess any distinctive method to yield substantial truths, but rather that philosophy should proceed in an a posteriori fashion, preferably “in tandem with” empirical studies,

broadly understood, in the natural and social sciences. The substantive form Railton characterizes in the following manner:

A substantive naturalist advances a philosophical account of some domain of human language or practice that provides an interpretation of its central concepts in terms amenable to empirical inquiry (N and P p 156)

The two forms of naturalism do not necessarily go together. This is not exactly the traditional way of making divisions in value theory, but it seems like the naturalisms proposed by Railton are on the formal side of the formal/substantial distinction, and on the synthetic side of the analytic/synthetic distinction. What Railton calls “substantive naturalism” is closer to formal naturalism as tradition (i.e. Moore, and the entourage) has it. But take note that he does not say that substantive naturalism gives any *analysis* of central concepts, but an *interpretation* in terms “amenable to empirical inquiry”. What comes out of such a venture is probably not intended to close any open questions. I would here like to embrace naturalism about value in both the Railtonian senses. In my doing so, I will go against the grain of most contemporary forms of naturalism, it seems.

Naturalism in most of its contemporary forms is recognisable by an array of features that seems to work as *restraints* on what can be properly construed as naturalism at all. In this paper these restraints are questioned as to make way for other forms of naturalism. It thus side-steps some of the traditional problems for naturalism and turns the attention to others, to my view more promising ones. Naturalism, as the title promises, is “unleashed” from some of these restraints. The restraints are

- The emphasis on value *judgment*
- Locating value in the intentional *object*
- The constraint on *mind-dependence*
- The emphasis on the *analysis* of ‘value’
- Supervenience and reasons

All of these represent unnecessary restrictions on naturalism about value, and most of them are respected by the naturalistic theories around. I’m not here concerned with showing that there is anything *wrong* with the theories that *do* respect these restraints (although I will certainly hint that there is), I’m merely pointing out that theories violating them could still be legitimate cases of naturalism, and fairly promising ones at that. Respecting these restraints is optional for the naturalist. I will also try to show that these restraints represents important theoretical *choices* for the naturalist, the decisions of which can be theoretically or otherwise motivated, and the making of which will determine what problems you will henceforth have to deal with. It will also influence what type of problem you take value-theory to be concerned with. Resent cases of naturalism being rather uniform in these particular choices, the debate in naturalism has focused on a subset of problems, unnecessarily restricting its range. New approaches (like affective theories) have been tested out on this limited set of problems instead of, as I would prefer, changing the fundamental issues.

The first restraint, the focus on value *judgment* in value theory is contrasted with a renewed interest on value *experience*. The *data* of value-theory are investigated, and the question about value *knowledge* is looked in to.

The second restraint, focusing on the intentional objects of value judgments and value experiences is replaced with a closer look on the states whose objects they are.

The third constraint turns on the unfortunate argument that value, if real and ontologically respectable, must be mind-independent. The controversial point about the ontological status of experiences should not be taken as a recommendation of keeping clear of it in matters of value.

The fourth and most general unleashing turns from a focus on analysing ‘value’ and value judgments to matters of more explanatory work. Naturalism, it is proposed, is backed up by a more “nuts and bolts” approach (methodological) to whatever it is naturalism about.

Fifth, the consensus seems to be that value *supervenes* on natural properties. In particular, value is believed to supervene on the natural properties offered as reasons when a value judgment is being justified. When the first and second restraints are dropped, this claim is undermined. I will not deny supervenience, but I will question its relevance and how it has traditionally been put to use.

Roughly, what I wish to show is that if these restraints are dropped, possibilities open up to deal with matters like the naturalistic fallacy (it is only construable as a fallacy if some of the restraints are accepted) and the open question argument. Besides, there is explanatory force to be latched on to. It broadens the outlook for naturalism. Everything that speaks against it, it seems, are things that speaks against it only if one accept one or more of the restraints above.

Now, one might want to hold on to some of the restraints listed. For my part, I only see it as a requisite that naturalism explains the general features of such a strange thing as value, all else are just matters of conventions, and misleading ones at that. But a case might be made that the restraints, especially the one about making value-judgments uncomplicatedly true or false, is an absolute must for naturalism. In that case, I propose that the project undertaken here is about *naturalizing* value. It is broadly naturalistic in outlook, and, if I’m not mistaken, if its not naturalism, it is because it is even more naturalistic than that.

Value Data: from experience to judgment

There is nothing surprising or especially new in the observation that analytical philosophy turns to propositions when in doubt. Dealing with propositions is, to a large if not exclusive extent, what analytical philosophy *does*. So when it comes to matters of value, an obvious way to start is to scrutinize the behaviour of value *judgments*. After all, they are comparatively well-behaved, proposition-like entities. But value judgments are not the data of value-theory. This section will deal with the matter of value-data, and how they relate to value judgments.

If naturalism is true, value theory is *about* something natural, and if there is something natural to be about, there better be *data* of some sort or other. Data is needed to justify natural judgments. These data just *cannot* be the existence of value-judgments and their internal relations merely. Reason cannot provide the *content* of a theory about the good. Graham Oddie takes note of this in his recent book “Value, Reality and Desire” (2005) and argues convincingly that the value data we got are *value-experiences*. Further, he claims that those value experiences are *desires*, or at least that value-experience is a determinate of which desire is a determinate.

Whereas value judgments might be more reliable than value experiences, for instance in being more consistent over time, *value experiences as such are most likely prior to value-judgments*

as such. If value-experiences are the data we have access to, the relationship between them and value-judgments seems important¹.

The behaviour of value judgments is not our main interest. It would be if it were the only indication we had as to what value might be, but it isn't. This is far too frequently forgotten in contemporary value-theory. While not necessarily primarily concerned with value judgments, the naturalist must be able to say that the natural property picked out is what makes value-judgments true. But value-judgments are not taken as the primary explanandum. The explanandum, will be *value*, and our access to value is not value-judgments, it is the data and what the data implies, tracks or supports. When *this* is done, the truth and falsity of value-judgments is a free lunch.

Now, most naturalists realise that we need to look at desires (as pro-attitude and/or as experience of value) if we are to find out anything about value. It is in explicating the role of these experiences and attitudes naturalists come to differ.

A caveat: As naturalists, we want to know what makes value-judgments true, but we are not necessarily concerned with making *our* typical value judgments true. It is not the business of naturalism *per se* to construct an as coherent theory as possible under the assumption that our favourite value judgments are true. Even on a deflationary theory, this does not mean that undermining our confidence and explaining evaluations away implies that evaluations are *never* true. It just means that it is perhaps not exactly as we tend to think. If, indeed, we tend to think anything very specific about these matters, something we might very well doubt.

From value data to value knowledge

If there is such a state as the experience of the goodness of P, then, by analogy with the perceptual case, it would give me a reason to believe that P is good. (Oddie (2005) p 41)

We know about value, Oddie says, primarily by things seeming to be good to us. "Seeming to be good" translates roughly into "being desired". To desire something is to have that thing appearing to you as something good (this is a rather common view these days, but the experience usually spans over a rather wider array than mere desires²). But there is no *necessary* connection between believing that something is good and desiring it. There is nothing *incoherent* in saying that "I believe that x is good, but I do not desire x". But still, it is a bit awkward. If desire is the experience of value, and is a defeasible reason for belief in the goodness of x, this awkwardness sits very well indeed with the theory.

The experience of value is understood as our *acquaintance* with value, and while this is not our only source for getting to know value (we might be told about it, we might infer it, we might invoke standards and instruments previously developed etc.), experiencing it is necessary for understanding its most central features. Experiencing value is on this account not a type of knowledge of value³, it is rather one of our (defeasible) *reasons* for making value judgments. Experience is not knowledge, but experience is necessary in order for us to have (at least one type) of knowledge, namely "acquaintance-knowledge"; knowledge of that experience⁴.

¹ Although there is probably no necessary relation, the latter seem to require conceptual capabilities that the former doesn't.

² Helm, Johnston, most emotion theorists

³ See Kevin Mulligan (1998) who argues that it never could be

⁴ See Frank Jackson's "Epiphenomenal Qualia" where he invokes "Mary the colour-scientist", a person who knows all there is to know about colour in physical terms but is colour blind. We might construct a "Mary, the value theoretician" case, a person who can communicate, predict and in general deal competently with value-

By being our acquaintance with value, Oddie says, desires play the same role in value-judgment as perception does in perceptual judgment.

Value experiences *track* value, they *pick* value properties *out*, but they are not infallible. There are occasions where value-judgment goes against value-experience. How should this be dealt with? Easily, on Oddie's view: the same way I can see red without believing it to be red, if I've got contrary evidence of a reliable nature. "It looks red, but I don't believe it is" is not incoherent, nor is the delightfully wicked proposition "It feels good, but isn't". The lack of incoherence in these statements does nothing, of course, to rebut the fact that feeling good or looking red is precisely the type of evidence we base our judgements on.

The great thing about having desires in this (data) position is that it gives you a strong connection between value and motivation, while staying clear of strong internalism. If desires are simultaneously the experience of value and the type of motivating, dispositional state it is usually thought to be, there is no wonder that knowledge of value tend to motivate. Further, since the relationship experience – judgment is not yet clear, you can argue for the connection being exactly as strong as you want or need it to be. As Railton pointed out (in "Facts and Value"), what naturalism needs in order to answer its critics is a way to show that its conception of value is not "intolerably alienated" from our evaluative practices. By treating motivational states as value data, this is elegantly avoided.

The Turn

Experience of value plays a rather ungrateful role on most theories of value. In contrast to the realist/naturalist approach of treating value-experiences as *evidence* and thus giving reason for value judgments, non-cognitivism understands evaluative judgements as having an *expressive* or *prescriptive* nature. Value-experience, then, can be cast in diametrically opposite roles: either as *evidence* for there being something the value-judgment can be true about, or as that which is *expressed*, which guarantees the *sincerity* of the judgment.

Oddie observes that it is necessary, if value-theory shall have any serious content at all, to account for value-experience. He argues for a realistic theory which he takes these experiences to be in accordance with⁵. But there is yet another reply to this observation: To identify value with value experience. One theory that does this is Hedonism, but other mental state theories might as well. Experience of value, I guess, would be all intrinsically valenced experiences, i.e. the emotions (according to appraisal theories, "being valenced" would be the determinable, desire, then, would only be an experience of value if it "felt good". This is not precisely what Oddie intended). Let's call these theories cases of "mental state naturalism".

Value experiences can be tried at various positions and in various capacities in order to make sense of value. But comparatively seldom, these days, it is tried at the most central position: as value itself, as "the good". The suggestion here is precisely that: to treat value experience as the good, and not merely as "a good", as Moore would have it. This, I claim, is not to commit any "naturalistic fallacy"; it is making a characteristically bold theoretical decision towards naturalism.

Experiences of value are responsive to natural features of the thing experienced. But they are not mere representations of those features. They are not necessarily to be understood as experiences *of* the value of the object, they are experiences of *value*, caused by the object, only given the psychological state of being responsive to that object in the right (i.e. sufficient

concepts but does not experience value. In both these cases, we ask our selves: is there anything about the relevant domain that Mary does not *know*?

⁵ He prefers realism to naturalism due to the radically disjunctive nature of states desired, and there being no indication that *this* would change under idealisation.

to secure the effect) way. To what the experienced value actually belongs is not disclosed by the experience itself. Nothing in the felt nature of these experiences ensures that the value experienced belongs to the object (as we shall return to in the next section). Actually, nothing in their felt nature *implies* how they should be interpreted at all. The suggestion for interpretation here is that there being experiences of value suggests that value is *identical* to the characteristic property of these experiences. True value-judgements then are those that attribute value to these experiences, and to them only.

Johnston (in “the authority of affect”) considers this way of placing value in experience and dismisses it in a symptomatic way, by calling it names like “pornographic”. He accuses it with being too self-indulgent. But this objection is misdirected: having experience as value does not recommend any particular way of life, it merely says that the value in that life will be dependent on the experienced had, no matter how these ultimately come about. Johnston (and Oddie) pursues the idea that value-experiences *disclose* value in the world, and Johnston’s claim that affect has “authority” amounts to denying any attempt of deflating these mental states ability to detect value in the world. The answer here is to restrict the authority of affect to self-rule, as it were: affect discloses the value of affect.

Still, there hardly is any doubt that the experience of value actually works as Oddie suggests in the forming of value-judgments. A reliable causal connection between objects and features of objects and experiences of value will give rise to a type of “knowledge” of the evaluative features of external objects. What the mental state naturalist then does is to deny that this is what is important. Indeed, he would suggest that this is misleading if we want to find out the “deep truth” of the matter, and not just write the history of particular value standards. This history, ultimately, can establish truths about instrumental values only.

Does the explanation of experiences of value need to postulate the existence of value? A common challenge to naturalism/realism is that if it doesn’t, naturalism/realism is not supported by the evidence. But if value is inherent to the experience, this changes. Obviously the experience of value must be mentioned in an explanation of the experience of value.

The goodness of value experience (a bit about substantial mental state naturalism)

Now, no one denies that value experiences can be good, can be very good indeed. Mental state naturalism has an ally in this fact. Moore’s challenge, to grant that value-experience (pleasure) is *a* good, but not *the* good can here be answered: Is there anything that supports the view that the goodness of value-experience is in any way similar to the alleged value of external things? To feel good is not merely to assign value to an experience the way you assign value to a painting or a friendship or to the outcome of an election⁶. Mental-state naturalism identifies value by taking “feeling good” as its primary explanandum and to explain other values using this as a device. Substantial value theory thus heavily informs formal value theory. Indeed, on this form of mental state naturalism, they are not taken to be very separate projects at all. This is far from the usual naturalistic technology of identifying value by way of some construction from desired desires under some idealised description. Doing so avoids identifying value with an impractically disjunctive property. It also avoids the many pitfalls of dispositionalism.

⁶ C.I. Lewis (Knowledge and Valuation)

Identifying value: state and object

The next step consists in giving up a rigid attachment to the *objects* of pro-attitudes as the bearers of value. This attachment has informed most present day forms of naturalism, as we shall see. Giving it up is a natural step to take after loosening the attachment to value judgements.

Most of our judgments of value are about things other than experiences; things towards which we have certain attitudes and towards which we recommend others to take up similar attitudes. It would indeed be strange, and most likely rather fruitless, to recommend others to instantiate certain mental states (“Oh, I felt the most delightful pleasure yesterday, you really ought to try it” just doesn’t cut it. Recommending the wine/chocolate/piece of music that caused it, on the other hand, just might): if we want to recommend anything, our best bet is on the objects that made us feel like we did. But this should not mislead us into placing value in those objects. The “intentionality” of experiences of value might *encourage* us to do so, but for a number of reasons, the considerations of which will be the subject of this section, we might want to resist it. Doing so is one of the guiding lights for the type of naturalism I want to propose.

Merit and affordance

We tend to think that some objects that we care about actually *merits* positive emotional assessment. There are certain things that make the object *suitable* for that kind of responses; that make it a *fitting object for a pro-attitude*. As we shall see below (reasons and supervenience) this way of putting tends to focus on the features in virtue of which the attitude is taken up. A natural way for naturalism to go, then, is to isolate these features and ascribe value to the bearer of those features.

If we we’re good Millians, on the other hand, we could account for this by saying that the object is an *instrumental* good, it tends to cause value-experiences. The reason this won’t do is if we want to say that it *should* be enjoyed, that it still merits assessment even when we happen to fail to do so. The good, Moore famously remarked, is not what *is* desired, but what *ought* to be desired.

This can be amended if instead of “merit” we use the term “affordance”. The object *affords* being appreciated; attention paid to it, and its “good-making characteristics” is likely to be a rewarding (i.e. value-) experience.

But what an object affords is not an inherent property of that object: it is a property that inheres in that object only *given* a specified subject. Given a sufficiently specified subject, it would be sufficient to describe the object to secure the effect, i.e. the responding attitude or value-experience. It would then be quite possible to put the value in the object. But is this the most plausible understanding? Would it not be as plausible to say that it was *instrumental* to value under those circumstances? Specifying a subject rigidifies the relevant features of the object, but the *reason* why we need such rigidifying is to ensure the response. How can this be, if it is not the response that is important? Why is it important to *appreciate* the good that we have? (It might be an *additional* value, you might say).

When we *recommend* objects, we do not just trust that the objects will do its instrumental business, we often recommend aspects of it, things that should be attended to. We do this because these are the things that make the thing good, and it makes the thing good because those are the things attention to which will tend to be rewarding. We have a hen and egg situation going here, and naturalism can put its money on either of these. *My* business is to

point this out: there is a choice to be made here, both alternatives in which are consistent with being a value naturalist.⁷

The systematic displacement of value, its implications for formal value-theory

There is a host of reasons why we should come to think that *things* are good. And in a laudably democratic and autonomy respecting spirit, most value-theories respect this. This has import, not just for the substantial question of what has value, but also for the formal question of what value *is*. Here, then, there are clear connections between substantial and formal value theory: obviously, if *things* have it, value cannot *be* a certain type of experience, and vice versa: if value is a certain type of experience, (other) things cannot have it.

Nothing in the experience of value actually *puts* the value in the object of that experience. Let's call the non mental state type of naturalism the *object interpretation* of naturalism. The two main problems with the object interpretation can be simply put like this:

1. The objects picked out seem to have little or nothing in common.
2. Our pro-attitudes pick out things that we hesitate to call valuable.

These are problems of a very general kind: how can naturalism account for both the universality and the motivational element of value? The latter part is dealt with by including *some* relation to attitudes. The natural property that value is identified with is almost always done so with reference to pro-attitudes. It's not hard to see why this is so.

The role that attitudes and experiences play on the object interpretation of naturalism varies. The attitudes can *detect* or *project* the value in or on to the world. It can be part of the constitutive ground, or it can be part of the supervenience basis of value.

The first position is rather *realist* than *naturalist* in nature, which has its reasons (typically, internalistic considerations make it hard to identify value with any natural property that is not, at least dispositionally, essentially connected to some mental state of the agent).

How then should we understand this position? How deal with all the notions of necessity, irrational desires, the diversity of pro-attitudes etc.? Modern naturalists have concentrated on this set of problems (Railton, Jackson type functionalism, Copp, Sayre-McCord, Sturgeon etc.?), the mere outline of which I'd like to hint at here.

The simplest form of pro-attitude centred naturalism would identify value with being the object of a pro-attitude. This would indeed solve the first problem, all these objects would have something (although not intrinsic) in common. But it would not solve the second. Then people have gone on to say that "value" refers not to the properties picked out by *actual* experiences, but by the experiences as they *should* be; by the pro-attitudes of some idealised subject. Or, on a theory ascribing value to well-being, value belongs to the things an idealised subject would want a particular person to have, if concerned with his/her well-being.

Naturalists usually want to be able to say that the things we find valuable *are* valuable, and that their valuableness consists in their having the natural properties they have. Some or other elaboration from this seems present in *all* forms of modern naturalism. It usually affords disqualifying annoying evaluations by inserting some version or other of *idealisation* in the subject whose valuations are to be trusted.

⁷ I also happen to recommend the value experience hen (or the egg...) reading of value, even though I probably would recommend you to attend to the egg (or the hen) in the everyday situation. Value theory is not about what we should attend to in our everyday life, it's about *value*, and value is not any typical everyday concern of ours. It's even closer to home.

Now, why would this give *any* reliable indication to what is valuable, if not some evaluative feature is already baked into what we take “idealisation” to involve? Idealising a subject is one way of fixing it so as to ensure a certain type of reactions to a specific type of object. As mentioned above, the reason why this might seem like a good idea is because it is actually the reactions that matters. *If* value is value experience, and let’s not rule it out just yet, we are not guaranteed to get clearer on this point as we go along. Idealised subjects might have great taste, but they are not necessarily good value-theoreticians.

Now, the Millian form of naturalism famously has a way of disqualifying evaluations (in reply to 2) by appealing to what we could call understandable cases of misattribution. We are likely, so this argument goes, to confuse instrumental value for intrinsic value, in particular when the instrumental value is reliably productive of intrinsic ones. This type seems to have gotten lost in modern value-theory. I would here like to make a case for bringing it back (a similar approach can be found in Railton’s “Naturalism and Prescriptivity”, although he does not endorse it himself), and reinforcing it with Brandtian/Railtonian psychological conditioning (further backed up by the affective sciences: hence the “nuts and bolts” variety⁸). I suggest that this would be the mental state naturalists best bet in dealing with the second objection above.

Why has this argument, or “explanatory device”, been neglected? One of the reasons is probably that it is made obsolete by the methods employed in modern naturalistic theories: *mistakes* are presumably ruled out by the idealisation process that they put the character appealed to through.

The object interpreter has to answer the following problem: How could being valuable be *universally* motivating if it just is having a set of natural properties? What property could be identified with value in a way that ensures this desirable feature? If “Goodness” is to be one and the same thing in every possible world, (even if it does not *pick out* the same natural properties in every possible world) we need to specify it as a “Role-property” that can assign value to objects and still be essentially bound up with pro-attitudes. This, one might conclude, is doing naturalism “the hard way”.

Putting the value in the value-experience (that has some pro-element in itself) undercuts most of these problems. The object is more or less irrelevant to the value of the experience, and there is no problem with identifying a universal feature in virtue of which all experiences of value are good. The motivational feature is already secured. There is though, a possible version of problem 2, were people wants to hesitate to call certain value-experiences valuable. But note that the reason for this is almost always to be found in the object or in some constituent (like malice) of the mental state concomitant with the experience. For the remainder, if any, the Millian instrumental/intrinsic argument is readily employable.

Respecting the autonomy of people amounts to respecting what they want. And the fact is that people want things, they want states of affairs to obtain. They want other things than mental states. We are not, so goes the intuitive appeal, systematically deceived: we know what is good for us, and we know what is *good*. Admittedly, we sometimes loose track and get it wrong, but how could we possibly be wrong nearly *all* the time, as suggested if value does not in fact typically belong to the object of pro-attitudes. This is something to explain, surely, but this observation does not undermine the naturalist’s project. It *is* the naturalist’s project, when

⁸ I will actually *not* be arguing extensively for this here.

taken in the sense opened for here. What we're losing is the possibility of accounting for some of our favourite values *directly*, by simply assigning value to them. This is a challenge for mental-state naturalism, surely, but every theoretical choice brings its own host of problems. As I noted in the introduction, no plausible theory of value will be unproblematical.

Mental state naturalism of the sort suggested still can respect what individuals want. *What* is good for you, and what you should ideally get, will depend on your characteristics. But *this* fact, the *goodness* resulting when you get what is good for you, will not be thus dependent. On mental state naturalism, this is not because what you get is what an idealised version of you would want you to have; it is because it causes the experience of value. In that fashion, respect for the preferences of persons goes together with universality of value.

It is important to note that the mental state naturalism proposed here should not be taken as the outcome of any of the meta-ethical theories proposed by the modern naturalists. It is not the fact that idealised agents would desire to desire experiences of value only that makes those experiences good. It is not what is meant by saying that it is good either. The explanation goes the other way around, if it goes around at all. The theory proposed does not differ merely in the substantial *outcome* of a universally applied naturalistic method; it differs *methodologically* from those theories, by questioning the “restraints” guiding the object interpretation.

Mind-Dependence

Simplifying matters a bit, we can say that there are two options for the ontological status of value. Either it is a *sui generis* property, or it is somehow *dependent*. The matter of dependency lies at the heart of the realist/irrealist debate. Oddie outlined what he called a “Robust Realism”, a theory of value that accords with a list of requirements. Fairly early on this list we find the requirement of “Mind-independence”. To be robustly real, the value of a thing should not be dependent on the mental states of any agent. Some writers have thought that the fact of mind dependence of value undermines the whole realist project. If value is mind-dependent, it cannot be a *sui generis* property, and since value *is* mind dependent, it is not a *sui generis* property.

Sharon Street (forthcoming) takes it as the *defining* claim of value realism that there are at least some evaluative facts or truths that hold independently of all our evaluative attitudes. This is suggested as a formal restraint, obviously present to rule out theories like projectivism, and quasi-realism. But it has some unfortunate substantial consequences, if the dependency relation appealed to is not qualified.

First of all, realists obviously should not be barred from saying that value bearers can be or involve mental states, like evaluative attitudes⁹. This is not the nature of Oddie's requirement, rather *the property of value itself* can not be mind-dependent. But why couldn't value itself depend on minds? What is so bad about evaluative attitudes that make dependence on *them* particularly bad? It seems very odd. If there is a plausible dependency restraint, it must be a *general* restraint. Further, *some* sort of dependence should be allowable for realists, or else the realist would be bound to pose his/her property as a *fundamental* one, and very few wants to

⁹ Street also acknowledge that the requirement does not rule out the possibility that the value of a character, say, depends in part on his/her evaluative attitudes.

do that¹⁰. Presumably, there are *some* sorts of dependence that would make a theory claiming it unrealistic, but this does not mean that *all* types of dependence are tainted in this way.

Take a typical modern naturalist proposal like Peter Railton's. Roughly, what is good is what some one under ideal circumstances would want to want. If this is all that is meant by "good", the theory presented is not robustly realistic.

Most naturalist theories around, as we have said, want to keep the double connection to attitudes on the one hand and to the object of those attitudes on the other. A typical way of putting it is that the value *belongs* to the object by *supervening* on a particular set of its natural properties, *what* properties it supervenes on depends on (some refined version of) the attitudes of persons.

In contrast, Oddie treats desires as fallible *data*, were the data could fail in *tracking* value. This is what failing *amounts* to, according to this realist¹¹. Idealisation makes us more likely to get things right, so refinement of desires have a role to play on Oddie's account as well, but that is not the constitutive role played by idealisation on naturalistic theories.

To be realist in Oddie's sense is to accept that value is a property not ontologically dependent on the existence of minds of any particular kind¹². The reason seems to be that mind-dependence would lead to idealism, or, at best, "phenomenalism", about value. But, actually, if it does, it does it only if the object interpretation is accepted.

Typically, response-dependency accounts want to explain value as involving a subjective response without thereby explaining it away. The subjective part, as we have seen, is needed for internalistic reasons.

Naturalists typically fix the reference of "value" by appeal to response, and then use the properties referred to as supervenience basis of value (more on this below).

The restraint on mind-dependence is intended to bar the type of solutions were the object and properties valuable are those that are the objects of pro-attitudes, merely. It can not suffice to be experienced as valuable. Being such is not to have any robustly real property. Neither does being the object of ideal pro-attitudes, if Oddie's argument about properties is correct.

Realism in most its forms wants to claim that value is a property that exists independently of any observer. The role of observation is merely to find this property, detect it, as it were, and point us squarely at it. Value, on the most hardcore versions of realism, is *mind-independent*. Naturalism has largely followed this lead, putting more or less emphasis on the role of observation in axiology, the properties thusly picked out exist independently of the process of picking them out. But this won't do as a *criterion* for naturalism.

Mind-independence, I fear, is avoided due to the many faults of *projectivism*, but mind-dependence need not involve any projectivistic claims. If value were projected, it would not be *revealed*. The experiences would not *disclose* or *track* value, as Johnston and Oddie would have it. Johnston pointed out that projectivists must ultimately explain value *away*, debunking the "authority of affect" that a detectivist can take at face-value. On the account offered here, what is debunked or explained away is merely the *apparent* value "projected", but the value

¹⁰ About the sui-generis / supervenience problem, see McNaughton and Rawlings recent "Naturalism and normativity. See also the section on reasons and supervenience below.

¹¹ Note that as mental state naturalists, we can still say that experiences of value, desire etc. are deficient in one way or another. By being different from what we would want them to be, for instance.

¹² The radical disjunctive class of natural properties picked out by desires Oddie takes as a reason to embrace *realism*, as a non-reductionist view on value properties. He claims that he would be a naturalist if this set was *not* radically disjunctive, but converging towards some natural property

itself, as the basis of this projective practice, is a robust natural property. I'd say value is *not* projected, if anything, value is in fact doing the projecting. The fact that these experiences do not always bring attention to themselves does not mean that they are not what 'value' in fact *track*. The mind-independence restraint suffers from a lack of theoretical imagination, and a failure to realise the possible role played by responses in value-theory, namely the lead.

The mind-independence restraint is understandable when taken in conjunction with the former ones dealt with, but it fails if naturalism is conceived broadly, as it should. Notice that this holds when mental states are not just *valuable*, but *constitutive* or *identical* of the value-property itself. Nothing in the nature of naturalism rules this out.

. If value inheres not in the object, not in the object given the attitude, but in the mental states themselves, there is no reason why value would not qualify as "robustly real". If, in other words, the dependency relation is *identity*, neither projection nor non-cognitivist "quasifying" of value properties follows.

Note that this is not idealism (if, of course, you are not idealist about experience), nor is it subjectivism. It is not even a "sensible" one in Wiggins old phrase - it's not *that* kind of dependence, Railton on the "in virtue of our subjectivity" in terms of which things can (objectively) *matter*.

The point of this section has been to show that even if some sorts of mind-dependence *does* undermine the realist/naturalist status of value, this restriction must be qualified as too not rule out mental-state theories of value. *Realism(not even the robust sort) should have no quarrel with mind-dependence if the dependence consists in the relevant mind-states being constitutive of value.*

Analysis and explanation, two ways of doing naturalistic value-theory

The received opinion is that naturalism suffered a rather severe beating from G.E. Moore (in "Principia Ethica" and other publications). The open question argument and the charge of committing the "naturalistic fallacy" have been taken to drive the two last nails in the coffin of this vainglorious approach.

Ever since, though, different ways of coping with these problems has cropped up and very few people (naturalists) today find the open question argument to be very damaging to the naturalist cause. And while there might be naturalistic fallacies to be committed, naturalists are not bound to commit them. Naturalism could be synthetically true, or it might even be analytically true, while we might just fail to see it (Railton's version is synthetic, Prior noted that analytically true statements need not be a priori). In short, naturalist in general takes a rather relaxed attitude towards these arguments.

The best argument by far is the naturalistic fallacy: to be good is not a matter of having mere natural properties. Natural things can be good, and be so in virtue of their natural properties, but being good, so goes this argument, is not a matter of having those properties. You might *think* it is, but really, it isn't.

Some of the reasons why this fallacy have been thought to be detected is found in the above: we want to have things being good by their having the natural properties that they do, at the same time we want something in addition, we want motivation to enter somehow. This additional feature, the non-naturalist say, requires some non-reducible value property.

Pointing out a possible fallacy, though, (and for all I know, some weak-hearted naturalists might be convinced by it), is not to establish that a fallacy have been committed. A fully acceptable answer to the naturalistic fallacy is to merely state that it is not a fallacy. To

identify value, not just the valuable, is *precisely* what the naturalist intends to do. It might still be *wrong*, but it is not founded on a mere *mistake*¹³.

The guiding light of this section is that we should not pre-judge what kind of a problem value-theory is actually trying to solve. If we let the *data* be our guidance, naturalism has a few alternatives as were to go, theory-wise. The trend in modern naturalism (with the very miniscule exception of Peter Railton not speaking seriously) is to put the emphasis on a naturalistic way of doing value *semantics* (Jackson, Sayre-McCord, Copp, Boyd etc.). Here, I'll get into a non-semantically inclined way of doing naturalistic value-theory. It's the "nuts and bolts" variety mentioned in the introduction.

Substantial and formal axiology

Here is a way of reasoning, typical of the naïve naturalist: If everything that has value has some natural property in common, why not say that this property *is* the value property, and call it a day? What could possibly be the point of denying it? If there is more than one property turning up, we might want to distinguish between the role-property of being valuable and the realiser-properties that instantiate that role. Proper naturalists claim that the *role* is a natural one, whereas it is near universally acknowledged that the realisers are natural properties. Naturalism proper is not just the claim that all things valuable are things natural. Naturalism proper is the further claim that *to be* valuable is itself a natural property.

The *analytical/synthetic* distinction in analysis is not the main battle-ground between naturalists as I conceive it. Although it's quite certain that mental-state naturalists should not be arguing for analytical naturalism, it's not really arguing for *synthetic* naturalism either. It is not foremost concerned with *meaning*. It seems to me that the two questions about what value *is* and what 'value' *means* has been far too frequently conflated. This said, what is offered by the mental state naturalists is a kind of *diagnosis* of the concept by identifying the property doing all the actual work in the domain. It's a diagnosis of value naturalism as usually conceived, and offers an alternative, which take substantial value-questions as prior, or at least on a par with the formal ones, ontologically speaking. This approach is quite sympathetic to what Railton called "substantial naturalism".

But it has closer affinities with the methodological view as well. How can we argue for such a stance? First of all we could say that conceptual analysis has had its chance, now it's time to test something else. But there is an article of faith element here, that should be acknowledged. As I said, we shouldn't prejudge what type of a problem we have with value. As long as we are not sure whether it is a semantic one, conceptual analysis should be just one of the approaches working on the problem. Turning from analysis to explanation, is one of the theoretical choices open for the naturalist that forms the type of problem dealt with.

Explanation

In the Railtonian distinction mentioned in the introduction, substantial naturalism is about the concept(s) of value, but it is not about analysis, but about the *interpretation* of central terms with help of what ever is up to the task. The *methodological* naturalistic claim is that we can do better than just sit around thinking up conceptual analyses. We can actually do some *explanation*. This must be done conceptually to some extent, but some of it can be done by attending to, perhaps even contributing to, the relevant sciences. Psychology, mainly.

The important thing about this "methodological" naturalism (it is substantive too in that the method tries to make sense of value semantics as well) is that its prime motive is to explain

¹³ As naturalists, we want to be wrong, if we are wrong, in a much more dramatic way

what value is thought to explain. It tries to explain how we come to believe that things are intrinsically motivating, why we tend to seek them out and revise our attitudes with regard to them, and so on. Such a project might come up in favour of non-cognitivism (as Sharon Street thinks), but if some feature keep coming up in these explanations, naturalists might be excused if they want to say that that feature is what value is. Mental-state naturalism does just that. Actually investigating what lies at the bottom of motivation and justification is the method. Naturalising value-theory is the ideology, the favoured *approach*. Much of this paper has been concerned with paving the way for it.

This shift is probably the most radical of the questioned restraints, and the most likely to put this version out of the naturalism business. The shift to an explanatory approach, though, is supported by the questioning of the other “restraints”. Turning to explanation is a natural way of coping with the questions raised by the earlier theoretical choices. Before concluding this business and evaluating the result, let’s throw a quick glance at two mainstays in the value theoretical debate: *Reasons* and *Supervenience*.

Reasons and Supervenience

You often hear that value supervenes on natural properties of objects. You also hear that the natural properties of objects give us reason to promote them. Value, then, might naturally be thought to supervene on those properties that give you reason to value it. If this were true, our evaluations would be true when justified, i.e. when there are sufficient, non-defeated reasons for holding them. This would occur when value actually supervenes on the thing valued, since the presence of reasons ensures the presence of the supervenience base. We would then have rather well-behaved instances of value-knowledge.

This should look familiar from the modern, non-mental state types of naturalism known as the enemy in this paper. Reason and supervenience are tightly knitted on most naturalistic theories of value. The last unleashing to be done in this paper concerns this alliance.

On the account we have been concerned with making room for, the mental-state naturalism, value supervenes on natural properties in the least interesting sense of all: by being *identical* to the property of value experiences. Reasons on this account can be explained as by-products, but not as constituents, of value. They are *practically* connected to value, pointing towards that which is *instrumental* for bringing it about. This, of course, is not to deny that reasons has to do with value, or that value somehow supervene, but supervenience is not held to be a very relevant issue. Emphasising the supervenience of value is trivial at best, misleading at worst. The account presented also denies that the key to value lies in the relation to reason. The key to reasons, on the other hand, might very well be found in the relation to value. There is a reason to bring something about *because* it will lead to valuable experiences.

When arguing for the supervenience of value, the usual appeal is to the features that *make* things good. It’s the colouring of that painting, the delicacy of that brush-stroke, that makes it a good one. When justifying our evaluative judgments and feelings, we often *point out* just what it is that we base our judgment on. Since neither you nor I have any ready access to the psychological state of the other, we settle for pointing to things, try to make each other come around to seeing things our way. As we think they *should* be seen, *merit* to be seen, no doubt. (Our reasons for valuing are, one could say, mostly of the “wrong kind”, value wise).

None of the properties cited as reasons for evaluation are good candidates for *being the* value property, and the *disjunction* of them is open-ended and seems to lack any distinctive pattern. So, many have concluded, they are not value, they are merely the supervenience basis for value, and the unity of value is not to be found in the supervenience founding properties, but in their relation to something other, such as the desires of some idealised desirer or what have

you. This, I would say, mystifies more than it clarifies. The alternative answer is to retreat into non-cognitivism, and offer some deflationary account for how come we believe what we do (as quasi-realists does).

But naturalists do not have to flatter evaluators by saying that they are *nearly* getting things right. It does not even have to flatter other naturalists. The severely disjointed set of good-making characteristics appealed to in everyday evaluative justification was taken by Oddie to speak in favour for *realism* rather than naturalism, since the set does not correspond to any properly natural property¹⁴. Another way to deal with this unwholesome disjunctiveness, as we have seen, is to turn to mental state naturalism, i.e. once again by turning the back on the “object-interpretation” of naturalism. To question the relation between reasons and supervenience amounts to questioning our ability to get things right by reasoning alone. Reasons typically are pragmatic, instrumental; they correspond to ways of dealing with the world. And as we have seen, we are systematically led to believe that things other than mental states are good. This is practical for creatures like us, but it does not necessarily latch on to any deep fact about the nature of value.

Value Naturalism Unleashed

This concludes the unleashing. We have been dealing with a row of decisions for the naturalistically inclined value-theoretician to make. These decisions correspond to ways of reasoning, to theoretical commitments, and to problems you have to solve.

There are many ways of being a value naturalist. The only universal commitment of value naturalism is that the value it speaks of must be a natural property, accounted for without substantial reference to any non-natural properties. All else is based on theoretical decisions, and these decisions reflects what type of problem the philosophy of value is dealing with.

In this paper I have been concerned with making room for something I’ve been calling “mental-state naturalism”. I’ve suggested it as a theory that loosens up all of the restraints dealt with, but, of course, a mental-state type of naturalism can be conceived that *respects* almost all of these restraints as well.

Such a theory would retain the emphasis on value judgments and take its favoured mental states as the proper objects of attitudes, it could *not* accept unqualified mind-independence though, but it could take conceptual analysis, analytic or synthetic, as its preferred method, and it could say that the reason founding properties are the same as the supervenience basing ones. This would be a clear cut case of mental state naturalism. So why am I kidnapping it to this outer fringe, guerrilla-warfare-like existence?

It seems clear to me that naturalists do not have to accept the restraints mentioned, and that mental-state naturalists in particular shouldn’t. As such, naturalism is not committed to the way of framing problems formulated by naturalist theories with a wholly different agenda, and, as I’ve said, it is not even committed to solving the same problems. The person who places the value in the external objects of attitudes, for instance, has fundamentally *different* problems to deal with than the one that places it in the attitude laden experience itself. The account for motivation is one of these things.

The unleashing here affords other approaches to value than tradition has it. It opens up for value-theory being more than an isolated discipline, sometimes informed by philosophical semantics and economics. Rather, it puts it in touch with the philosophy of mind, with psychology and the cognitive and affective sciences. And from there on there is no telling

¹⁴ Frank Jackson (1998) has no problem with this, though, the disjointedness merely means that *talk* of moral properties might be indispensable: the ethical might be infinitely disjunctive and patternless.

were it might go (sociology? Political science?) For instance, instead of settling it from the beginning, the connection between value judgments and motivation is here something that we can *find out*.

This is what methodological naturalism is all about: not prejudging what type of problem we are dealing with, and not pre-judging what type of method could solve the problems that arise. In my view, naturalism has been far too restricted by the considerations mentioned, and I believe that mental state naturalism in particular has a lot to gain by changing the subject in this way.

[*References*

Peter Railton
Frank Jackson
Graham Oddie
Sharon Street
Copp
Richard Boyd
David Wiggins
Nicholas Sturgeon
Geoffrey Sayre-McCord
Arthur Prior
G.E. Moore
Kevin Mulligan
Mark Johnston
C.I. Lewis
Bennet Helm]