***Ben Davies. Unpublished draft. Last updated 28 September 2015.***

**1. Street’s Darwinian Dilemma for value realism**

Sharon Street (2006) argues that value realists face a ‘Darwinian Dilemma’ that casts doubt on their view. In Street’s terms (110), value realists are those who claim that there are some evaluative facts which “hold independently of all our evaluative attitudes”. A subset of value realism, also vulnerable to the dilemma, is moral realism, which makes analogous claims about moral facts. Street notes that most contemporary realists accept that evolutionary pressures have played a significant role in shaping our evaluative attitudes, which in turn ground our evaluative beliefs (119-120). If realists accept this claim, they have two options regarding the relationship between these evolutionary processes and the postulated evaluative facts.

The first option is that there is no relation, and we have no reason to think that any of our particular evaluative beliefs is true; it would be pure chance if a belief formed by processes that bore no reliable relation to the evaluative facts coincided with them. This does not mean that there will be no overlap, since we might get lucky; but this does not tell us whichof our evaluative beliefs has happened upon truth by chance. As such, we have no justification for any individual evaluative belief. On this view, even if realism is correct, it gives us no guidance in action or belief. This view also undermines at least one potential line of support for realism: the fact that some of our value judgements *seem* as if they correspond to some objective truth. There might be other reasons to be a realist, but finding an argument that does not involve this appeal somewhere may be difficult.

Alternatively, realists could claim that the relevant evolutionary processes *did* bear some relation to the evaluative facts. The relation Street considers (125) is that they ‘tracked’ the evaluative facts, such that we evolved the attitudes we did at least in part because they were true. As Street says, in advancing such an explanation, realists are offering an account that is open to assessment by scientific standards. But, she argues, the tracking account fails to be the most persuasive explanation of our evaluative attitudes when subjected to such standards. There is a more plausible explanation of why we evolved the attitudes we did: Street’s own “adaptive link account” (127) says that we evolved them because they were evolutionarily beneficial, not because they had any relation to independent evaluative facts. Individuals who felt the urge to kill their offspring as soon as ­they were born, for instance, would do poorly in evolutionary terms compared with nurturing individuals, and creatures that developed negative attitudes to such an action would have greater evolutionary success. Moreover, perceiving a useful activity or response as *demanded* by reasons external to the agent, rather than simply as a whim, gives it more motivational force than merely desiring it.[[1]](#footnote-1)

There is no additionalevolutionary advantage to having evolutionarily useful evaluative beliefs *also* track evaluative truths. So an appeal to evaluative truth is explanatorily otiose. Evaluative realism (whether according to a tracking account or otherwise) is also less parsimonious than an adaptive link account, since it posits independent evaluative facts where anti-realist accounts need not do so. Realist accounts are also explanatorily vague in the processes and relations that they posit; there is no explanation in the tracking account of how we arrive at evaluative facts, or why we would evolve to do so (129-134). These latter two problems seem likely to affect all realist accounts, unless they can give some scientifically plausible explanation of what ‘evaluative facts’ amount to, and the processes by which we access them. Street concludes that since evaluative realism does not feature in the best explanation of our evaluative beliefs, so we ought to reject the second horn.

**2. The ‘rationalist’ defence**

Katarzyna de Lazari-Radek and Peter Singer (LRS) have argued (2014)[[2]](#footnote-2) that a form of moral[[3]](#footnote-3)realism heavily inspired by Sidgwick (1907) can escape Street’s dilemma.[[4]](#footnote-4) For the majority of moral beliefs, we take Street’s first horn: evolution has not tracked the truth for most of our moral beliefs, and many of those beliefs are explained by the adaptive link account, or some other non-truth tracking account; so most of our moral beliefs are unjustified (181).

However, at least one quite fundamental moral belief is immune evolutionary debunking (182). This reverses the burden of argument; if the anti-realist appeal to evolutionary advantage cannot explain certain moral beliefs, and a realist story can, realists can claim explanatory advantage. To ground this fundamental moral belief, LRS note (119) that Sidgwick identifies two “self-evident” claims. The first I will call Universe Neutrality:

**Universe Neutrality:** “...the good of any one individual is of no more importance, from the point of view (if I may say so) of the Universe, than the good of any other”.

The second I will call Value Neutrality:

**Value Neutrality**: “...as a rational being I am bound to aim at the good generally...not merely at a particular part of it”.

From these two claims, Sidgwick claims to deduce the moral proposition that LRS (2012: 17) call Universal Benevolence:

**Universal Benevolence:** each person is “morally bound to regard the good of any other individual as much as his own, except in so far as he judges it to be less, when impartially viewed, or less certainly knowable or attainable by him”.

For Universal Benevolence, LRS offer an *indirect* tracking account, which does not face the problems of scientific implausibility that come from a direct tracking account. First, they claim that there is no direct evolutionary benefit to believing Universal Benevolence; so Street’s adaptive link account is not plausible for this belief. A creature fully motivated by Universal Benevolence would give no additional weight to her own survival, or that of her biological relatives, damaging her evolutionary fitness (182-3). So the kind of story Street offers about many moral beliefs is apparently unavailable.

Second, LRS claim to have an explanation of how evolutionary processes tracked the truth of Universal Benevolence. That tracking is not scientifically implausible because it is indirect. Sidgwick argues that the truth of Universal Benevolence is a deliverance of reason, based on the self-evident axioms of Universe Neutrality and Value Neutrality. If we reflect properly on those axioms, we will simply come to see that they are true, and entail Universal Benevolence. Evolutionary advantage comes not from a belief in Universal Benevolence itself, but from the rational capacity or capacities by which we come to that belief. They offer an analogy with complex mathematical beliefs (182): there is no evolutionary benefit from a belief in Fermat’s last theorem. But there is evolutionary benefit from a rational capacity that enables us to form this belief; the very same capacities are likely involved in forming evolutionarily useful complex beliefs.

Street does consider the possibility that realists could claim that our ability to grasp evaluative truths is “the byproduct or outgrowth of some other capacity which *does* have an explanation in terms of natural selection” (2006: 142). She suggests that realists will either have to make the implausible claim that this capacity involves a basic ability to grasp evaluative truths, or that a capacity that was selected for on the basis of factors that had nothing to do with evaluative truth just happenedto pick out the evaluative truth.

This is a reasonable response with regard to most capacities, for which the ability to pick out moral truths would be a rather unlikely side-effect. However, in the case of rationality LRS (183) appeal to the idea of parsimony that Street uses in motivating the Darwinian Dilemma. Reasoning is an extremely useful tool in evolutionary terms, but it is more likely that creatures would develop a quite generalrational ability, which will pick out some evolutionarily neutral (e.g. complex mathematics), and even disadvantageous (e.g. Universal Benevolence) truths, than a mechanism that would only arrive at the very specific rational operations that aided our survival (especially since the content of that class might change over time). Creatures that developed a reasoning-like ability that only picked out evolutionarily useful facts might fare better; but this is unlikely to develop, especially once a species starts down the more general path. Evolution is a satisficing mechanism, not an optimising one.

Of course, this assumes that the intuitive appeal of Universe Neutrality and Value Neutrality – the bases of the argument to Universal Benevolence – really are self-evident. LRS lay out three “elements in the process of establishing that an intuition has the highest possible degree of reliability” (195):

 1. Careful reflection leading to a conviction of self-evidence;

 2. Independent agreement of other careful thinkers; and

 3. The absence of a plausible explanation of the intuition as the outcome of an evolutionary or other non-truth-tracking process.

If Universe Neutrality and Value Neutrality pass all three criteria, we have good reason to view it as justified. We should note at this stage, however, that LRS’ realist story still potentially suffers some of the putative shortcomings that Street offers against all realist theories. First, it postulates independent evaluative facts and properties, so it may be open to the charge of lacking parsimony. And although LRS’ Sidgwickian appeal to reason does give us an explanation in some sense of how true moral beliefs fit into an evolutionary story, it is still extremely vague when it comes to exactly *how* it is that our rational capacity enables us to alight on Universe Neutrality and Value Neutrality. We do not reason *to* them from other premises (as we do Universal Benevolence); rather, we merely ‘see’ that they are true. But this appeal to brute rational insight is unclear at best.

The real strength of LRS’ argument, then, comes in the absence of a plausible evolutionary explanation. It is not that the rationalist explanation is necessarily a particularly compelling story, but that in the absence of a convincing anti-realist story it is the best available explanation of our attitudes. This means that an anti-realist explanation with as much or greater plausibility will have an explanatory advantage in scientific terms, since it will not face the additional problems of parsimony and vagueness of LRS’ explanation.

In Section 3, I argue that LRS’ reasoning from Universe Neutrality and Value Neutrality to Universal Benevolence fails. The argument can only be rescued at a cost to LRS’ claims that its premises are self-evident, and by relying on a broader evaluative realism that fails the Darwinian Dilemma.

**3. The universe doesn’t care**

LRS’ argument for Universal Benevolence relies on the following reasoning. Since my good is (self-evidently) of no more value from the point of view of the universe than anyone else’s (Universe Neutrality), and I am (self-evidently) rationally bound to aim at the good in general (Value Neutrality), I am morally bound to regard everyone else’s good as equal in importance with my own.

There are several potential criticisms of this reasoning,[[5]](#footnote-5) but I will focus on just one. This is the unstated assumption that once we have established that every person’s good is of equal importance, as Universe Neutrality has it, we have automatically established that it is part of ‘the good’ that features in Value Neutrality. If that were true, the injunction to aim at the good generally entails a requirement to treat each person’s good – including my own – with equal priority, as per Universal Benevolence. However, neither LRS nor Sidgwick offer any argument to the effect that my good and your good being of equal importance *with one another* entails that they are part of ‘the good’.

LRS might suggest that no such argument is necessary because to talk of ‘my good’ simply is to talk of a part of the general good that happens to belong to me. On this view it is impossible that something should be good for me but not part of the general good. But this assumption seems to me to be baseless.

I assume that to be part of ‘the good’ is to be of objective value. This then aligns with the most obvious reading of ‘the point of view of the universe’, which is simply as a shorthand for objectivity.[[6]](#footnote-6) Under this interpretation, the argument says that my good is of no greater objective value than yours, and that I am bound as a rational being to aim at all instances of objective value. But nothing in this expanded view entails that my good, or anyone’s, is of any objective value. And there are alternative ways of characterising the individual good such that this connection does not follow necessarily. One might, for instance, speak of something’s being ‘good for me’ when it serves my interests; but there is no need to think that something’s serving my interests is itself objectively good, or even that the satisfaction I derive from having my interests served is objectively good. Things might in this sense be ‘objectively’ good for me – if what is in my interests depends on something other than my subjective attitudes – and yet not part of the objective good. Alternatively, perhaps things can be *subjectively* good for me – such that something’s being good for me is reducible to my holding some attitude or mental state towards it[[7]](#footnote-7) – but there is in addition properly objective value to some other states of affairs that have nothing to do with how things are for us. In trying to establish justification for basic evaluative attitudes, LRS cannot assume these possibilities away.

LRS may insist that these alternative explanations do not properly capture the idea of an individual’s good; they describe only a kind of pseudo or apparent good. They have two options for such an argument. First, they might claim that it is simply self-evident that individuals’ good is part of the good. But the possibility of alternative relationships, such as those outlined above, undermines such an unargued appeal. While a claim to self-evidence need not mean that one sees the truth of a claim as soon as one understands it, the disagreement of other ‘careful thinkers’ (to put it in LRS’ terms) suggests a need for further argument.[[8]](#footnote-8)

Alternatively, they might appeal to Universe Neutrality. If each individual’s good is of equal value from the point of view of the universe (which I have taken to mean ‘objectively’) perhaps this implies that each person’s good is objectively valuable. And since Universe Neutrality is supposedly self-evident, this might seem to provide a self-evident basis for the belief that our good is part of the objective good.

It certainly seems deeply plausible that the only perspective from which my good could plausibly matter more than some other individual’s good is a parochial one (i.e. my own or that of someone who has some particular relation to me). Certainly my own good matters more *to me* than the good of many other people, in the fairly weak sense that I care about it more. But when I think about the idea that my good could matter more than other people’s, the only reasons I can conceive of for such a view are parochial.

However, this acknowledgement does not come anywhere close to agreeing that Universe Neutrality as LRS present it is self-evident. More strongly, it *cannot* be self-evident because it is insufficiently determinate, being ambiguous between two contradictory readings. One way that everyone’s good could have equal objective value is indeed if everyone’s good has some equal, positive objective value. But there is an alternative interpretation of Universe Neutrality. This is that everyone’s good is of equal objective importance because everyone’s good is of noobjective importance at all: the universe doesn’t care. This is the reading under which Universe Neutrality seems persuasive to me. My good does not objectively matter more than yours does because neither of our goods matters objectively.[[9]](#footnote-9) This is not self-evident; but nor is the former interpretation, that each person’s good has some positive objective value.

Claiming that it is self-evident that each person’s good is part of the objective good would be at odds with LRS’ own discussion of the argument for Universal Benevolence. In introducing the argument (134), LRS say that “taking the point of view of the universe is an essential part of the ground for [Universal Benevolence]...without that, it would be possible to take our own good as the general good”, excluding everyone else’s, and so assume that it is only our own good at which we should aim generally. This implies that if we accepted Value Neutrality alone we could acknowledge *that other people have goods*,but not think that their good was part of the general good, and so assume that we should be rational egoists, pursuing only the greatest possible benefit for ourselves over a lifetime.

This scenario does not involve a failure to recognise that things are goodfor other people. And it would be odd for LRS to agree with Sidgwick that we can recognise that other people have their own good, but not be rationally compelled to include it in the general good, and yet to also insist that we cannot think that *all* people, ourselves included, can have individual goods without all of those individual goods being part of the general good. If it is plausible that nobody’s good *but mine* has objective value (or at least, sufficiently non-obviously wrong that we need an additional premise to reject it), why should the addition of my own good to the ‘valueless’ category make such a stark difference? If it were self-evident that an individual’s good is part of the general good, LRS would not need to appeal to Universe Neutrality in the argument for Universal Benevolence; it would be enough to note that each individual has their own good, and that we are rationally required to aim at the good generally.

This means that Universe Neutrality is not sufficient, even taken jointly with Value Neutrality, to get us to Universal Benevolence. If the point of view of the universe is evaluatively indifferent, such that nobody’s good has any value from that perspective, this would imply that no individuals’ good is part of the general good. But then the maxim to ‘aim at the good generally’ gives us no advice on how to consider our own and others’ good. In deciding how to weigh up our interests against those of other people, we would have to appeal to more parochial concerns that do not contain the ‘essential’ perspective of the universe.

So LRS need each individual’s good (or, given Universe Neutrality, just one individual’s good) to have objective value, or to matter from the point of view of the universe. I have argued that an appeal to self-evidence is not persuasive in this regard, while Universe Neutrality is insufficiently determinate to provide the support that LRS require. So LRS need to provide a direct argument for the claim that our good is also objectively good, and so part of ‘the good’ that we are bound to pursue generally. As I will suggest in the following section, LRS may be seen as offering such an argument, but this argument itself relies on evaluative realism, and is thus vulnerable to the Darwinian Dilemma.

**4. An independent argument for the objective value of individuals’ good**

Prior to their discussion of Universal Benevolence in (2014), LRS offer a brief defence of the view that reasons for action are objective(‘reasons objectivism’), based on rejecting the subjectivist alternative that reasons and value are always “conditional on our having certain desires and wants” (2014: 45).[[10]](#footnote-10) Since the case on which LRS’ argument for reasons objectivism rests concerns an individual’s good, if this argument is plausible then it also supports the view that individuals’ good has objective value, and generates objective reasons. Combined with Universe Neutrality, this would imply that *all* individuals’ goods have objective value, and the original argument would be successful.

LRS’ defence of reasons objectivism is based solely on Derek Parfit’s hypothetical case of a person with ‘future Tuesday indifference’. Parfit (1984: 123-124) imagines a person – call him S – who resembles the rest of us except that he is indifferent to pain on future, although not present, Tuesdays. Street (2009) offers a useful fleshing out of this agent: S is told by his dentist that he can have a dental procedure next Tuesday or Wednesday, but that there will be a delay in the delivery of anaesthetic such that the Tuesday operation would be conducted without pain relief. S will currently prefer the Tuesday procedure if there are anycosts to the Wednesday procedure, even minor inconvenience, since he does not care about pain on future Tuesdays. However, when Tuesday arrives S’s preferences will change; indeed, he will likely be anxious enough to reschedule the appointment, since his attitudes towards pain on *present* Tuesdays match ours, and the operation will be agonising. S also knows himself well, and therefore knows when he makes the appointment that he will likely cancel it. But he is equally indifferent to the inconvenience that this will cause him, since it occurs on a future Tuesday. We can expand the case to make S’s apparent irrationality even clearer: if the dentist’s office regularly has the late Tuesday delivery, this pattern of scheduling and rescheduling seems likely to continue without end.

S should probably switch to a less chaotically organised dentist. But LRS restrict themselves to two claims, which they take to be sufficient to defeat reasons subjectivism. The first is that subjectivists must endorse S’s pattern of behaviour. The second is that S is “obviously crazy”, which I take simply to mean that S obviously has strong reasons not to book the Tuesday appointment, whatever his feelings, and no reasons to do so. Since subjectivists must endorse behaviour that an agent obviously has reason not to engage in, and no reason to engage in, we ought to accept that some things are good and bad for individuals in a way that is entirely independent from their present desires, and that these facts generate reasons for action that are also independent of an individual’s subjective mental states at any particular time. Presumably, the reason that S has to avoid the Tuesday appointment is that it will be painful, and that the pain will be bad for him, and this provides an objective reason to avoid an action even if one does not currently care about the future pain. If LRS’ argument is successful, then, they have shown that our future good gives us objective reasons for action; although this is not *quite* an argument that our future good has objective value, the argumentative leap is small enough to ignore here.

Why does S have reason to avoid the Tuesday appointment, despite his indifference to the pain he will feel? One explanation, typified in Street’s extension of the case, is that it leads him to a broader pattern of behaviour that is against his own interests. But it is important to note that subjectivists can just as easily insist that S has strong reasons to avoid the Tuesday appointment on these grounds. After all, being dragged by his own inconsistent behaviour into a time-wasting, toothache-worsening game of chicken with himself in the future is not something S *currently* wants; he is only indifferent to this stuff on future Tuesdays, not all the time. So long as S cares sufficiently about the inconvenience (and likely toothache) he will experience on future Mondays, Wednesdays, and so on, he has strong subjective reasons to commit himself, perhaps against *one* of his current inclinations, to a course of action, even if that course of action holds no intrinsic appeal to him.

On the other hand, LRS and Parfit’s case seems to rest on something more basic: the view that indifference to future Tuesday pain provides S with no reasons because it is arbitrary. On this view, we need not appeal to the broader consequences for S to insist that he has reason to avoid the Tuesday appointment; the pain he will feel is sufficient, given the presence of better options. Subjectivists seem unable to declare that S’s pain in itself provides him with reasons to avoid it. So if it does provide S with reasons, then subjectivism is undermined.

Before I explain why this argument fails, let me reconstruct the dialectic so far. LRS claim to have identified a moral belief (Universal Benevolence) which avoids Street’s dilemma because although not evolutionarily useful itself, it is reached by rational deliberation from self-evident premises (Universe Neutrality and Value Neutrality). I claim the argument fails because Universe Neutrality does not tell us that any individual’s good matters objectively, and so does not identify individual good as a target for Value Neutrality. I suggested that LRS might respond by offering an independent argument that our good is objectively good and so properly included in the scope of Value Neutrality. The best grounds they have is by utilising their existing argument that even someone who is indifferent to pain on future Tuesdays has objective reason to avoid such pain (even aside from its repercussions at other times) because such indifference is arbitrary, whereas pain is objectively bad. Since S’s pain in the future is bad whether or not he cares about it, this would suggest that S’s good is objectively good; combining this with Universe neutrality tells us that, because everyone’s good is equally valuable from the point of view of the universe, everyone’s good is objectively good.

In the next section, I note that the claim that such indifference is arbitrary is itself a realist evaluative claim. It rests on a distinction between things that ‘really’matter, and things which only appear to matter: pain matters, and which day of the week it occurs on does not. Section 5 points out that this realist intuition is itself vulnerable to the Darwinian Dilemma. A premise which is vulnerable to the very dilemma its argument seeks to escape undermines that argument.

**5. Why future Tuesday cases are problematic**

The reason it seems so obvious that S has reason to avoid the Tuesday appointment despite his indifference is that none of us *are* future Tuesday indifferent; for us, the fact that something will hurt us seems to demand certain responses, while the fact that something will occur on a future Tuesday demands no such response in itself. These are themselves evaluative beliefs: pain matters objectively; days of the week, absent some further context, do not. If we are to accept these intuitions then, by LRS’ lights, they should meet the three criteria for ‘highly reliable’ intuitions outlined in Section 2, one of which is that there is no plausible alternative non-truth tracking evolutionary account of the intuition. And if the argument is to avoid the Darwinian Dilemma, there should be no *better* explanation of our intuitive response to the future Tuesdays case that does not invoke objectivism.

However, such an account is available. Section 1 referenced the moral anti-realist argument that seeing certain acts as morally called for provides additional motivational force for agents than merely desiring them, or thinking them instrumentally useful to evolutionary purposes. A similar case is available here: it would be motivationally valuable for creatures to evolve such that they feel, and then believe, that certain things *really are* valuable and perhaps even called for, rather than merely thinking that we subjectively care about and value things. This sense would enable creatures that held it to bridge motivational gaps such as onsets of ennui and depression, or the short-term temptations that can prevent us from fulfilling longer term commitments. Think back to S, who is indifferent to future Tuesdays; a creature that evolved to be indifferent to certain stages of the future and hence failed to plan would not last very long in evolutionary terms.[[11]](#footnote-11) A creature that felt such indifference, but also felt that it had objective reasons to care about its future, would at least go some of the way toward overcoming that indifference. Since evolutionary fitness is concerned in part with long-term survival, escaping such short-term distractions would have been fitness promoting. Such a creature would do better from an evolutionary point of view than one which merely pursued whatever they most wanted in an instrumentally rational way because it would feel as though it had *binding* reasons to avoid distraction, temptation and weak will, reasons that transcended its own parochial concerns in precisely the way that LRS suppose.

Recall that this is not supposed to be an evolutionary explanation of a belief in Universal Benevolence itself, but of an independent argument for objectivity of reasons that LRS require to get their argument for Universal Benevolence going. But before I discuss my suggestion further, it is worth returning briefly to another important element of that same argument. This is the claim that we cannot appeal to an evolutionary story to explain our belief in Universal Benevolence, because such a belief would have been evolutionarily damaging, not helpful. An individual who was sole moral motivation was to give equal weight to the claims of all would not care much for themselves or for their children. But it is less obvious that it would be evolutionarily disadvantageous to come to a *partial* commitment to Universal Benevolence. An individual who felt the pull of universal benevolence, but felt far more strongly motivated by more parochial concerns – in other words, a creature like us – might only sometimes act against the interests of itself or its kin, not obviously threatening evolutionary advantage.

Of course, my explanation of out tendency to ascribe objectivity to our evaluations is only a story. But at least with respect to LRS’ criteria for reliable intuitions, my task is undemanding. LRS’s third reliability criterion is a lack of *any* plausible evolutionary, non-truth tracking explanations for an intuition. I have suggested that there is a plausible evolutionary explanation of our intuitive belief in objective value and reasons. Since the only support LRS offer in favour of reasons objectivism is intuitive, and that intuition fails the third criterion of reliability that they offer, their defence of reasons objectivism is inadequate. So if this is to be our independent defence of a positive version of Universe Neutrality, which Section 3 argued is necessary for the Sidgwickian argument, we lack a compelling case for Universal Benevolence.

LRS’ intuitive argument for reasons objectivism is also vulnerable to the Darwinian Dilemma. The dilemma requires not just that there is some plausible evolutionary explanation, but that it is more plausible than an explanation that invokes value realism, such that the latter does not figure in our best explanations of our attitudes. Insofar as both my explanation and LRS’ are ‘just-so’ stories, we may seem to be on a par here.

But it is questionable whether LRS’ appeal to rational reflection really does constitute an *explanation* of our attitudes. Consider again the additional advantages of anti-realist positions according to the Darwinian Dilemma, as discussed at the end of Section 2. Anti-realism is more parsimonious than realism, since the latter posits objective evaluative truths, whereas evolutionary accounts posit the kinds of things to which we are already committed. Anti-realist accounts are also less vague than realist accounts. Street says that realists fail to explain why believing evaluative truths would be evolutionarily advantageous. LRS can claim to have met this challenge; the beliefs themselves are not advantageous, but they come from a capacity which is. But they would still owe us an explanation of *how* reasoning puts us in touch with the evaluative facts, since they offer no explanation of how our reasoning capacity reliably delivers the belief in an objective evaluative fact (i.e. that each individual’s good is objectively valuable), or the objective reasons that derive from it. While ordinary uses of rationality put us in touch with features of the world – such as how relatively fast the bus and train will get me home – and apply them to particular ends, reasoning to evaluative truths does not work like this; realists need some alternative story about what is going on. Saying that something is a deliverance of reason is not an explanation of *how* reason operates to give us access to evaluative facts or reasons.[[12]](#footnote-12)

Evolutionary anti-realists, on the other hand, have an explanation of how a belief in reasons objectivism might come about in terms of relatively well-understood processes; they can appeal to the idea of Darwinian evolutionary selection and, as I have suggested, the view that a belief in objective reasons would be evolutionarily useful in terms of overcoming weakness of will. That is not to say that a failure to offer such an explanation automatically rules realism out; as I suggested in Section 2, if realism faced no otherwise plausible competition, LRS could declare that they had the best theory by virtue of the absence of plausible competitors; even if it is vague, the appeal to deliverances of reason might be better than *no* explanation. But if my explanation is at all plausible, then LRS’ realism *does* face such competition, and so loses its place as a default explanation.

Indeed, LRS seem to tacitly accept this line of argument in their treatment of other moral beliefs. One *could* offer a rationalist explanation for various common-sense moral beliefs, claiming that they seem self-evident to us and so must be deliverances of reason. LRS reject such an appeal because of the availability of a plausible evolutionary explanation for these beliefs. This suggests that the only reason we have to regard any particular evaluative belief as plausibly reflecting objective facts – at least insofar as we derive that belief from our intuitive responses – is the idea that our response *could not* plausibly be the result of a non truth-tracking evolutionary adaption.

LRS might appeal to the phenomenology of thinking about evaluative claims, and note that it typically *seems* to us that we have reached these conclusions by using reason; shouldn’t we take such phenomological data at face value? But this argument would apply just as much to other moral beliefs; and it is untrue that an anti-realist explanation cannot also explain the phenomenological data. Humans are certainly successful ‘reasons-finders’: we are adept at seeing genuine connections between various propositions and facts. But we are also voracious *reasons-seekers*. We have a strong tendency to seek out standards by which to define our lives and validate our projects and commitments. The thought that our existence and our most treasured commitments are ‘objectively’ valueless is a frightening and upsetting thought. So the fact that we should seek objective reasons about what most matters to us is hardly surprising. Perhaps our ability to successfully detect reasons comes as a broader package which also delivers a number of false positives.[[13]](#footnote-13) If there is an evolutionary advantage in having an over-active rather than an under-active reasons-seeking capacity, (it is more important to see reasons where there are none than to risk missing useful ones) and there is evolutionary gain in a belief in the objectivity of value, we can explain our sense that normative reasons are objective without appealing to the objective truth of that claim. So even if a belief in Universal Beneficence really is best explained “as the outcome of our capacity to reason” (193), this may not vindicate realism, because our capacity to reason may be liable to misfire. Our intuitive sense that we have reached a conclusion through reasoning is liable to mislead.

**6. Conclusion**

LRS make the strong claim that “there is no plausible explanation of [Universal Beneficence]...as the direct outcome of an evolutionary process, nor is there any other obvious, non-truth tracking explanation” (193). Rather, it is a self-evident conclusion of two self-evident truths: Universe Neutrality and Value Neutrality. But Universe Neutrality, if left sufficiently vague to be self-evident, does not play the role in an argument for Universal Benevolence that LRS want it to. Extending the argument with an independent defence of objective reasons moves the argument beyond claims of self-evidence and, at least on the view that LRS offer, relies on a broader evaluative realism. The intuitive ground for this extension is susceptible to plausible evolutionary explanations. As such, it fails LRS’ third criterion for highly reliable intuitions. Moreover, this defence does worse than anti-realist stories in terms of clarity and parsimony, and so does not feature in the best explanations of our intuitions and beliefs. As such, it is vulnerable to the Dilemma. A failure to make such an argument leaves Universe Neutrality ambiguous between LRS’ interpretation, and a nihilist view under which Universe Neutrality gives us no reason to apply Value Neutrality to individuals’ good, as we must in order to derive Universal Benevolence.

References

De Lazari-Radek, K. and Singer, P. (2014). *The Point of View of the Universe: Sidgwick & Contemporary Ethics.* Oxford: Oxford University Press.

Dorsey, D. (2012). Subjectivism without Desire. *Philosophical Review*, 121(3), 407-422

Field, H. (1980) *Science Without Numbers* New Jersey: Princeton University Press

Kahane, G. (2014). Evolution and Impartiality *Ethics*,124(2), 327-341.

Joyce, R. (2006) *The Evolution of Morality* Massachusetts: MIT Press.

Nichols, S. (2006) *Sentimental Rules: On the Natural Foundations of Moral Judgement* Oxford: OUP.

Parfit, D. (1984). *Reasons and Persons.* Oxford: OUP.

Sidgwick, H. (1907). *The Methods of Ethics*, 7th ed. London: Macmillan.

Street, S. (2006). A Darwinian Dilemma for Realist Theories of Value. *Philosophical Studies*,127(1), 109-166.

Street, S. (2009). In Defense of Future Tuesday Indifference: Ideally Coherent Eccentrics and the Contingency of What Matters. *Philosophical Issues*,19(1), 273-298.

1. See also Joyce (2006); Nichols (2006)*.*  [↑](#footnote-ref-1)
2. All bracketed page references will be to this work unless otherwise stated. [↑](#footnote-ref-2)
3. It is important to stress that this is explicitly a defence of realism about a particular moral claim, to be outlined below. I will return to the broader category of evaluative realism in Section 4. [↑](#footnote-ref-3)
4. I do not aim here to show that the dilemma is immune to all challenges, only that insofar as it is compelling, which LRS seem to accept for many moral beliefs, the rationalist argument is not a way around it. [↑](#footnote-ref-4)
5. For instance, one might question the move from a rational injunction in Value Neutrality to a moral injunction in Universal Benevolence. [↑](#footnote-ref-5)
6. At some points LRS seem to take the idea of the point of view of the universe more literally. For instance, they suggest (134-5) that we must adopt the point of view of the universe when evaluating actions, describing it as “a point of view outside one’s own dispositions, projects and affections”; since merely adopting someone else’s parochial perspective would not get us equality of value, I assume this is meant to be a perspective which lies, in a sense, outside all parochial perspectives. I am sceptical about the possibility of such a perspective, taken literally, but my comments about the indeterminacy of Universe Neutrality apply to this interpretation just as readily. [↑](#footnote-ref-6)
7. See, e.g. Dorsey (2012). [↑](#footnote-ref-7)
8. Even if it is self-evident, so that my argument is undermined, LRS face an alternative problem outlined by Kahane (2014). Kahane notes that even if the argument for Universal Benevolence succeeds formally, it does not show that anything actually *is* objectively good, so that at most we get a hypothetical injunction that *if* something is good for us, we should be impartial with respect to it between different people. This differs from my criticism: Kahane is concerned with the indeterminacy of what counts as individuals’ good; I am concerned with the question of whether individuals’ good has any value from the point of view of the universe, and so whether it falls under ‘the good’. [↑](#footnote-ref-8)
9. Indeed, it may be that nothing has objective value; from the point of view of the universe, perhaps nothing matters. But I do not need to defend this view for my argument to go through. [↑](#footnote-ref-9)
10. I think we should add other mental states such as evaluations to this list, but I will accept this definition for the sake of argument. [↑](#footnote-ref-10)
11. Presumably, the evaluative attitude would have to reference something less socially contextual than the idea of ‘Tuesdays’. [↑](#footnote-ref-11)
12. LRS appeal to an analogy with our knowledge of mathematical truths, which are also supposed to be self-evident deliverances of reason. But this does not clarify the process; it merely shows that other kinds of knowledge face the problem. And indeed, some people take an anti-realist position on mathematical knowledge e.g. Field (1980). [↑](#footnote-ref-12)
13. We can see this pattern in other of our cognitive abilities; think of our extremely useful facial recognition system that is reliable in discovering actual faces, but also sees deities burned onto toast, or a man in the moon. [↑](#footnote-ref-13)