

# Epistemic Welfare Bads and Other Failures of Reason

Antti Kauppinen

For *Midwest Studies in Philosophy*, special issue on ill-being, ed. Gwen Bradford

Final draft, November 26, 2022

One of the many terrible things cults and totalitarian states do to children is keeping them in ignorance, preventing them from learning about crucial aspects of the world and understanding it, or even worse, instilling a severely distorted conception of reality. It is not that the children are necessarily unhappy because of it, or that they always fail in their practical endeavors, either then or as adults. In theory, after all, a guru or a dictator might be genuinely benevolent. Indeed, to engage in some philosophical fantasy, people who have ended up ignorant in one way or another might even accidentally end up pursuing genuinely valuable ends and successfully realizing them. Still, I think most of us would agree that there is something very important missing from or otherwise wrong with the lives of such people. To put it in technical terms, such lives are blighted by lacking *epistemic welfare goods* like understanding and by containing *epistemic welfare bads* like false sense of understanding.

What I've just said is hopefully intuitively plausible. But it's less clear what the epistemic welfare goods and bads are and why they are good or bad for us apart from their consequences. One possibility, endorsed by many objective list theories of well-being, is that things like knowledge simply are in themselves good for us, just as pleasure might be, and no further explanation can be given. But it's not just disappointing if explanations come to an end so very soon, but also quite implausible when we start to think about epistemic welfare bads. Our beliefs might fall short in various ways – they might be false, unjustified, justified but luckily true, concern trivial matters, yield a partial picture, or stand in isolation

from each other. Without some kind of unified explanation of their relative value and disvalue for us, we end up relying on a jumble of intuitions.

The explanatory story I want to tell about epistemic welfare goods and bads is a new twist on one of the oldest ones on record. It begins with the basic ideas of *perfectionism* about well-being, the view that we do well or badly when we flourish or unflourish as the kind of beings we fundamentally are. On my variant, flourishing amounts to successfully realizing the formal aims implicit in the exercise of our fundamental capacities. I argue that the relevantly fundamental capacities are those whose operation defines who we are and consequently what fits or suits us, rather than those that form a “human nature”, as the more traditional conception has it.

However, in agreement with the traditional view, I hold that one of those self-defining capacities is *reason*. Reason, as I’ll understand it, is roughly speaking our power to shape our beliefs and intentions into coherent bodies that are well-grounded in available evidence. Because reason asks “why?”, it takes us beyond appearances and inclinations and puts us in a better position to grasp reality and make autonomous choices. Its formal aim, or the internal standard of success that is implicit in its use, is bifurcated, because it operates on mental states with opposing directions of fit. Practical reason, I’ll argue, formally aims at competently realizing self-chosen valuable ends that are in harmony with each other. In the best case, these amount to valuable achievements that give meaning to our lives. Crucially for our purposes, I argue that theoretical reason formally aims at competently grasping fundamental enough subject matters to form a comprehensive, deep, and unified explanatory model of the world that allows for prediction, seeing how things hang together, and answering why-questions. Such epistemic flourishing thus comes apart from agential flourishing in terms of achievement, contrary to some recent virtue epistemological claims.

According to my perfectionist account, then, competent inquiry that yields understanding of fundamental issues is the basic epistemic welfare good, since it constitutes success by the internal standards of theoretical reason. The absence of such understanding in our lives is one kind of epistemic welfare bad, privation of a welfare good. But are there robust epistemic welfare bads, that is, epistemic performances that in themselves make our lives bad for us to some extent? Taking a cue from Shelly Kagan's (2014) observation that the bads corresponding to complex goods come in many varieties, I distinguish between several different kinds of failure of both practical and theoretical reason. I argue that some of them are indeed robustly bad for us. Practical and theoretical failures differ in this respect, however. Among other things, I suggest that while the end we pursue matters more to agential flourishing than the means we take, in the theoretical case, *how* we inquire matters more for epistemic flourishing than *what* we inquire into, and grasping even unimportant matters can involve some excellence of reason.

### **1. Flourishing and the Formal Aim of Reason**

The list of theories of well-being that can explain the intuition that at least some kind of knowledge or understanding is in itself good for us is surprisingly short. After all, on any *subjectivist* view, it will be a contingent matter what kind of things are welfare goods for individuals, since it depends on their attitudes (typically, intrinsic desires or values) (Heathwood 2006, Tiberius 2018). So we can safely set aside such views as explanations of understanding (or anything else) as an inherent welfare good, insofar as we think understanding is good for you regardless of your attitudes. Insofar as there is a difference between understanding and having the *experience* of understanding, as there certainly seems to be – we sometimes have the illusion of understanding something, after all – *experientialist*

views of welfare like hedonism are also ruled out, even if we were to grant the implausible view that understanding is always pleasant.

We might, then, consider an objective list theory of well-being. Such a list could certainly include understanding – indeed, knowledge is a staple of objective list accounts (e.g. Parfit 1984, Fletcher 2013). Alas, the downside is that a list – a disjunction of separate welfare goods – provides us with no deeper explanation of *why* understanding might be a welfare good. No wonder it is an important part of the project of objective list theorists to argue that all explanations come to an end, and that the end comes quickly in the case of well-being (see especially Fletcher 2013). However, even an objective list theorist may be able to offer *some* unifying account of at least some welfare goods. In this vein, Thomas Hurka (2020, 591) argues that knowledge and achievement instantiate the same value, that of *rational connection to reality*, in one direction or another (the mind matches the world or the world matches the mind). He can thus be seen as offering a little shorter objective list (though not that short, since his includes also goods like friendship, virtue, and pleasure in Hurka 2010). Still, it's natural to want more – what's in it for me to have a rational connection to the world, when lucky success would get me what I want just as well?

Quite quickly, then, we're left with just one promising candidate among traditional theories of well-being: *perfectionism*, which promises to give a principled answer to the question about the welfare value of understanding and achievement in terms of *flourishing as the kind of beings we fundamentally are*. To distinguish clearly between different possible variants of perfectionism, I will break it down to a number of theses, starting with the most abstract basic idea (Kauppinen forthcoming a):

#### *Explanatory Perfectionism*

What is in itself good for a subject S are the things that constitute flourishing as the kind of being S fundamentally is. What is in itself bad for S are the things that constitute unflourishing as the kind of being S fundamentally is.

Given this basic starting point, any variant of perfectionism will involve some view of what we fundamentally are, or our *nature*, as well as a view of what *flourishing* or thriving as such a being amounts to. For traditional human nature perfectionists, we are fundamentally human beings, and flourishing consists of developing the capacities that are essential to being human and exercising them well by their own internal standards (see e.g. Hurka 1993). Reason has traditionally had pride of place among such capacities. This obviously goes back to Aristotle, who famously held that "the characteristic activity of a human being" is "activity of the soul and actions in accordance with reason", and that we flourish when we "carry this out well and nobly" over a complete life (NE 1098a).

While human nature perfectionism can pretty straightforwardly explain why it is good for us to contemplate fundamental truths or why it is bad for our intellectual development to be stunted regardless of whether it makes us feel bad, it does face some familiar and fundamental challenges. First, from the perspective of contemporary biology, there is no such thing as human nature or essence (as noted by Kitcher 1999, among others), and even if we grant that we can meaningfully talk about such things, it seems that there are some distinctively human capacities whose development and exercise does not seem to benefit us. Second, there appears to be a very significant gap between being *a good human being* and leading a life that is good for *you* (e.g. Sumner 1996, Bradford 2017). As subjectivists about well-being are wont to argue, for something to be in itself good for us, it must fit or suit *our own nature* rather than our nature as a member of a kind (Rosati 1996, Dorsey 2017).

So I believe that what perfectionists need is a conception of who we fundamentally are that doesn't involve questionable empirical or metaphysical assumptions and that consists of capacities whose proper use is plausibly linked with our own good. A good starting point for this is the subjectivists' own alternative, according to which our own nature consists in our desires and values that define where we ourselves stand and consequently what kind of things suit us. One somewhat underappreciated problem for subjectivism is that it seems our desires and even values can be alien to us, for example when they result from manipulation or indoctrination. In discussions about autonomy, it is pretty standard to understand manipulation in terms of bypassing our rational capacities (e.g. Mele 2008). As a first pass, then, we might say that our desires and values are our own if they *do* result from our own exercise of rational capacities (which need not involve explicit reasoning). This, then, suggests that who we fundamentally are is not to be understood in terms of any *states* we may have but in terms of our *activities*, or of how we *respond* to what is given to us or what we find ourselves with, such as appearances and desires. As Kantians like Christine Korsgaard (2018) emphasize, while the responses of most other animals may be governed by instinct, we are condemned to take a normative stance to such natural inputs – to treat some things as reasons for theoretical or practical commitments, beliefs or intentions, or (as I would add) implicitly evaluate them in terms of valenced feelings.

My thesis, then, is that our own nature, certainly in the sense that matters for the fit or suitability that is essential to determining which things are welfare goods and bads, is defined in part by *how we exercise such capacities for normative self-governance and valenced feeling*, rather than the values or desires that may result from doing so. *Reason*, as I'm going to understand it, is precisely the capacity for normative self-governance. I thus defend the following claim (see again Kauppinen forthcoming a):

*The Subjective Nature Thesis*

Our fundamental capacities in the sense relevant for well-being are those whose functioning defines who we are, or our *self-defining capacities*. In the adult human case, they include at least the practical and theoretical rationality and the capacity for valenced experience.

Unlike human nature perfectionism, this view of what we're fundamentally like doesn't appeal to dubious biology, but rather to plausible views about the nature of our self. It is only the first step in reformulating perfectionism, however, since it also requires giving an account of what constitutes *flourishing* as the kind of beings we are. Here the traditional core idea was that we fare well when we *exercise our fundamental capacities successfully by their own standards*. I think it is a good starting point. But there are several ways of understanding both successful exercise and the particular standards involved. Sometimes perfectionists speak as if reasoning well, for example, constituted success in exercising reason. While this is an understandable thought, I believe it is important to emphasize that the exercise of our capacities contains a *telos* or formal *aim* of bringing about some outcome – for example, forming the right intention doesn't yet constitute success in using practical reason, which (I'll argue) aims to effect the right changes in the world. I'm thus in full agreement with George Sher when he says that “what has inherent value is not the mere exercise of a fundamental capacity, but rather its successful exercise as measured by the achievement of its defining goal” (1997, 202; cf. Bradford 2021 on the value of ‘proper outputs’ of capacities). So here's my thesis about flourishing:

*The Telic Interpretation of Flourishing and Unflourishing*

Flourishing consists in successfully realizing the formal aims implicit in the functioning of our fundamental capacities to a sufficient degree. Unflourishing in

some respect consists in frustrating a formal aim, or realizing it to an insufficient degree.

So to find out what it is for us to flourish in terms of our capacity for reason, we must ask ourselves: What are reason's own standards? When do we succeed at normative self-governance? To answer these questions, we need to think about what reason *is*. I'm going to use the term rather broadly for our power to respond to reasons, including both implicit uses, such as subtly adjusting our credence in the possibility of rain in response to the shape and colour of clouds without conscious effort, and explicit reasoning (see also Kauppinen 2021a). Evidently, we can succeed or fail at exercising it – and we can fail at exercising reason even if we are, by dint of good luck, successful at our personal goals. The exercise of reason thus seems to come with internal standards. That's why we can say it has an overarching formal aim in addition to what our particular ends may be.

To come to grips with this formal aim, let's first consider what happens when we explicitly reason about what to do. Suppose I'm reasoning about whether to buy a dog for my son for his birthday. I want to get him something that he likes and that would help him grow to be a good person, and getting a dog occurs to me as an option. If I reason about it, I obviously don't just go for it, as I might pick a can of Coke among many, but consider the pros and cons of having a dog in the house and their relative weight (cf. Enoch 2010, 71–75). A big part of this deliberation involves considering how having a dog would bear on my existing commitments to other things, which impose constraints on how much time and space there is for a pet in our family. While the person-level goal of my reasoning is just to settle the dog question, for the process to amount to reasoning, it must be implicitly guided by the aim of getting it right, and getting it right non-accidentally, in virtue of having weighed the reasons correctly within the constraints of my ongoing commitments, so that the



choice will fit in with my overall ‘plan’ in life. On the basis of similar considerations, George Sher says that “basing one's decision on one's weightiest combination of reasons is the generic aim of all practical deliberation” (1997, 205). (To distinguish my richer conception of the internal standards of reason from Sher's, I'll talk about formal rather than generic aims.)

So at the most abstract level, the formal aim of practical reason is responding correctly to reasons for action. But we can be a bit more specific if we distinguish between different features of intentional action. An intentional action always has an *end* (or ‘goal’ – I'll use these terms interchangeably) to which we take some *means*, which results in some *outcome*. The end of the action is something that is more or less *worth realizing*, and something that more or less fits together with the other ends we have. The means we choose to take are more or less efficient and economical in realizing the end, given our credences and beliefs about empirical facts – for short, our choice of means is more or less *competent*. And finally, the outcome is that either the end is *brought about* or it isn't, and if it is brought about, its realization is more or less due to the exercise of our competence rather than factors beyond our control, and thus more or less attributable to us rather than good outcome luck. Bearing in mind that the formal aim of reason is something that tells us what the point of subjecting our inclinations and desires to systematic scrutiny is, we can say on the basis of these distinctions that the formal aim of practical reason is *to competently realize a harmonious set of self-chosen worthwhile ends*.

How about reason in its theoretical use? Clearly, a major part of what we're doing when we're engaged in theoretical reasoning is trying to figure out what is true. But just as clearly, what's distinctive of using reason to acquire true beliefs is trying to figure out *why* something is true and how it fits in with the rest of what we believe. We are – that is to say, we *must be*, in order to be engaged in reasoning – aiming to *secure* the truth of what we

believe and fit our beliefs into a *coherent* picture of the world, or a part of it. And of course, we're trying to figure out the truth about *something*, some particular subject matter that is a part of our whole picture of the world. In parallel with reasoning about action, we can thus distinguish three elements with respect to reasoning about how things are: the subject matter, ways of forming beliefs about the subject matter, and the resulting state that match or mismatch between mind and world.

What does reason aim at in terms of these three elements? It should be helpful to examine the possibilities in some more detail. To begin with, the subject matter can be more or less *fundamental*. One way in which A can be more fundamental than B is that truths about B hold in virtue of truths about A, but not vice versa. Another way is that grasping A can serve as the basis for explaining more or fewer truths, directly or indirectly. This is not easy to quantify exactly – after all, already the identity conditions of ‘a subject matter’ are vague and context-dependent (Is physics a subject matter? Astrophysics? Planetary physics? Physics of Mars? And how many truths are there about any of these things?). Nevertheless, it's intuitively clear that understanding human biology puts you in a position to answer more questions than understanding baseball, and some truths about baseball hold in virtue of human biology (that is, certain facts about biology explain why the rules are the way they are, why individuals with certain physical characteristics do well in the game, and so on), while the opposite is not the case.

Second, we can reason about the subject matter about it more or less *competently*, by relying on more or less reliable methods of recognizing and responding to reasons. What is involved in such reasoning varies depending on the nature of the subject matter – good reasoning about metaphysics likely involves different methods from reasoning about physics. Here we can understand reasoning narrowly, in terms of making inferences from what we already believe or from the evidence we possess, or broadly, as a process of rational

inquiry that may include gathering new evidence, experiments, deference, and cooperation. The latter captures excellent use of reason better – if the project is figuring out what there is and why, why restrict ourselves to a base that consists only in what’s already in our minds? At the same time, it’s important to note that not all competent ways of forming beliefs amount to using reason. It’s nothing like reasoning when you step outside and form the belief that it’s raining on the basis of perception. However, here we must remember that the gap between reasoning and perceptual belief is narrowed if you have the kind of background beliefs about the reliability of your perceptual faculties that epistemological internalists require for justified belief. After all, you will then in principle be capable of giving an argument to support your perceptual belief.

Finally, reasoning or reason-governed inquiry will conclude with some output, unless it is somehow interrupted. The output can be a true or false belief (“Yes, there is a bear behind the tree”). But when the subject matter is broader than an individual factual question, the results may be correspondingly broader. They may include a set of related beliefs - consider, for example, what you believe about the events of 9/11, if you’ve ever looked into the matter. (It wasn’t an inside job.) Such beliefs can be mutually supportive, or fail to be such. Finally, beyond a set of beliefs, rational inquiry can result in *grasp* of the subject matter, or what is sometimes called *objectual understanding* (Carter and Gordon 2014). This is what is at issue when we talk about understanding particle physics, for example, or understanding a car engine. Now, a part of understanding car engines is knowing, or at least truly believing, a body of propositions concerning them – pistons do this, carburetors do that. But that is clearly not enough. One must grasp how these facts are related, perhaps in particular the various explanatory and dependency relations among them (e.g. Kvanvig 2003, 18). Many emphasize that this involves grasping the truth of various counterfactuals – were the fuel ignited at this point, the piston would move like that, and that would move the

crankshaft so and so. Here ‘grasping’ is often thought to go in some way beyond belief or knowledge, and involve an *ability* to draw inferences about similar and counterfactual cases and possibly also follow and give explanations of why the relevant dependency relations obtain (Grimm 2006, Hills 2015). So if you understand car engines, you’re able to infer that if the choke had opened fully, the air pressure in the carburetor’s venturi tube had been lower and the fuel flow consequently higher. Such ability is obviously useful in constructing and repairing engines.

Once we consider these options, it seems obvious what the aspiration of theoretical reason as such is. The point of using reason – regardless of the particular purpose for which we use it – is to contribute to forming a mutually supportive body of beliefs that are not only true but also competently formed and held, and thus gaining an ability to answer questions, at least for ourselves, about what there is and what depends on what. Because of the built-in striving for explanatory depth and breadth, theoretical reason aims inquiry towards the more fundamental rather than the less fundamental. Here’s how I would summarize this in a thesis:

*The Formal Aim of Theoretical Reason*

In its theoretical use, it is constitutive of exercising reason to aim at competently grasping a fundamental enough subject matter to contribute to a comprehensive and coherent explanatory model of what there is.<sup>1</sup>

Let’s take stock of where we’ve arrived. The reason why we’re interested in the formal aims inherent in the operation of our self-defining capacities, including reason, is that for a perfectionist, they yield a standard for flourishing and thus well-being. It follows from what

---

<sup>1</sup> I say ‘model’ rather than ‘picture’, since talk of models is more suggestive of being capable of counterfactual inferences (see e.g. Pearl and Mackenzie 2018).

I've just argued for that we fare well as practical reasoners to the extent that we competently realize a harmonious set of worthwhile ends. This, then, is one of the welfare goods, the things that are in themselves good for us, setting consequences aside – call it *agential flourishing*. What about theoretical reason? The preceding considerations support the following thesis:

*Perfectionist Epistemic Welfare Goods (PEWG)*

One of the basic welfare goods is competently grasping a fundamental enough subject matter that contributes to having a comprehensive and coherent explanatory model of what there is.<sup>2</sup>

Again, for the perfectionist, this sort of rich objectual understanding is good for us, because it constitutes a part of flourishing as the kind of beings we fundamentally are, or in terms of the capacities that define who we are (on my particular variant of it). For short, I'll call it *epistemic flourishing*, which, I emphasize, abbreviates *doing well in epistemic respects* rather than *doing epistemically well*, that is, meeting whatever the correct epistemic standards are for belief and other states. (Thus, for example, perceptual knowledge of my immediate environment might constitute doing epistemically well without contributing to epistemic flourishing.) To go back to Hurka's claim I discussed earlier, this is the

---

<sup>2</sup> This thesis, formulated in terms of a kind of understanding, contrasts with what is probably the most detailed account in the perfectionist tradition, defended by Thomas Hurka (1993; 2010; 2020). Briefly, in his most recent work, Hurka argues that knowledge is the more valuable the more general it is in two respects, intrinsically or relatively. The *intrinsic* generality of a truth "turns on its extent, or how far the state of affairs it affirms extends in space, time, and the number of objects it involves, so more general truths include more and less general ones less" (2020, 600). *Relative* generality is a matter of "how many other truths you've used this one to explain and therefore understand" (ibid., 601). I don't think intrinsic generality has any role in epistemic flourishing – knowing the salt content of the Atlantic is in itself no better for me than knowing the salt content of a rock pool in West Cork. Relative generality comes closer to what I think makes a subject matter fundamental, with the important difference that on my view, you don't need to actually believe the truths that grasping the subject matter allows you to explain.

perfectionist's answer to the question of what is good for us about having a "rational connection to reality".

## 2. Understanding and Achievement

In the last section, I argued that certain kind of knowledge and understanding regarding explanatorily fundamental matters is in itself good for us, if we accept a plausible version of perfectionism about well-being. As I see it, fit between theory and intuition should increase our confidence in each of them. The theory is further supported if we can show that the explanation it offers is superior to competing accounts. Here, I will concentrate on one recently prominent competitor, namely the view that knowledge and understanding are finally valuable as *achievements*. This type of view is notably defended by (reliabilist) virtue epistemologists, who believe that knowledge *as such* is a kind of achievement, a performance that is successful sufficiently through competence rather than luck (Sosa 2007, Greco 2010). Combining this account of the *nature* of knowledge with the thesis that all achievements are finally valuable either intrinsically or prudentially yields a simple explanation of final intrinsic or prudential *value* of knowledge. It should be noted that it's not clear whether virtue epistemologists themselves want to claim that knowledge is finally valuable in either of these ways – for example, in places Sosa emphasizes that his claims concern final *epistemic value*, which may come apart from other kinds of value (e.g. Sosa 2007, 88). But elsewhere Sosa himself suggests in a perfectionist spirit that through some kinds of knowledge, understood as an achievement, "you flourish as a *rational* animal" (Sosa 2021, 6). So it is of interest whether this type of account could account for the possible prudential value of epistemic goods. This requires exploring what makes achievements prudentially valuable.

How does the virtue epistemological account compare with my view? To begin with, I'll grant for the sake of argument that knowledge is at least apt belief, or belief that is true because of competence that is part of the agent's epistemic character, or whatever specifications virtue epistemologists want to add, so that it can be considered as a kind of achievement.<sup>3</sup> Under this assumption, the main questions will concern the value of achievement. First, is it really the case that all achievements as such are finally prudentially valuable? Hardly so. Even the most liberal defenders of the value of achievement, such as Simon Keller (2004), only argue that it is good for agents to realize their goals, something that *they* aim at. Even if we think that *beliefs* 'aim' at truth or knowledge, it doesn't follow that *agents* necessarily do. And in any case, more credible accounts of the value of achievement hold that for achievements to be valuable, they must be *difficult*, as Gwen Bradford (2015a) emphasizes, or perhaps in some different way *test the limits* of a capacity worth exercising, as Sukaina Hirji (2019) argues. Only such 'capital-A achievements' (or 'A-achievements', as I'll call them for short) are "endeavors that are particularly noteworthy in some respect, and evoke a sense of awe, admiration, and of being impressed" (Bradford 2015a, 4), which makes them plausible candidates for having final prudential value. And few if any apt beliefs will count as A-achievements (Bradford 2015b) – consider easy perceptual or testimonial knowledge (Lackey 2007).

In response to this problem with easy epistemic achievements, Duncan Pritchard (2010) proposes that instead of knowledge, we should think of *understanding why* as the finally valuable epistemic good, since it is more plausibly an A-achievement. After all, it

---

<sup>3</sup> For criticisms of virtue epistemology, see e.g. Lackey 2007, Pritchard 2010, and Hirvelä 2019. For my purposes, these arguments and refinements made in response can be set aside – after all, the virtue epistemologist can always restrict her value thesis to those instances of knowledge that are (at least) apt beliefs.

may not be easy to understand why bond yields rise when bond prices fall, for example.<sup>4</sup> But as Adam Carter and Emma Gordon observe, sometimes understanding why is no A-achievement either – it’s pretty easy to understand why the dryer doesn’t work, if one sees that it’s unplugged (2014, 5). As they point out, objectual understanding is a better candidate for A-achievement. More precisely, at least *some* instances of objectual understanding no doubt constitute A-achievements – it’s definitely hard to come to understand crypto markets, for example, or I’d be a lot richer. The same goes for certain kinds of knowledge. So if it is the case that all A-achievements are finally prudentially valuable, we have here an alternative explanation of the final prudential value of some epistemic states, while others will be relegated to having purely instrumental value.

But are all A-achievements in themselves good? It is indeed a standard view in the literature on achievement, taken by both Hurka (2010; 2020) and Bradford (2015a). Both hold that A-achievements have *intrinsic* value, and in later work, Bradford (2021) suggests that A-achievement is a basic *welfare good*. But is this true? Here it is important to focus on what makes something an A-achievement. As Bradford emphasizes, an achievement involves both a *process* that aims at something and the *product* of that process when it is successful. (Note that the ‘product’ may itself be an activity or performance, like singing a song.) On her view, roughly speaking, what is key to A-achievements is that the product is competently caused by the agent and the process involved in this is *difficult*, which she spells out further in terms of requiring a high degree of total intense effort (Bradford 2015a, 50). (There are various ways to get to difficulty *tout court* from being difficult *for* some individuals, but this issue need not detain us here.<sup>5</sup>)

---

<sup>4</sup> It’s because yield is based on the bond’s interest payments divided by its market price, or so I’ve learned from Investopedia: <https://www.investopedia.com/ask/answers/061715/how-bond-yield-affected-monetary-policy.asp>

<sup>5</sup> For difficulty, see Bradford 2015a; von Kriegstein 2019; Hirji 2019; Kauppinen forthcoming b.



Importantly, for Bradford, the product of the process need not itself be objectively valuable for the achievement to be finally (prudentially) valuable: “There are many particularly valuable and impressive achievements with products that have no value on their own. Climbing Mt Everest is a perfect example” (2015a, 86). While Hurka’s account of the process appeals to complexity rather than effort, he agrees that the product of A-achievements need not be of value: “A process of pursuing an end can have value even if the end itself doesn’t; this is clearest in games and sports” (2020, 593). For Hurka, the value of achieving a goal depends in part on its place in the hierarchy of goals – in particular, the more subgoals must be achieved to realize a goal, the more value realizing the top goal has. So for him, “part of the worth of climbing Everest ... is the many different actions it requires” (2020, 605).

It is certainly a commonsense view that climbing Mount Everest is an A-achievement, a notable event that can play a momentous role in one’s life. But I’m skeptical of whether such A-achievements without valuable product are *finally* prudentially valuable. Once we carefully rule out confounders, it’s hard to see why it would be better *for me* to make an effort and bring about an outcome by way of exercising my abilities if that outcome is itself worthless. One obvious confounder is that especially when we make an effort to achieve something, we get emotionally invested in realizing it, and consequently feel positive emotions if we do achieve it and negative ones if we don’t. An experience characterized by such responses, surely, is finally good or bad for us, which may make it hard to see that the achievement itself isn’t. What we must ask, then, is whether it would be good for me to succeed at getting to Mount Everest if I wasn’t in any way pleased to get there, and wouldn’t mind at all if I didn’t get there.

Perhaps more importantly, however, we must also rule out another, less obvious confounder, namely the *aesthetic* value of skillfully striving to overcome obstacles,

highlighted by C. Thi Nguyen (2019) in his account of games as exercises of alternate agency. Nguyen emphasizes the positive aesthetic experience of a “harmony between self and challenge” in a good game, but I think we should also say that there is something aesthetically valuable – something *beautiful* – about accomplishing the difficult task of climbing up a mountain, especially when it involves skillfully coordinating many different actions. If this is right, climbing Mt Everest can indeed have a kind of final value, depending on how you get there, but it will be aesthetic rather than prudential. And insofar it *does* have prudential value independently of how it makes you feel, it will have it because its ‘product’ – in this case a beautiful climb, not getting to the top as such – has aesthetic value. In contrast, an inelegant, stupid, laborious, and joyless climb up to the top, however difficult and complex, will not be in any way good for you, even though it will be an A-achievement by Bradford’s and Hurka’s criteria.

If what I’ve just argued is right, knowledge and understanding won’t be finally prudentially good even when they’re A-achievements, at least simply in virtue of being such. However, I do believe that even if not all A-achievements are finally good, there is a truth in the immediate vicinity. Recall what I said above about the formal aim of *practical reason*: the point of using it is competently realizing a harmonious set of worthwhile ends. This means that among other things, it is finally good for you to bring about an objectively valuable outcome by way of exercising your abilities to overcome challenges (where, again, the outcome may not be separate from the activity, as in the case of governing well). In such cases, you will precisely be exercising your rational competence to a high degree, and thereby bringing about a worthwhile end. This is part of what it is for you to flourish as the kind of being you fundamentally are, since it amounts to successfully exercising one of your self-defining capacities. It is no coincidence that this is also the kind of thing that makes a person’s life meaningful – not just in the sense that it makes it *feel* meaningful (again,

climbing a mountain might do that), but that it actually does give a point to her efforts (cf. Wolf 2010, Kauppinen 2021b).

Note that if the ‘product’ of achievement is that which culminates the activity, it does not need to be *finally* valuable to be *objectively* valuable. If you invent a drug that cures many people’s long Covid, it is objectively *instrumentally* valuable, because it brings about finally valuable relief from suffering. If you master the piano after hard work, it has what Joseph Raz (2011) calls *facilitative* value, since it makes it possible for you to perform something that has final aesthetic value (even before you perform). These are the kind of efforts that are worth making, and being successful at them is in itself good for you. (To be sure, we can say that we bring these things about *for the sake of* a finally valuable outcome, which we might also think of as the product of the achievement.<sup>6</sup>) Whether an achievement with a more valuable product is greater *as an achievement* is debatable, but what matters for our purposes is just that it has more prudential value.<sup>7</sup>

What does this alternative perfectionist account of the value of achievement as a realization of the formal aim of practical reason mean for the value of epistemic achievements? Well, there are certainly epistemic performances that qualify as valuable practical achievements by the above criteria. Suppose, for example, that you come to understand the circulation of blood as a result of sustained inquiry. This understanding puts you in a position to give good medical advice and treat diseases – it is a facilitative objective good. So by the above criteria, arriving at such understanding by yourself is a finally prudentially valuable achievement, and something that in itself makes you better off. (It is less of an achievement to come to the same understanding as a result of being taught at a

---

<sup>6</sup> I thank Gwen Bradford for clarification on this point.

<sup>7</sup> If we think that a greater achievement can have less prudential value than a smaller one, what Bradford (2015a) labels the “essential value” of an achievement (the value it has in virtue of the features that make it an achievement) comes apart from its prudential value.

medical school, though it may be equally or more instrumentally useful.) The same goes for many other possible objects of understanding and knowledge. For the purposes of possible action, self-understanding and understanding one's social and physical environment will be particularly valuable. In a very different way, the same goes for moral understanding and understanding other sorts of evaluative and normative domains, which are crucial not only for selecting which means to take but also which *ends* to pursue.

What follows if it is the case that not all achievements nor even all A-achievements are finally prudentially valuable, and only achievements whose outcome is objectively valuable are? First of all, this means that we can explain the welfare contribution of only a small subset of knowledge and understanding in terms of achievement. And in these cases, the explanation is the same as for other goods of *practical* reason, cast in terms of realizing something of objective value by using reason, even though in this case it is one's own epistemic state that has objective value in virtue of what it's apt to lead to. Second, to be clear about the kind of prudential value involved, we need to distinguish between the *practical* achievement that consists in *coming to understand* some challenging subject matter, and the *theoretical* success that consists in *understanding* itself, or holding the subject matter in our grasp. While the achievement account can in the best case explain the former, it can't explain the latter, since it is not a process whose product is objectively valuable – it's not a process at all, but something like an ability, or a set of dispositions and beliefs. PEWG, in contrast, does say that some forms of understanding and knowledge gained via rational inquiry are in themselves good for us. Here we might think of something like gaining a novel understanding of cell biology as a result of carefully designed experiments. This sort of understanding puts us in a position to answer many why-questions and unify many explanations at a deep level, and thus fulfills the *telos* of theoretical reason to a high degree.

Third, there are of course many instances in which coming to understand an explanatorily fundamental matter does have objective practical value, as in the case of understanding medicine or one's own society. In such cases, coming to understand may be a valuable practical achievement, and the understanding itself will be an instance of an epistemic welfare good. Consequently, other things being equal, it is better for you to study a fundamental subject matter that also has practical value – say Earth geology rather than exoplanet geology. But of course other things are rarely equal. For one thing, because practically important subject matters have naturally been explored the most intensely, there is less room for major practical achievements involving creative use of reason. For another, it tends to be a long way from fundamental subject matters to practically valuable applications. So in practice, there will be tradeoffs between valuable epistemic achievements and epistemic welfare goods, or in other words, between practical and theoretical excellence. I will not try to formulate any method for weighing them here.<sup>8</sup>

Finally, an important consequence of explaining epistemic welfare goods in terms of success in the use of theoretical reason is that true beliefs, justified beliefs, and knowledge as such will *not* be welfare goods. Here the implications of my version of perfectionism differ from objective list theorists like Parfit and Hurka. We can, after all, come to know things without using reason, such as by perception and testimony – though insofar as knowledge involves internalist justification, the true beliefs that constitute it will plausibly qualify as minimally rational as well. True beliefs and knowledge are, to be sure, typically instrumentally good. But on my view, they don't in themselves benefit the subject, unless they are *constituents* of understanding a fundamental enough subject matter. Again, this does not rule out their having *epistemic* final value, which means that what I've said here is

---

<sup>8</sup> For discussion of such trade-offs, see Hurka 1993, ch. 7. I add a few things in Kauppinen 2009.

orthogonal to many debates about the value question in epistemology (for which see e.g. Zagzebski 1996).

### 3. Epistemic Welfare Bads as Failures of Reason

As I said in the introduction, though my aim in this paper is to discuss epistemic welfare bads, it is best approached once we're armed with a theory of epistemic welfare goods. Having sketched such an account, we are finally in a position to focus on prudentially bad epistemic states and performances. There are two preliminary points to consider. As Shelly Kagan's (2014) groundbreaking work on ill-being has made clear, we must distinguish between two ways of being bad. Some things are bad in the sense that their presence entails the absence of a welfare good, or *privatively bad*. Not being happy is privatively bad for you. If we represent someone's well-being numerically, privative bads contribute zero to someone's well-being. Other things are bad in themselves, or *robustly bad*. Unhappiness is a robust bad. Its contribution to someone's well-being could be represented by a negative number. So one thing to bear in mind is that there can be both privative and robust epistemic welfare bads.

Second, as Kagan also importantly emphasizes, whenever a welfare good has a complex structure, as understanding or achievement does, there are many ways to fall short of it, which may not all be equally bad (2014, 277ff). For example, Kagan observes that if we think of knowledge as the epistemic good, lacking a belief, having a justified false belief, and have a false belief against one's evidence are all ways of falling short of it. Yet intuitively, they're not all equally bad. But how can we rank their badness in a principled way, and which of them are robustly bad and which merely privatively? Kagan himself holds that false beliefs are bad for you, though not as bad as unjustified false beliefs. Hurka, in contrast, maintains that false beliefs are *not* robustly bad, and are no worse than the absence

of belief about the topic (2020, 599). On his view, however, unjustified false beliefs are robustly bad, because the believer will have “flouted rather than followed rational norms” (ibid.). For Hurka, this is a case of a kind of organic unity – the disvalue of unjustified false beliefs isn’t the sum of the disvalue of unjustified belief and the disvalue of false belief, since otherwise unjustified false belief and unjustified true belief would be just as bad, given that truth and falsity as such have no value or disvalue for him.

So, objective list theorists disagree about the badness of epistemic performances that fall short of knowledge or understanding. But as list theorists, they will have little to say about why a particular view of epistemic welfare bads would be better than another, other one ordering being more intuitive. This is where the perfectionist account of epistemic welfare goods I’ve just defended should be helpful.

### *3.1 Failures of Practical Reason*

As I did above in considering the successful use of reason, I’ll start with the parallel case of failures of practical reason. First, though, we need to be clear about the relevant notion of failure. Suppose it would be good for me to build a dam, but I don’t make any effort to do so, maybe because I don’t realize it would be good for me. In the sense I’ll be focusing on, this does not constitute a failure, but just the absence of a potential good. If building a dam is the best thing I could do and I have access to all the relevant evidence, we can indeed say that failure to pursue this goal is an error of practical reason. But such absence of pursuit of good is very different from failed pursuits, whether their aim is good or bad. I will focus here on the latter.

I argued that the formal aim of practical reason is competently (with respect to *means*) realizing (as an *outcome*) a harmonious set of worthwhile ends (as the *goal* in a narrow sense). For simplicity, I’ll set aside here the kind of failure that consists in undermining other

ends that one is committed to, and focus on a single activity. An agent can then fail to be successful by the internal standards of practical reason in terms of either the means, outcome, or goal, or a combination of these. Here are some of the most salient options:

- a) S competently pursues a sufficiently valuable goal, luckily realizes it.
- b) S incompetently pursues a sufficiently valuable goal, luckily realizes it.
- c) S fails at pursuing a sufficiently valuable goal in spite of competent pursuit.
- d) S fails at pursuing a sufficiently valuable goal because of incompetent pursuit.
- e) S competently pursues an insufficiently valuable goal and realizes it because of competence.
- f) S competently pursues an insufficiently valuable goal, luckily realizes it.
- g) S incompetently pursues an insufficiently valuable goal, luckily realizes it.
- h) S fails at pursuing an insufficiently valuable goal in spite of competent pursuit.
- i) S fails at pursuing an insufficiently valuable goal because of incompetent pursuit.

How can we rank these failed performances in terms of their final prudential value by the standard given by the formal aim of practical reason? As a preliminary, note that as the difference between *a* and *c*, for example, highlights, the success of any endeavour always depends on circumstances beyond our control being sufficiently favourable. (Even success “because of” competence isn’t due to competence *alone*, as virtue epistemologists acknowledge.) For short, I will say that the world can be hospitable or inhospitable. In the perfectionist framework I’ve sketched, then, the obvious starting point is that S can exercise reason in aiming or in pursuit, or both, or can fail to do so, and that the world can be hospitable or inhospitable. To keep things manageable, I’ll assume in the following that when S pursues a sufficiently valuable goal, it’s because she’s rightly responding to reasons



for pursuing it. Similarly, I'll assume that competent pursuit involves responsiveness to instrumental reasons. Neither of these need always be the case.

With these assumptions, in the scenarios above, S manifests reason in aiming in *a*, *b*, *c*, and *d*. She manifests reason in pursuit in *a*, *c*, *e*, *f*, and *h*. And finally, the world is hospitable in *a*, *b*, *e*, *f*, and *g*. Just eyeballing these scores for performances suggests immediately that *a* comes closest to the meeting the formal aim of practical reason (the only thing that is missing is that her success is too lucky to count as manifesting reason), while *i* is the farthest from it. To rank the options in between, we need additional principles about the relative importance of these formal criteria for success and failure. In the case of practical reason, it seems to me that setting the right goal is more important for flourishing than pursuing it competently. One principled explanation for this might be that our goals say more about who we are than how we pursue them, so insofar as flourishing has to do with exercising self-defining capacities, how we set them counts for more. Second, competent pursuit also says more about us than the world being hospitable – indeed, success in realizing a goal doesn't so much require good luck when the pursuit is competent, but just the absence of bad luck. And the kind of 'good luck' that leads to success that isn't due to competence even if one has exercised one's competence – Sosa's (2007) second gust of wind that blows a competently fired arrow back on course after an unlucky gust of wind has thrown it off – doesn't detract much from prudential value. Third, another plausible principle is that it's good for us for the world to be hospitable only when our goal is worthwhile in the first place. That is to say that success in the narrow sense of realizing our goals is valuable only as a component of organic unity composed of competent success and valuable goal – it's not that we add the independent value of realizing just any goal to the value of competent pursuit, but rather that once you pursue a worthwhile goal, you're better off realizing it than not doing so.

The following performances meet two of the three criteria: *b*, *c*, *e*, and *f*. Of these, *b* manifests reason in aiming and *c*, *e*, and *f* in pursuit. The world is hospitable in *b*, *e*, and *f*. In addition, *e* is the only scenario of success because of competence, which I glossed above as manifesting reason in success in the narrow sense. Given the above principles, I'm pretty confident that *b* and *c* contribute more to S's flourishing than *e* or *f*, and that *e* is better for her than *f*. That's because the right goal is more important than competent pursuit or even success because of competence. Since in *c* S manifests reason both in aiming and in pursuit in spite of failing due to bad luck, my principles suggest that it comes closer to realizing the formal aim of practical reason than *b*, which is only luckily successful. This means a ranking of  $c > b > e > f$  for these options.

The remaining options are *d*, *h*, and *g*, where *d* manifests reason in aiming and *h* in pursuit, while the world is hospitable in *g*. Here the ranking is straightforward according to my principles:  $d > h > g$ . We've thus arrived at the following ranking of failures of practical reason, in order of the least bad to the most bad:

- 1) S competently pursues a sufficiently valuable goal, luckily realizes it. (*a*)
- 2) S fails at pursuing a sufficiently valuable goal in spite of competent pursuit. (*c*)
- 3) S incompetently pursues a sufficiently valuable goal, luckily realizes it. (*b*)
- 4) S competently pursues an insufficiently valuable goal and realizes it because of competence. (*e*)
- 5) S competently pursues an insufficiently valuable goal, luckily realizes it. (*f*)
- 6) S fails at pursuing a sufficiently valuable goal because of incompetent pursuit. (*d*)
- 7) S fails at pursuing an insufficiently valuable goal in spite of competent pursuit. (*h*)
- 8) S incompetently pursues an insufficiently valuable goal, luckily realizes it. (*g*)
- 9) S fails at pursuing an insufficiently valuable goal because of incompetent pursuit. (*i*)

One remaining question is where, if anywhere, on this list we move from privative prudential badness (absence of contributors to well-being) to robust prudential badness (presence of contributors to ill-being). As Jason Raibley (this issue) argues, it's by no means obvious that there are robust welfare bads corresponding to each welfare good. But it does seem to me that at least *i* is robustly bad. How can we tell? Here's the litmus test: setting consequences aside, are you worse off having engaged in pursuit in the first place than if you had refrained from doing so? After all, not pursuing an end whose competent realization would have benefited you is a paradigm case of absence of goodness. And *i* certainly intuitively meets that test. *g* is not far behind – after all, it involves bringing about, by sheer luck, a goal one shouldn't have pursued in the first place. What is there to be said for having engaged in such pursuit over not doing anything, if we set aside all possible consequences?

Is there some principled criterion we could formulate? It is a little tricky, given the multidimensionality of success and failure. On the one hand, we have the clear cases of success and failure where everything goes right or wrong. (Let's not forget, though, that not every success is equally good, since some sufficiently valuable goals are more valuable than others, and being due to competence, too, is a matter of degree.) On the other hand, in between them we have *partial* success and partial failure. One way to describe things is to say that any partial success is *also* a partial failure, which makes it hard to say when we cross over to the robust side. But I think we can place the accent on one or the other. Consider making a competent effort at settling a pay dispute even if you ultimately fail due to factors beyond your control (type *c* scenario). For this to count as a competent effort in the first place, you must have had a well-thought-out, realistic plan that you were able to carry out, and you must indeed have carried it out to some extent – you've persuasively addressed the parties, brought them at the same table, proposed a reasonable compromise – only to be undermined by some contingency like a meddling relative. This is clearly a partial

*success* in terms of the formal aim of practical reason, and something that contributes positively to agential flourishing.

In contrast, consider *g* again. Here your insufficiently valuable aim might be destroying an eagle's nest, and your plan involves throwing paper planes at it. By sheer luck, one of them hits a bear you didn't realize was sleeping in the nearby bushes, and as it instinctively swings its powerful arm around, it happens to shake the tree so that the nest falls down. Success! But still, a partial failure of practical reason. While realizing one's goal is part of the formal aim of practical reason, the point about the value of such success being an organic unity was that it only contributes to agential flourishing if you had sufficient reason to pursue the goal in the first place. What if you had pursued the bad aim competently, say by skillfully throwing an appropriate size rock just in the right direction from a nearby ledge, but failed because of a freak accident of someone flying a drone in the way (scenario *h*)? Here we have a mix of partial failure and partial success that no longer seems robustly bad. So perhaps the principle is this: engaging in an activity constitutes agential unflourishing as a robust welfare bad when the agent has no sufficient reason either to adopt the end or take the means.

### *3.2 Failures of Theoretical Reason*

Let us now turn to theoretical reason and epistemic welfare bads. I argued above that true beliefs and knowledge as such are not welfare goods. By the same token, I don't believe that false beliefs or ignorance (as absence of knowledge) are welfare bads. They don't necessarily constitute failures of theoretical reason. For the same reasons as in the practical case, the absence of reasoning about something it would be good to understand also won't constitute a failure of theoretical reason – if you're not using your reason for something, you can't fail at using it. Here there's an interesting asymmetry between the theoretical and the

practical, however. Suppose that it would be good for me to understand quantum physics and, contrary to fact, I could come to understand it without an excessive opportunity cost. In that case failing to pursue this epistemic welfare good would amount to an error of *practical reason*, since questions about which ends to pursue are practical questions. In this sense, we can say that the practical use of reason is primary. On this view, then, we can talk about failures of theoretical reason only in context of reasoning about or more broadly inquiry into some subject matter.

To see what such failures amount to, let's again recall that success in theoretical reason amounts to something like competently grasping a fundamental enough subject matter. Assuming that we are engaged in rational inquiry, there are again at least three ways in which this could go wrong – in terms of subject matter, conduct of inquiry, and success in grasping.<sup>9</sup> I'll talk about reasoning here, but as noted above, it should be understood broadly to incorporate rational inquiry in general and even unconscious uses of reason. Here are some salient (perhaps partial) failures of theoretical reason in these terms:

- a) S reasons incompetently about a fundamental enough subject matter, but luckily grasps it.
- b) In spite of competent reasoning about a fundamental enough subject matter, S fails to grasp it.
- c) S reasons incompetently about a fundamental enough subject matter and fails to grasp it.
- d) S reasons competently about a shallow subject matter and grasps it.
- e) S reasons incompetently about a shallow subject matter, but luckily grasps it.
- f) S reasons competently about a shallow subject matter, but fails to grasp it.

---

<sup>9</sup> For simplicity, I'll set aside the fourth aspect, the distinction between success because of competence and success that is competent but lucky, and assume that if competent efforts are successful, they are such because of competence.

g) S reasons incompetently about a shallow subject matter and fails to grasp it.

How should we rank these failures? Again, it seems pretty clear that *g* has nothing going for it, so it's the worst kind of failure, a borderline comic one. You tried to figure out whether Michael Douglas was born before or after his father, and couldn't work it out. To be sure, it could be even worse: you could have got it wrong and thought you got it right. These are the two different kinds of failure of grasping a subject matter: not getting an answer (and consequently suspending belief) and getting the wrong answer. The latter is clearly further from realizing the formal aim of theoretical reason and thus prudentially worse.

So the worst option is clear enough. But which failure is the best, or biggest partial success? Recall that in the case of practical reason, I argued that adopting the right ends counts for more than adopting the right means, since our ends reveal more about our selves or character than how we pursue them. But in the case of theoretical reason, I suspect it's the other way around. How we go about inquiring says more about our epistemic character than what we inquire into. To be sure, the thrust of theoretical reason is toward the more fundamental matters and thus answers to deeper and more broadly applicable answers to why-questions. But as part of the primacy of practical reason, what we inquire into is also legitimately influenced by practical interests. It's thus no wonder that how we conduct inquiry seems epistemically primary. Second, there's a major difference from the practical case when it comes to reason's involvement in success. The paradigmatic case of practical success is bringing about a state of affairs. But insofar as theoretical success is objectual understanding, and understanding requires an ability to make certain inferences, *reason is constitutively involved in epistemic success*. Correspondingly, grasping a subject matter makes an independent contribution to epistemic flourishing. No matter how you come to grasp astrophysics, if you do, you're manifesting theoretical excellence to some degree.

If we accept that competent inquiry and grasping a subject matter both display theoretical reason to a greater extent than what we look into, it is *d* rather than *a* that is the least bad of the failures above. That is, it's better for you to reason well and as a result come to a deep understanding of chess than reason badly but luckily come to grasp atomic physics equally well, even though chess is a comparatively shallow subject matter. But how does competent failure with respect to a fundamental subject matter compare with lucky success with a shallow subject matter – does inquiring well about a deep issue count for more than luckily coming to be able to reason well about something that doesn't explain much? (A complication here is that both inquiring well and understanding are a matter of degree.) For example, think about a scenario in which a scientist conducts a physics experiment using devices that she has excellent reason to be reliable, but which yield a set of random but plausible results, so that she ends up endorsing a deeply mistaken hypothesis. Contrast that with studying chess by observing a bunch of kids playing soccer in the park, which by way of an amazing coincidence triggers a chain of associations in your mind so that you gain a deep understanding of chess. Both of these are clearly cases of partial success. But if we use sympathy as a heuristic for which is worse for a subject, it seems that an outsider knowing the facts about both inquiry and its outcome should sympathize more with the scientist than you, since your very poor epistemic activity led to a positive epistemic outcome. This suggests that *e* is after all superior to *b* from the perspective of epistemic flourishing.

The remaining options are *f* and *c*, both of which fall short of grasping the subject matter. Since *f* involves competent inquiry, it is less of a failure of reason than *c*, in spite of its subject matter. So in all, the picture is as follows, from the best to the worst:

- 1) S reasons competently about a shallow subject matter and grasps it. (*d*)
- 2) S reasons incompetently about a fundamental enough subject matter, but luckily grasps it. (*a*)

- 3) S reasons incompetently about a shallow subject matter, but luckily grasps it. (*e*)
- 4) In spite of competent reasoning about a fundamental enough subject matter, S fails to grasp it. (*b*)
- 5) S reasons competently about a shallow subject matter, but fails to grasp it. (*f*)
- 6) S reasons incompetently about a fundamental enough subject matter and fails to grasp it. (*c*)
- 7) S reasons incompetently about a shallow subject matter and fails to grasp it. (*g*)

Again, we're faced with the question of which if any of these failures of reason are robustly rather than privatively bad. For reasons analogous to the practical case, both *c* and *g* seem to be robustly bad. We epistemically unflourish when neither the inquiry nor its outcome involves proper responsiveness to reasons, though it's less bad if we're at least exploring the right sort of thing, as in *c*. The other cases, I think, are better seen as partial successes with respect to the formal aim of theoretical reason, and thus at best privatively bad. Indeed, different imperfect forms of grasping a subject matter seem to me to be positively good for a subject.

## 5. Conclusion

I have argued that if we adopt a perfectionist theory of well-being according to which success in terms of the constitutive formal aims of reason is in itself good for us, we can explain why failure that consists of lack of understanding or merely apparent understanding is privatively or robustly bad for us. When understanding is demanding and has objective (practical) value, it's also true that coming to understand is in itself good for us in virtue of constituting a valuable achievement. One thing to add to earlier observations at this point is that since understanding fundamental enough subject matters is a welfare good, if *my*



coming to understand something fundamental but practically useless, such as the origins of the universe, will potentially help *others* come to understand it, too, making such discoveries counts as a valuable achievement by the same logic as making discoveries in medicine – both are apt to make the lives of others better in one respect. In this roundabout way, there are good practical reasons to engage in fundamental inquiry that benefits others *only* in terms of helping them gain understanding.

I can't pretend to have given a full or even a satisfactory defense of the perfectionism about well-being that my account of epistemic welfare goods and bads relies on. But I do think that its implications for these matters are independently plausible. Given a holistic conception of normative inquiry, this match should increase our confidence in both perfectionism and the intuitive judgments. To be sure, given the focus on the proper and successful use of reason rather than belief-forming capacities in general, the perfectionist account ends up saying that some common contenders for welfare bads, such as false beliefs and ignorance, are not in themselves bad for you unless they contribute to lack of understanding or misunderstanding. They are, of course, often instrumentally harmful regardless. I think this is in fact an intuitive conclusion once we clearly distinguish between prudential and epistemic value.

To wrap up, I want to briefly consider the implications of this account for three issues about epistemic welfare goods and bads. First, I started this paper by talking about the evils of depriving children of education and indoctrinating them. They are both apt to discourage inquiry into fundamental questions, and in the case of indoctrination, result in being content with false answers in case one does engage in reflection. In this way, indoctrination is likelier to result in unflourishing rather than just the absence of flourishing. But it's worth noting that not just any kind of education will suffice for flourishing in epistemic respects either, since merely knowing a set of important truths falls short of the

ideal of understanding set up by the formal aim of reason. This point resonates with Mill's unsurpassed defence of freedom of speech in *On Liberty*. As he notes, even if we happen to be taught all the right things about matters on which there is room for reasonable disagreement, without being exposed to the strongest counterarguments, ideally presented by their most passionate partisans, we're unlikely to acquire the ability to see and show where they go wrong (Mill 1859/2003, 115). Mill emphasizes that in the absence of access to grounds for preferring our own view to its alternatives and the ability to defend it, we will at least to some extent lack understanding of the subject matter – our grasp of it will be weak at best. And as he says, “this is not the way in which truth ought to be held by a rational being” (Mill 1859, 114). So from the perspective of flourishing in epistemic respects, Mill is right to hold that if we're unlucky enough not to encounter real-life opponents of our views, “it is indispensable to imagine them, and supply them with the strongest arguments which the most skilful devil's advocate can conjure up” (1859/2003, 116).

Second, it's worth noting that there may be exceptions to the rule that false beliefs and ignorance are only instrumentally harmful. Suppose that standing in certain friendly relationships to other people is in itself good for you. (For a perfectionist of my favourite kind, this requires the operation of our capacity to relate to other subjects as such to be a self-defining capacity, while traditional perfectionists can appeal to the sociality of human nature.<sup>10</sup>) Then if having certain true beliefs about other people, yourself, or the relationship is constitutive of the relationship itself, which seems plausible – how could you be friends with someone without correctly believing that they're favourably disposed toward you? – ignorance is privatively bad for you in virtue of constitutively and not merely causally depriving you of a welfare good (see also Baril in this issue).

---

<sup>10</sup> For the good of standing in right relations to others, see Laitinen forthcoming. I hope to develop this aspect of perfectionism elsewhere.

Finally, it will not come as a surprise that a perfectionist account ends up recommending what Socrates called the examined life, a life in which we manifest excellence of practical reason by dedicating some of our time to inquiring into what is truly important in a way that manifests excellence of theoretical reason, and consequently come to firmly grasp what really matters. It does not, as such, entail the famous Socratic thesis that an unexamined life is not worth living, since the absence of such epistemic welfare goods might be compensated for by other welfare goods. But it does suggest that for an unexamined life to be a good one, we have to be fabulously lucky in terms of those other goods, or perhaps blessed by the Gods.<sup>11</sup>

## References

- Baril, Anne (this issue). Doxastic Harm. *Midwest Studies in Philosophy*.
- Bradford, Gwen (2015a). *Achievement*. Oxford University Press.
- Bradford, Gwen (2015b). Knowledge, Achievement, and Manifestation. *Erkenntnis* 80:97-116.
- Bradford, Gwen (2017). Problems for Perfectionism. *Utilitas* 29 (3):344-364.
- Bradford, Gwen (2021). Perfectionist Bads. *Philosophical Quarterly* 71 (3):586-604.
- Carter, J. Adam & Gordon, Emma C. (2014). On Pritchard, Objectual Understanding and the Value Problem. *American Philosophical Quarterly*.
- Dorsey, Dale (2017). Why should Welfare 'Fit'? *Philosophical Quarterly* 67 (269):685-24
- Fletcher, Guy (2013). A Fresh Start for the Objective-List Theory of Well-Being. *Utilitas* 25 (2):206-220.

---

<sup>11</sup> The Gods blessed me in the form of comments and challenges by Gwen Bradford, Jaakko Hirvelä, Maria Lasonen-Aarnio, Max Lewis, and Lilian O'Brien.

- Greco, John (2010). *Achieving Knowledge: A Virtue-Theoretic Account of Epistemic Normativity*. Cambridge University Press.
- Grimm, Stephen R. (2006). Is Understanding a Species of Knowledge? *British Journal for the Philosophy of Science* 57 (3):515-535.
- Heathwood, Chris (2005). The Problem of Defective Desires. *Australasian Journal of Philosophy* 83 (4):487–504.
- Hills, Alison (2015). Understanding Why. *Noûs* 49 (2):661-688.
- Hirji, Sukaina (2019). Not Always Worth the Effort: Difficulty and the Value of Achievement. *Pacific Philosophical Quarterly* 100 (2):525-548.
- Hirvelä, Jaakko (2019). Knowing Without Having The Competence to Do So. *Thought: A Journal of Philosophy* 8 (2):110-118.
- Hurka, Thomas (1993). *Perfectionism*. Oxford University Press.
- Hurka, Thomas (2010). *The Best Things in Life: A Guide to What Really Matters*. Oxford University Press.
- Hurka, Thomas (2020). The Parallel Goods of Knowledge and Achievement. *Erkenntnis* 85 (3):589-608.
- Hurka, Thomas (2021). Against ‘Good for’/‘Well-Being’, for ‘Simply Good’. *Philosophical Quarterly* 71 (4):803-22.
- Kagan, Shelly (2014). An Introduction to Ill-Being. *Oxford Studies in Normative Ethics* 4: 261-88.
- Kauppinen, Antti (2009). Working Hard and Kicking Back: The Case for Diachronic Perfectionism. *Journal of Ethics and Social Philosophy* (1):1-10.
- Kauppinen, Antti (2021a). Rationality as the Rule of Reason. *Noûs* 55 (3), 538–559.
- Kauppinen, Antti (2021b). Against Seizing the Day. *Oxford Studies in Normative Ethics* 11: 91-111.

- Kauppinen, Antti (forthcoming a). Telic Perfectionism and the Problem of Pain. In Mauro Rossi and Christine Tappolet (eds.), *Perspectives on Ill-Being*. Oxford University Press.
- Kauppinen, Antti (forthcoming b). Creativity, Spontaneity, and Merit. In Alex King and Christy Mac Uidhir (eds.), *Philosophy and Art: New Essays at the Intersection*. Oxford University Press.
- Keller, Simon (2004). Welfare and the achievement of goals. *Philosophical Studies* 121 (1):27-41.
- Kitcher, Philip (1999). Essence and Perfection. *Ethics* 110 (1):59-83.
- Kraut, Richard (2007). *What is Good and Why: The Ethics of Well-Being*. Harvard University Press.
- von Kriegstein, Hasko (2019). On Being Difficult: Towards an Account of the Nature of Difficulty. *Philosophical Studies* 176 (1):45-64.
- Lackey, Jennifer (2007). Why We Don't Deserve Credit for Everything We Know. *Synthese* 158 (3):345-361.
- Laitinen, Arto (forthcoming). Mutual Recognition and Well-Being: What Is it for Relational Selves to Thrive? In Onni Hirvonen and Heikki Koskinen (eds.) *Theory and Practice of Recognition*. London: Routledge.
- Mill, John Stuart (1859/2003). *Utilitarianism and On Liberty*. Edited by Mary Warnock. Blackwell.
- Mele, Alfred R. (2008). Manipulation, Compatibilism, and Moral Responsibility. *The Journal of Ethics* 12 (3-4):263-286.
- Nguyen, C. Thi (2019). Games and the Art of Agency. *Philosophical Review* 128 (4):423-462.
- Parfit, Derek (1984). *Reasons and Persons*. Oxford University Press.

- Pritchard, Duncan (2010). *The Nature and Value of Knowledge: Three Investigations: Three Investigations*. Oxford University Press.
- Raibley, Jason (this issue). Less than Zero. *Midwest Studies in Philosophy*.
- Rosati, Connie S. (1996). Internalism and the good for a person. *Ethics* 106 (2):297-326.
- Setiya, Kieran (2017). *Midlife: A Philosophical Guide*. Princeton University Press.
- Sher, George (1997). *Beyond Neutrality: Perfectionism and Politics*. Cambridge: Cambridge University Press.
- Sosa, Ernest (2007). *A Virtue Epistemology: Apt Belief and Reflective Knowledge*. Oxford University Press.
- Sosa, Ernest (2021). *Epistemic Explanations*. Oxford University Press.
- Sumner, L. W. (1996). *Welfare, Happiness, and Ethics*. Oxford University Press.
- Tiberius, Valerie (2018). *Well-Being as Value Fulfillment: How We Can Help Each Other to Live Well*. Oxford University Press.
- Zagzebski, Linda Trinkaus (1996). *Virtues of the Mind: An Inquiry Into the Nature of Virtue and the Ethical Foundations of Knowledge*. Cambridge, England: Cambridge University Press.