

# Perspectives and Good Dispositions

Maria Lasonen-Aarnio

First Draft

Rutgers Epistemology Conference

## I. Miners, the Access Diagnosis, and Perspectivism

### Miners

Let me begin with well-known type of case, and the broad kind of view it has been taken to motivate:

**Miners.** 10 miners are trapped in one of two shafts (shaft 1 or shaft 2), and floodwaters are rising. You must decide which shaft to block before finding out where the miners are. They are no more likely, given your evidence, to be in 1 or 2. You are able to block the water from reaching one of the shafts, but you don't have enough sandbags to block both. If you manage to completely block the shaft where the miners are, they are all saved; if you block the other shaft completely, they all drown. If you do nothing, letting both of the shafts will halfway with water, one miner will drown in any case.<sup>1</sup>

What ought you to do?<sup>2</sup> Such cases are puzzling, many think, for we seem to be pulled in two different directions. It is clear what the *best* option is in this case: blocking whatever shaft the miners are in, thereby saving all of them. On some views, this gives rise to an objective sense of 'ought': if all the miners are in shaft 1, then you objectively ought to block shaft 1. However, that cannot be the end of the story. For everyone seems to agree that you would merit some kind of positive evaluation for letting both shafts flood halfway, and at the very least, you'd have an excuse for doing so, or be blameless for doing so. And were you to choose to block shaft 1, perhaps you could be blamed.<sup>3</sup> Further, many think that appeal to excuses and blame is not enough: it is true to say that you ought to do nothing, letting both shafts flood halfway.<sup>4</sup> That is what you ought to do, even if you know that it is not the best option.<sup>5</sup>

Indeed, though not everyone agrees, there is a long tradition of distinguishing between objective and subjective senses of 'ought', or objective and subjective notions of moral rightness.<sup>6</sup> Similarly, many distinguish between the ought of advisability and an ought of rationality.<sup>7</sup> According to a view popular in epistemology, a parallel distinction is needed when considering what we ought

---

<sup>1</sup> E.g. Regan (1980: 265), Parfit (1988). There are many structurally similar cases, such as Jackson's (1991: 462-463) case of *Dr. Jill*.

<sup>2</sup> The question is standardly understood to be what you overall *morally* ought to do, and this is how I will understand it. Those who think there are more subjective and objective senses of 'ought' typically think that such senses arise for different kinds of oughts (e.g. Wedgwood 2016).

<sup>3</sup> See Olsen (2017, §4) for a discussion of the limits of the Objectivist view appealing to notions like blamelessness.

<sup>4</sup> For instance, Kolodny and MacFarlane (2010: 115) say that it is obvious – largely independently of one's moral views – that one ought to block neither shaft.

<sup>5</sup> Another kind of case often given in support of more subjective oughts is one in which a subject has every reason to think that she is doing what is objectively best, but in which her choice results in a tragic outcome. For instance, in Holly Smith's (2010: 64) *Twin Towers I* example, a security guard tells office workers to evacuate the building using stairs during 9/11, having every reason to think that that is the safest option. However, using the stairs takes too long, and tragically, everyone dies.

<sup>6</sup> This was a rather standard view among moral philosophers working in the first half of the twentieth-century. For some representative examples from the past half decade, see Jackson (1986), Gibbard (1990, 2005: 340), Oddie and Menzies (1992), Olsen (2017), Smith (2010)...

<sup>7</sup> For a recent example, see Schroeder (2018).

to believe. Perhaps it is objectively best to believe a proposition  $p$  just in case  $p$  is true. But consider a simple case like the following:

**Coin toss.** All you know is that a fair coin was just tossed in the next room. In fact, the coin landed heads. Ought you believe it landed heads, believe it landed tails, or suspend judgment about the matter?

Gibbard (2005), for instance, thinks that in the objective sense, you ought to believe the coin landed heads, since that is true, but what doxastic states you ought subjectively to have depends on your evidence. Some talk about objective and subjective senses of ‘justification’ instead.<sup>8</sup>

My concern below will be what the right account is of the more “subjective” kinds evaluations we seem to need in cases like *Miners*, and of why we need them in the first place. I will talk of *norms*, but for now leave it open whether these are normative in some more heavyweight sense, or whether they are better thought of as standards of evaluation.<sup>9</sup> While I have used *Miners* as a warmup case, I want to distance myself at the outset from the project of providing a truth-conditional semantics for natural language ‘ought’ -statements.<sup>10</sup> As will become clear below, the kind of puzzle raised by cases such as *Miners* that I want to focus on is a symptom of the following general phenomenon: for just about *any* norm, there are cases in which subjects who conform to it are negatively evaluable, and subjects who violate it are positively evaluable.

For now let me just point out that the need for a more “subjective” kind of evaluation does not essentially rest on consequentialist ways of thinking.<sup>11</sup> Take the deontological norm *Keep your promises!*. Assume that I promised my friend to return her book today. I go to my bookshelf, and exactly where I remember placing the book, I seem to see it. I pull it out, and return it to my friend. Unfortunately, a prankster put another book inside the hard cover of the book I had promised to return. I do not keep my promise, for I did not succeed in returning my friend’s book. Nevertheless, I seem to be positively evaluable. Similarly, knowledge first epistemologists will claim that in *Coin toss* you shouldn’t believe the coin landed heads simply because you are in no position to know this! Still, there will be cases in which we want to positively evaluate, even praise, beliefs that fail to constitute knowledge. Consider a subject with massively misleading evidence for  $p$ , who fails to know  $p$  because  $p$  is false. Or, consider a subject in a Gettier-case who fails to know  $p$  simply because her belief is susceptible to a problematic kind of epistemic luck.<sup>12</sup>

For now I will set aside norms governing belief, and focus on cases like *Miners*. I will begin by outlining a prominent diagnosis of just why we need more subjective norms, and an accompanying

---

<sup>8</sup> Feldman (1988) opens his paper with the sentence “A view widely held view by epistemologists is that there is a distinction between subjective and objective epistemic justification, analogous to the commonly drawn distinction subjective and objective justification (or rightness or obligation) in ethics”.

<sup>9</sup> Ultimately, I don’t claim to fully understand, nor subscribe to, the distinction often drawn between the normative and evaluative (e.g. Kolodny 2005). The project making sense of that distinction, as well as saying exactly what the status of my dispositional evaluations is, is a project for another occasion (and one that I take up in a book I am writing).

<sup>10</sup> Indeed, the case may be seen as raising several puzzles, one of which has to do with the semantics of ‘ought’ -statements. Note also that not only do I not undertake the project here, but I don’t subscribe to the project of understanding true ought-statements in terms of the notion of a reason. Consider cases in which an agent assessing your situation has more information than you, even though her information is still incomplete. In such situations, she might utter ‘you ought to  $\phi$ ’, for some act of  $\phi$ ’ing. Such seemingly true ought-statements are a problem for many views that endorse a distinction between subjective and objective senses of ‘ought’, for it looks like such true ought-statements cannot be understood in either of these ways. For instance, according to a dominant view, objective oughts depend on which action (choice) is favoured by the balance of *objective* reasons, and subjective oughts depend on which action (choice) is favoured by the balance of *subjective* reasons. For a discussion of this problem, see Kolodny and MacFarlane (2010) and Schroeder’s (2018) discussion of what he calls the *four-envelope* problem.

<sup>11</sup> Cf. Smith (2010).

<sup>12</sup> I have also argued (Lasonen-Aarnio 2010, *forthcoming* B) that there is such a thing as “unreasonable knowledge”: there are also cases in which a subject is negatively evaluable, perhaps even to be blamed for believing a proposition  $p$ , even if she knows  $p$ .

view of what those norms should look like, which I call **perspectivist**. According to this diagnosis, the problem with objective oughts, and accompanying norms like *Do what is best!*, essentially has to do with *epistemic access*: given your limited perspective on the world, you lack access to facts about exactly where the miners are. As will become clear, I don't think perspectivist views solve the problem, and I don't think that the access diagnosis quite hits the nail on the head. Indeed, the purpose of this paper will be to sketch an alternative diagnosis of why seem to need more "subjective" evaluations, and to put forth an alternative way of thinking about more subjective kinds of evaluations.

The problem with more objective norms, it is often argued, is that they make reference to facts that may not, in Jackson's (1991: 467) words, be "present to the agent's mind". I take a natural understanding of what it would take for a fact to be present to an agent's mind to be epistemic: the agent has some sort of epistemic access to the fact. According to this diagnosis, the heart of the problem is that we often lack access to the facts that the recommendations of more objective norms depend on.<sup>13</sup> (Indeed, sometimes more subjective senses of 'ought' are referred to as *information-relative* ones.<sup>14</sup>) For instance, just which action is best in *Miners* depends on exactly which shaft the miners are trapped in, and the subject in the case lacks access to this fact, for she doesn't know whether they are in shaft 1 or shaft 2. Here the notion of *guidance* is often brought in: objective norms don't provide us with adequate guidance about what to do.<sup>15</sup> And the reason for this is precisely that we often lack epistemic access to whether their application conditions obtain.<sup>16</sup>

Because I want to remain neutral on exactly what sort of epistemic access is required, I will characterize a very broad set of views – indeed, what I take to be a broad paradigm for thinking about more "subjective" kinds of evaluations – in terms of the notion of a *perspective*. An agent's perspective is whatever is, in the required sense, present to the agent's mind. The problem with more objective norms, the thought goes, is that they make reference to facts that may lie outside a subject's perspective. For instance, the miners are in fact all in shaft 1, but that fact is not represented by your perspective. Given this diagnosis, the seemingly inevitable remedy is to find norms that make recommendations that are functions of one's perspective.

According to many views, a perspective is a set of propositions. It may consist of the totality of propositions constituting a subject's *evidence*.<sup>17</sup> Or, it may consist of the totality of subjective *reasons* one has.<sup>18</sup> Or – and such a view is not incompatible with either of the previous ones – it may consist of the totality of propositions a subject *knows*, or the totality of propositions she is in a *position* to know. (Indeed, it is difficult to find a paper on the subject which doesn't point out that in *Miners* the subject doesn't *know* which shaft the miners are in.) Or, perhaps a subject's perspective

---

<sup>13</sup> According to Gibbard (1990: 343), "The basic normative precepts that ground a subjective *ought* are subjectively applicable – applicable in light of information the agent has", or in light of what she knows.

<sup>14</sup> E.g. Wedgwood (2016).

<sup>15</sup> See e.g. Gibbard (1990: 344), Jackson (1991: 466-7), and Sepielli (2012).

<sup>16</sup> The idea of guidance, and associating it with a kind of epistemic access, has a long history. As one fairly recent example, Smith (2010: 84) characterizes one of the attractions of the notion of subjective rightness as follows: "it could be used to identify a type of duty to which the agent has infallible access in his decision-making". For a discussion of guidance and epistemic access, see Hughes (?). However, we should be careful not to equate the need for more subjective norms with the need for norms that we can be guided by. Assume that I rationally but falsely believe all the miners to be in shaft 2: tragically, my evidence regarding the matter is massively misleading. On some views, in this situation I *can* be guided by the norm *Do what is best!* even if, in choosing to completely block 2, I fail to conform to it (cf Sepielli 2012 on confidence, and Smith's (2012) distinction between a narrow and a broad sense of "having the ability to use a principle to guide conduct"). In such a case, we still need some way of making sense of the fact that one is to be positively evaluated, despite failing to do what is best.

<sup>17</sup> For such evidentialist views, see Conee and Feldman ()

<sup>18</sup> One prominent way of understanding the distinction between subjective and objective oughts is in terms of subjective and objective reasons (e.g. Schroeder 2018). See Lord () and Kiesewetter () for views of subjective reasons formulated either in terms of knowledge, or in terms of what a subject is in a position to know.

consists of the totality of propositions she justifiably believes, or the totality of propositions she has justification to believe. On many views, only propositions believed by a subject can be part of her perspective. Many would agree, however, that thinking about a subject's perspective wholly in terms of her beliefs results in norms that are *too* subjective. Consider a subject in a case like *Miners* who happens to believe, as a result of some irrational process not in any way sensitive to where the miners in fact are, that all of the miners are in shaft I. Do we really want to say that such a subject ought to block shaft I?<sup>19</sup>

On some views, a subject's perspective consist of a subset of her mental states.<sup>20</sup> For instance, some think that there is a *sui generis* kind of mental state called a *seeming*. According to a phenomenal conservative variant of perspectivism, perspectives consist of the totality of one's seeming-states.<sup>21</sup> Or, it might consist of the totality of a subject's mental states that are "phenomenally conscious" to her at a time.<sup>22</sup> But even views that think of perspectives more propositionally typically take them to supervene on (or to be fixed by in some other way by) a certain class of a subject's mental states. For instance, according to many internalist views, one's perspective (e.g. one's evidence) supervenes on one's non-factive mental states.<sup>23</sup> Externalists might think that one's perspective supervenes instead on the mental state of knowing, or on mental states that include factive perceptual states.<sup>24</sup>

We don't always represent a candidate fact as either obtaining or not obtaining. One might think that we need a view of perspectives that can reflect our degrees of uncertainty. For instance, there is a difference between failing to represent a fact about the weather in London right now, simply lacking any opinion about the matter, thinking it 50% likely to be raining, and thinking it 90% likely to be raining. One might think that perspectives either are, or at least give rise to, probability functions. On one possible view, a subject's perspective is her subjective credence function, though again, such a view appears too subjective (consider a subject who is irrationally certain that the miners are all in shaft I).<sup>25</sup> Perhaps we should look to credences one *ought* to have<sup>26</sup>, or the credences a reasonable subject in one's position would have. Or perhaps, instead, we should look at evidential probability functions, which the totality of one's evidence gives rise to: one's evidential probability function at a time *t* is a prior probability function conditionalized on the totality of one's evidence at *t*.<sup>27</sup>

Though there are significant differences between different conceptions of perspectives, they all have important features in common. Perspectives are closely connected with some class of mental states: indeed, it is precisely because some class of one's mental states gives rise to one's perspective that perspectives capture what is present to an agent's mind. Perspectives are *representational*. According to many views, a perspective is a set of true propositions corresponding to a subset of the facts – namely, those facts that a subject has access to (e.g. those constituting her evidence, her reasons, or her knowledge). And mental states like *seemings* have propositional, representational contents. Further, even a probability function can be seen as a degreed representation of the world. For instance, it is precisely because it is not likely enough on the relevant probability function in

---

<sup>19</sup> Cf. Gibbard (2005: 346), who gives an example of someone who has a true but unjustified belief about the exact number of eggs in Ann Arbor.

<sup>20</sup> Pollock & Cruz () think some (foundational) reasons are mental states, while others are propositional.

<sup>21</sup> Numerous questions arise, of course: the totality of one's seemings at a time *t*, or in one's history, or of a select subset of the totality of seemings in one's history?

<sup>22</sup> Cf. Sepielli, who (2014: 524) also explicitly talks of *perspectives*.

<sup>23</sup> E.g. Conee and Feldman

<sup>24</sup> Williamson (2000), Pritchard (?)

<sup>25</sup> E.g. Jackson (1991).

<sup>26</sup> Parfit (1988) characterizes subjective rightness in terms of what one has reason to *believe*, and Gibbard (2005: 346) in terms of what one ought to believe.

<sup>27</sup> For such a view of evidential probabilities, see Williamson (2000).

*Miners* that the miners are all in shaft 1 that the correct subjective norm doesn't tell one to block shaft 1.

However one thinks of perspectives, they can be **incomplete**: not all facts are represented by one's perspective. And however one thinks of perspectives, they can be **misleading**. Even if one's perspective consists exclusively of true propositions, true propositions can support (e.g. by making likely) falsehoods. For instance, if many normally reliable testifiers I have no reason to distrust tell me that the miners are all in shaft 2, then true propositions regarding what these testifiers tell me can make likely the falsehood that the miners are all in shaft 2.

Just about all views that endorse subjective oughts in response to cases like *Miners* appeal to perspectivist norms. According to a prevalent way of thinking, what we subjectively ought to do is maximize some quantity, where that quantity is a function of

- (i) some probability function, and
- (ii) a value function assigning objective values to the available options.

According to a prominent class of views, one subjectively ought to choose the action that maximizes *expected value* (or an action among those with highest expected value) by the lights of the relevant probability function.<sup>28</sup> For instance, according to Jackson's (1991) consequentialism, one ought to *maximize expected moral utility*. According to Parfit (1988), the subjectively right act is one that *maximizes expected goodness*. What makes the subjective ought subjective is that the relevant probability function is fixed by (or is identical to) the subject's perspective. According to a simple view, it is just a subjective credence function. But as we have seen, subjective credences seem too subjective. Similarly, when what is at issue is a moral ought, the relevant value function cannot just be a matter of subjective utility, derived from a subject's actual preferences: what if the subject prefers to kill as many people as possible? We need a more objective value function. Hence, it should be clear that though this view of subjective oughts has a decision-theoretic structure, what is at issue is not just maximization of subjective utility.

This is, of course, not the only possible view of subjective oughts, but it is the most popular.<sup>29</sup> I will try to avoid in-house debates as far as possible. Instead, I will sketch what I take to be fundamental problems for any perspectivist norms. While I will focus on the norm *maximize expected value*, it will hopefully be clear that the problems raised are not specific to this norm.

My aim in this paper will be to outline an alternative diagnosis of why, it would seem, you ought not to block shaft 1 in *Miners* – or, setting oughts aside, why you would at least be negatively evaluable for doing so, and positively evaluable for letting both shafts flood halfway. In a nutshell, whether one is to be positively evaluated from the dispositional perspective for doing something (i.e. choosing a given course of action, forming a given belief, etc.) is a matter of whether one is manifesting *good dispositions*. There are several approaches to determining whether a disposition is good. According to the approach I favour, a disposition is good just in case it is at least as good as feasible alternatives. The norms generated by the dispositional perspective are *manifest (one of) the best feasible dispositions*, and *do what would manifest (one of the) the best feasible dispositions*. Epistemic access and perspectives are closely connected with good dispositions: often the best feasible dispositions are sensitive to our perspectives (e.g. our evidence). My main complaint with the access diagnosis is that

---

<sup>28</sup> There are a range of debates here. What exactly is the quantity to be maximized? Such debates connect with what the right decision theory is: evidential or causal decision theory; risk-weighted value, etc.? I won't take a stance on these in-house debates.

<sup>29</sup> Alternatively, one might propose, for instance, that one ought to do what is subjectively most likely to maximize actual value. Such a proposal has come under many attacks (e.g. Parfit 1988, Jackson 1991). For a good overview and criticism of various proposals, see Smith (2010).

it does not, as I argue below, get to the very heart of the matter, for there are cases in which access and good dispositions come apart.

One important feature of my view is that it unifies the practical and theoretical/epistemic domains. The main contenders for objective epistemic value are *truth* and *knowledge*. One believes as one objectively ought just in case one's belief is true, or constitutes knowledge. The dispositional evaluative perspective can be straightforwardly applied to evaluate beliefs: my belief is to be positively evaluated just in case it is the manifestation of the best feasible dispositions. The view also straightforwardly applies to doxastic states other than belief. Just like believing, suspending judgment, for instance, can be a manifestation of the best feasible dispositions.

In what follows, I will first outline the dispositional evaluative perspective (§2). I will then argue that perspectivist norms do not solve the problem, for perspectivist norms are susceptible to miner-type cases (§3). Finally (§4), I will offer an alternative, dispositional diagnosis of the original problem with more "objective" norms like *Do what is best!*, and argue that unlike perspectivist norms, my preferred dispositional norm is not susceptible to the same problem.

## 2. Dispositional evaluations

### Ways

We need to evaluate doxastic states, choices, and actions, in a manner that is sensitive to the way in which those doxastic states are formed and retained, the choices are made, and the actions performed. The epistemology literature is replete with proposals for how to understand evaluations sensitive to, in particular, ways of forming beliefs: a belief is formed in a good way if it is properly based on sufficiently good, undefeated *reasons* or *evidence*, if it is the output of a *reliable process*, or if it is formed by a *reliable method*. In this connection epistemologists often talk of *doxastically justified* beliefs. But we need to be able to make similar kinds of evaluations in the practical realm. For instance, an agent might make what is in some sense the right choice, but make it in a bad way – she might do the right thing "for the wrong reasons".<sup>30</sup>

My proposal is to let the dispositions that manifest themselves as one's  $\phi$ 'ing (coming to believe something, retaining a belief, making a choice) to do the work of identifying these ways.<sup>31</sup> The goodness of a disposition depends on how successful its manifestations are across a range of relevant counterfactual cases – or, to be more precise, on how successful its manifestations are compared with the manifestations of feasible alternative dispositions.

Since my aim here is to paint a big picture, I won't take on questions regarding the metaphysics of dispositions. I will simply assume – like I think everyone should – that there are dispositions of a wide variety of different kinds, and that inanimate objects and rational subjects alike are constantly manifesting them.<sup>32</sup> I will assume that we can find natural dispositions at every level of reality, not just the microphysical level. For instance, the kinds of dispositions of interest when evaluating beliefs and choices will be at the psychological level, broadly understood. My starting point will be that when an action, doxastic transition, or doxastic state can be in some sense attributed to an agent, it is the manifestation of some of her dispositions. Sometimes things more or less happen to us: if, for instance, my mind comes to be controlled by some external force causing me to be in a neural state that makes for believing  $p$ , then, believing  $p$  is not something properly

---

<sup>30</sup> *Kant's shopkeeper example etc.*

<sup>31</sup> In principle, the theoretical role played by dispositions in my account could be played by something else, such as *methods*, *rules* or *strategies*. We would then evaluate a belief (choice, action) by asking whether beliefs formed by that method or rule are successful across relevant, somewhat normal counterfactual cases. Thanks to Bart Streumer for discussion.

<sup>32</sup> A distinction going back to Aristotle is sometimes drawn between dispositions like solubility and rational powers, but my notion of a disposition is broad enough to encompass both: indeed, the dispositions manifesting as choices often involve exercising our rational capacities. For all I say here, some of the relevant dispositions might involve responsiveness to reasons.

attributable to me. In this case, we are not even in the sphere of the kind of evaluation I want to outline. But when things don't merely happen to us, we are manifesting some dispositions.<sup>33</sup>

### *The framework*

Dispositional evaluations depend on what we take the relevant success to be. Believing that one's lottery ticket will lose just based on the odds may be the manifestation of a good disposition if epistemic success is a matter of believing truly, but not if it is a matter of knowing. Moral choices will be evaluated very differently depending on whether acting successfully is a matter of doing things with good consequences, of conforming to deontological rules, or of something else entirely. For now I will set aside these questions, and present the basic dispositional evaluative framework. It should be clear, however, that it can be implemented in very different ways.

Assume that a subject  $\phi$ 's (comes to believe a proposition  $p$ , chooses to block shaft A, etc.) in the actual case, and that her  $\phi$ 'ing is the manifestation of some disposition D.<sup>34</sup> What is it for D to be good in the relevant sense? We can split how the goodness of a disposition is determined into three steps.

- (1) Take the total set **S** of counterfactual cases in which the disposition manifests itself. **Weight** these according to *relevance*.
- (2) Determine an overall **score** of the disposition across these counterfactual cases. The score of a disposition is a function of two things: a **value** function assigning values to the manifestations of the relevant dispositions in cases within S, and the *weighting* of the cases according to relevance.
- (3) Say how the overall relevant kind of **goodness** of a disposition is related to its score.

Let me discuss these in turn.

**(1)** I think that dispositional evaluations are highly malleable and context-dependent: what counts as relevant can easily shift depending on our focus and context as evaluators. Indeed, what will be at issue is a contextually determined metric of relevance. Nevertheless, there are general structural points to be made.

The first thing to emphasize is that relevance should not be understood in terms of any relation such that a case trivially bears that relation to itself (to a maximal degree). Relevance is not a matter of *relevant similarity*: counterfactual cases in **S** are not weighted according to how similar they are to the actual case. Similarly, relevance is not a matter of what could easily have occurred, for what actually occurs could trivially have easily occurred. This is because the case a subject is actually in might be deviant or *abnormal* and hence, it – and cases very much like it – may be irrelevant when evaluating the goodness of dispositions. I won't here offer a theory of what such normality consists in, but even lacking a precise account, we have some initial grasp on what sorts of cases count as normal. Assume, for instance, that I am wandering around in a foreign city, and have a perceptual experience as of its being sunny outside as a result of looking into a covered courtyard creating the illusion of the outdoors, complete with fake sunshine. There is something abnormal about experiencing such an intricately crafted illusion. Even if that is the case I am actually in, it might

---

<sup>33</sup> This should not be conflated with the false claim that actions attributable to an agent always arise out of *habit*, or that we only ever do things we are generally disposed to do.

<sup>34</sup> For the purposes of stating my view, I will assume for simplicity that there is just one disposition. It might, however, be more realistic to assume that one's  $\phi$ 'ing is the joint manifestation of multiple dispositions. In that case, at step (1) we should consider the total set of cases in which *all* of these dispositions jointly manifest themselves.

be assigned very little weight for the purposes of dispositional evaluations: the dispositions I manifest in coming to believe that it is sunny are still good, for they tend to manifest as successful belief in somewhat normal cases.

Here is a toy model that at least comes close to how I think about relevance. As evaluators, we consider some features of a case to be idiosyncratic, while holding others fixed. Take, for instance, a given shot fired in a Saturday football game of 6-year olds. We consider various features or parameters of the very case in which the shot is made. Some we hold fixed: the operation of usual forces like gravity, the absence of hurricanes or snow blizzards, the skill level of the players, how fast they run, etc. Given the context of evaluation, we might only be interested in goodness in games of similarly skilled 6-year olds: it doesn't matter that the dispositions manifested would be highly unlikely to manifest as a successful shot in a game in which the goalie was a world cup player. Other features we consider idiosyncracies of the particular case at hand: the configuration in which audience members are seated, the current facial expressions of the various players, random gusts of wind, the colour of the goalie's socks, and their uninformed guess as to which corner the player will aim for.

We consider, then, counterfactual cases (within **S**) in which those features we are holding fixed obtain – we could think of these features as fixing a contextually determined *type* of case – while letting other features vary. For instance, the relevant type might be a Saturday football game of a certain team of 6-year olds in somewhat normal summer weather; or it might be a worldcup game against a given team. The counterfactual cases are then weighted according to how typical or normal they are qua instances of that type: we assign more weight to cases the more *normal* or *typical* they are qua instances of the relevant type.<sup>35</sup> The relevance weighting might be binary: either a case instantiates the right type and is normal, or it is not. Or relevance could come in degrees, cases instantiating the relevant type being assigned more weight the more relevant they are.

(2) The score of a disposition, representing how well the disposition does across counterfactual cases in which it manifests itself, is fixed by the weighting of these cases according to relevance, and a *value* function. The value function assigns values (positive or negative) to manifestations of the disposition in counterfactual cases. The nature of the value function will, of course, depend on what the relevant success under consideration is. In some cases all that matters is whether a manifestation is a success or failure – coming close to succeeding doesn't get assigned any extra value (e.g. penalty shots in football). In other cases success itself comes in degree. When evaluative dispositions in cases like *Miners*, in the absence of further information, it is natural to think that one's choices are to be assigned values in counterfactual cases as a function of the number of lives saved.

I will assume that the score of a disposition is a weighted sum of the values of its manifestations in cases within **S**, the weightings being by relevance. (I think this is a natural view, though it is not the only possible one.)

(3) There are various possible approaches to determining the overall goodness or “success-conduciveness” of a disposition. According to the view that I think is the most promising way of capturing the kinds of more subjective evaluations we need in cases like *Miners*, dispositional goodness is a matter of manifesting the **best feasible** disposition (or one of the best feasible dispositions, if there is a tie), or at least coming sufficiently close. If  $D_1$  and  $D_2$  are both feasible alternatives in one's situation, then  $D_1$  is better  $D_2$  than just in case it has a higher score. According to the resulting view, one is to be positively evaluated just in case one manifests (one of) the best

---

<sup>35</sup> In my treatment of the so-called new evil demon problem, I have argued that what counts as normal often depends on the kind of world or case that a subject is *anchored* to – she might in fact be in a case that is abnormal for her (see Lasonen-Aarnio *forthcoming A*). Hence, “alien cognizers” living in environments that are abnormal by our lights can be evaluated positively from the dispositional perspective, for their dispositions fare well in counterfactual cases that are not too abnormal for *them*.



feasible disposition(s), or does (chooses, believes, etc.) whatever (one of) the best feasible disposition(s) would manifest itself as doing in one's situation. The resulting norms are *manifest (one of) the best feasible dispositions!*, and *do what would manifest (one of) the best feasible dispositions!*

Not any way of being disposed is a way we could be disposed. There are various constraints, of different strengths, on what dispositions we can have and manifest. Some are logical, some nomological, some metaphysical, some flow from broad features of our cognitive design. Though there is nothing metaphysically impossible about such a disposition, given the way we are, we cannot be disposed to only use safety belts on car rides that end in crashes, or to buy insurance only for trips that end up involving some sort of calamity. Climbers cannot be disposed to only check their knots when they have mis-tied them. Or, consider the fact that our total evidence is sometimes misleading: we sometimes have overwhelming evidence in favour of a proposition *p*, even though *p* is false. We couldn't be disposed to only take into account total evidence making a proposition very likely when that evidence is not misleading. In general, we cannot be disposed to be sensitive to our perspectives only when those perspectives are neither incomplete nor misleading.

Given the context-sensitivity of dispositional evaluations, feasibility in different contexts might involve different senses of "could". As a general rule, feasibility is constrained by our broad cognitive architecture, allowing for some idealizing. Consider the literature on heuristics and biases – for instance, the systematic patterns of fallacious probabilistic reasoning exhibited by humans. Psychologists tend to think that such errors flow from features of our cognitive design, which consists of different systems tacked together (e.g. what is known as System 1 and System 2). If a subject commits the so-called bankteller fallacy, thinking a conjunction to be likelier than its individual conjunct, then even if she manifests some dispositions typical of human beings, she is not manifesting good dispositions: there is a better available disposition – indeed, we are sometimes able to pick up on such errors ourselves.

Any *feasible alternative* disposition is one that would manifest itself as a relevant doxastic state, choice, or action in one's situation. For instance, relevant alternative dispositions in *Miners* must manifest in the situation at hand as making a choice among the available options.<sup>36</sup> (It is feasible for the subject in *Miners* to manifest a disposition to scratch her head, but this disposition is, of course, not an *alternative* in the relevant sense.) Further, these dispositions must be feasible: in a contextually relevant sense of "could", they are dispositions that could be manifested in one's situation.

As I said, I think the most promising approach to capturing the more subjective kinds of evaluations that cases like *Miners* call for is in terms of the *best feasible dispositions*. This is not to say, however, that it is irrelevant how good a disposition is in absolute terms. In some unfortunate circumstances, even the best feasible dispositions are connected to success only in a way that is not robust. For instance, if I face a rescue operation in a storm at sea, the very best thing to do might be to randomly toss a life raft out of the ship, hoping for the best. Similarly, a doctor in war zone who has access to nothing but one kind of antibiotic might be manifesting the best feasible disposition when she gives this antibiotic to a patient with a fever and undiagnosed mixed symptoms. The dispositional perspective correctly evaluates these actions positively. However, we may also care about manifesting dispositions that are connected with success in a *non-accidental* way, dispositions that robustly manifest as successes across a wide class of cases involving variations of the idiosyncratic-seeming features of the actual case. Dispositions that thus manifest as successes in an invariant way typically *explain* the success to a high degree. And whether one's dispositions explain one's success might, for instance, be relevant for whether, and the degree to which, a subject is to

---

<sup>36</sup> Assume that the subject in fact manifests disposition *D*<sub>1</sub>. Then, in the actual *case* she manifests *D*<sub>1</sub>, and no other disposition. Nevertheless, we can ask whether some other disposition that *could* have been manifested is better. This is why I talk about *situations*: when asking what the best feasible dispositions are in one's situation, we keep fixed other features about the actual case, but not facts about what dispositions (if any) one manifests. My account also assumes counterfactual facts regarding what choice (action, doxastic state, etc.) a given disposition *would* have manifested as in one's situation.

be *praised for succeeding*.<sup>37</sup> Indeed, it seems plausible to me that the more robust the connection between the disposition manifested by an agent and success, the more attributable the success is to the agent.<sup>38</sup> And the more attributable the success is to the agent, the more praise she deserves for it. Positive dispositional evaluations come in many flavors: not all cases of manifesting the best feasible dispositions are on a par.

On my view, we often deploy something very much like the dispositional evaluative perspective.<sup>39</sup> It is important to see that making dispositional evaluations does not require having access to the precise dispositions manifested by a subject, for considerations of feasibility provide a significant shortcut. We often have a good idea of what kinds of dispositions an agent could feasibly manifest in a given situation. If there is no good feasible disposition that would manifest itself as choosing (believing, acting) as the subject does, this is all we need to know in order to correctly negatively evaluate her choice (belief, action) from the dispositional perspective.

### *Shortcuts and Miners*

Consider a human subject in *Miners* who chooses to block shaft 1. This is objectively the best course of action, for all of the miners are saved. But could such a choice be a manifestation of a good disposition? We know that the subject has limited epistemic access to just where the miners are – indeed, she doesn't know where they are, and her evidence favours neither the hypothesis that they are in shaft 1 nor the hypothesis that they are in shaft 2. It is not feasible, in the situation described, to manifest a disposition that is sensitive to the exact location of the miners. Whatever the actual disposition manifested is, it cannot be one that *discriminates* between cases in which the miners are in one shaft and cases in which they are in the other shaft – or, more generally, that discriminates between cases depending on what the objectively best thing to do is. Whatever disposition an actual human subject might be manifesting by choosing to block shaft 1, then, would indiscriminately manifest itself across a range of counterfactual cases as tragically choosing to block the shaft the miners are all in. That is why, if one manages to block the shaft the miners are in, one's success seems accidental, merely lucky.

Hence, the problem with choosing to block shaft 1 is that any feasible disposition manifesting itself as blocking only one of the shafts manifests itself across a wide range of relevant counterfactual cases as blocking the shaft containing all 10 miners. There is, after all, presumably nothing abnormal about counterfactual cases in which the miners are stuck in one shaft rather than the other. And 9 lives being saved has significantly more objective value than all 10 lives being lost. A disposition that uniformly results in 9 lives being saved is seems better than one that haphazardly results in all 10 miners being saved in some cases, and all 10 miners dying in others.

At least very often, our best dispositions to act and make choices are sensitive to how things are in the world by being sensitive to our mental states, which in turn reflect our evidence. Indeed, the best feasible dispositions are often sensitive to our *perspectives*: to how we, or our evidence, represents the world as being. Indeed, for all I have said, it may even be that the best feasible

---

<sup>37</sup> We should be sensitive to the difference between praising a subject for doing something (e.g. making a given medical decision), and praising her for succeeding (e.g. curing a patient). See Johnson King (*forthcoming*).

<sup>38</sup> One might here be reminded by an idea put forth by many virtue epistemologists, which is that a success is creditable to a subject when her competence explains her success (REFERENCES Sosa, Greco...). My claim here is in the same ballpark, though manifesting the best feasible dispositions in one's situation does not require anything like a full-blown virtue or competence. Also, I reject, of course, the idea that *knowledge* requires manifesting good dispositions (see Lasonen-Aarnio 2010, *forthcoming B*).

<sup>39</sup> I have argued that it explains a large class of "data" that has been used to support internalism in epistemology, like intuitions about victims of deceit. For now my view solves the New Evil Demon Problem, see Lasonen-Aarnio (*forthcoming A*). For defeat, see Lasonen-Aarnio (2010) and *forthcoming B*.

dispositions that subjects in cases like *Miners* could manifest are sensitive to something like expected value – indeed, they might be dispositions to maximize expected value.

The dispositional view often mimicks perspectivism because the best feasible dispositions are often sensitive to our perspectives in specific ways. But the underlying normative stories are very different. Assume that I acquire evidence making a proposition  $p$  overwhelmingly likely (though not certain), and I come to believe  $p$  on the basis of that evidence. Why is my belief to be positively evaluated? According to one perspectivist, evidentialist norm, it is because it is a general normative fact that I ought to only form beliefs in propositions that are likely on my evidence. Here is an alternative, dispositional story: in coming to believe  $p$  on the basis of the evidence making  $p$  likely, I am manifesting the best feasible disposition in my current situation.

So far it might seem that my alternative, non-perspectival, dispositional view rides on the achievements of the best-developed perspectivist views. Often, I have conceded, the best feasible dispositions are sensitivity to one's perspective. So I now want to turn to problems faced by the perspectivist approach. The core problem with perspectivism is simple: it doesn't solve our original problem, for miner-type cases keep cropping up for perspectivist norms.<sup>40</sup>

### 3. Problems for Perspectivism

In the beginning I outlined a broad framework for thinking about subjective oughts – and, as a result, of the more “subjective” kinds of evaluative standards cases like *Miners* call for: what a subject subjectively ought to do depends on her *perspective*, where a perspective is a kind of representation of the world. Perspectives capture how things are by the lights of particular subjects. Unfortunately, perspectivism cannot escape the challenge raised by cases such as *Miners*. I think we need to shift our diagnosis of the original problem, and the framework used for solving it.

A perspective is part of the world. Just as there are facts regarding where miners are located, about what the effects of administering a given medication are, and about what the weather in London is like, there are facts about what one's perspective is in the first place, and what one ought to do by the lights of one's perspective. For instance, there are facts about what we believe, what we know, and what evidence we have. And one's perspective can take a perspective on such facts, one that can be incomplete or even misleading. One can lack epistemic access to these facts, just as one can lack epistemic access to facts about which shaft the miners are in.<sup>41</sup>

To make the problem vivid, consider the following case, in which one's access to mathematical facts about expectations is incomplete:

#### *Another Mining Disaster*

You often find yourself in situations involving mining disasters. To prepare, you spend your evenings analyzing particular scenarios, and calculating the expected values of various actions. You now find out there has been another accident. Luckily, just last night you calculated the expected values of the available actions in the very situation you now face. But alas, you have forgotten the exact results of those calculations! There is no time for calculations – if you don't act quickly, all miners will die with certainty. You face a choice between actions A, B, and C. You remember the following facts about your calculations from last night: one of A or B has the highest expected value (10), while the other

---

<sup>40</sup> I by no means think this is the *only* problem. Another issue is the possibility of good dispositions that simply bypass anything like one's perspective on the world...

<sup>41</sup> One might also lack access to what choices or actions are genuinely available to one. One might lack access to facts about values, and to facts about what the correct normative theories are in the first place! The discussion below is closely related to discussions of normative uncertainty. For the purposes of this essay, I am setting the problem of normative uncertainty aside.

one has the lowest expected value (0). You know, and are certain, that C has quite a good expected value (9), though not as high as that of the best of the three options (A or B). What should you do?<sup>42</sup>

What we have, in effect, is a generalized version of the miners problem: intuitively, the agent is positively evaluable for choosing option C, even though she knows that doing so does not maximize expected value. Exactly the same considerations that motivate introducing a more subjective ought in the original *Miners* case appear to show that norms like “maximize expected value!” are not subjective enough. What you in fact ought to do by the lights of your perspective regarding the mining situation you face seems to come apart from what you ought to do by the lights of your perspective regarding your perspective regarding the situation you face.

The problem raised by *Another Mining Disaster* stems from the fact that calculating expected utilities is itself not a trivial matter; non-ideal subjects don't always have access to such mathematical facts. But this is but one instance of the more general problem. Consider, for instance, the fact that relations of evidential support are often highly non-trivial. We can construct structurally similar cases in which instead of lacking access to mathematical facts about expectations, you lack access to facts about evidential support. Assume that you went through some painstakingly complicated reasoning earlier on to figure out how likely a salient proposition  $p$  is on your evidence (it is in fact very likely). Unfortunately, you have forgotten the result of that reasoning, and there is no time for thinking things through again. All you remember is that  $p$  is either very likely or very unlikely. If it is very likely, A has the highest expected value (10), B the lowest (0), and C has an expected value that is close to but not as high as A (9). If  $p$  is unlikely, B has the highest expected value (10), A the lowest (0), and C again has a rather high, though not maximal, expected value (9). In such a case, many would say that you ought hedge your bets by choosing option C, even though you know that it doesn't maximize expected value.

One reply to such cases is to appeal to some notion of an idealized agent – agents who have unlimited computational powers and access to all necessary, *a priori* facts such as facts about evidential support – and to insist that the relevant theory of subjective oughts only applies to such agents. But a theory of a more subjective kind of evaluation that doesn't even apply to us non-ideal agents would be spectacularly disappointing. And even more importantly, I don't think idealizing in these ways solves the problem. For perspectives don't always accurately represent facts about themselves – that is, the perspectives that perspectives take on themselves can be incomplete and even misleading. Let a perspective be *transparent* to itself if it has full access to facts about what it consists in. In so far as a perspective is a set of propositions, transparency requires that it is always certain on one's perspective exactly what this set of propositions is. It requires both a claim of positive and negative access. If  $p$  is part of one's perspective, then it is part of one's perspective that  $p$  is part of one's perspective. If  $p$  is not part of one's perspective, then it is part of one's perspective that  $p$  is not part of one's perspective. Such access conditions are extremely strong. But even the positive and negative access conditions wouldn't be enough to make perspectives transparent to themselves. To see this, consider the following example.

A constellation of Greek Gods (and only Greek Gods) is trapped in an escape room together, and you are to figure out whether, given their known powers, they will be able to escape. You know that *most* arbitrary subsets of Greek Gods could escape, given their reasonable level of intelligence and resourcefulness. For each God trapped in the room, you know of that God that they are trapped there (cf. *positive access*). For each God who not trapped, you know that of that God that they are not trapped (cf. *negative access*). As a matter of fact, it is Dionysios and Ares who are trapped. So you know that Dionysus is there, and that Ares is there. And, for instance, you

---

<sup>42</sup> Adapted from Kagan (2018: 155). Spencer and Wells (*forthcoming*) discuss a case with a similar structure, *The Fire*. See also Smith's (1988: 98-99) much earlier discussion of what she calls the 'Problem of Doubt'. And there is, of course, a large literature on related problems having to do with normative uncertainty.

know that Hera, Aphrodite and Zeus are not. Doesn't this give you knowledge of exactly what Gods are in the room – don't you then know that Dionysus and Ares are in the room, *and that no-one else is*? It does not. For assume that your knowledge of Greek Gods is rusty: you do your best to write down a list of each and every God, but for all you know, you have omitted someone! As a result, even though you have gone through your list, crossing out or underlining names, you cannot be sure that the two remaining underlined names constitute a full list of the contents of the room. You know that Dionysus and Ares would not be able to escape: they would end up getting drunk and fighting. But given an addition of almost *anyone* else, they could probably sort things out. So unfortunately, you do not know whether the trapped Gods would be able to escape.

The point of the above example was to demonstrate just how difficult it can be to achieve full transparency, whether about the exact contents of a room, or about the exact contents of one's perspective (for instance, of one's evidence). Just as one's perspective can be incomplete when it comes to the world, one's perspective can be incomplete when it comes to one's perspective, by failing to correctly represent true facts about what is and what isn't part of one's perspective. Not only might it be incomplete, but it might even be misleading when it comes to such facts. It is blatantly clear that transparency fails on more externalist theories of perspectives – for instance, views on which one's perspective consist of all and only the propositions one knows. It is more or less uncontroversial that I can fail to know *p*, without being in any position to know this. But internalism as such provides no guarantee against transparency failures. The currently most popular versions of internalism are formulated in terms of the notion of a *non-factive mental state*. Knowing *p* is not a non-factive mental state, but believing *p* is. Our own beliefs, however, aren't transparent to us. Empirical evidence confirms that we are very far from even having positive access to our own beliefs.<sup>43</sup> The non-transparency of our beliefs creates trouble for any view on which a proposition *p* is part of one's perspective *only if* one believes *p*.

I am not the first person to discuss these kinds of problems.<sup>44</sup> But proposed solutions normally assume some form of perspectivism. I take perspectivists to have three main lines of response. The first is to accept what I have said, and to concede that it shows that we need a plethora of different subjective oughts of different orders, or of different kinds of subjective evaluations.<sup>45</sup> The second is to deny that Miner-type situations could arise for perspectivist norms. This strategy could take one of two forms. The first denies that perspectives can be incomplete or misleading when it comes to facts about perspectives, or facts about what particular perspectivist norms recommend in one's situation. The second denies the intuition regarding *Another Mining Disaster*, or discounts its significance. The third is to try to defend some theory of subjective oughts that takes into account what perspective a perspective takes on itself, on various mathematical facts, facts about evidential support, etc. According to such a view, for instance, it may be that in the original *Miners* case one subjectively ought to maximize expected value, but that in iterated cases such as *Another Mining Disaster*, one ought instead to maximize expected expected value. Such views might either appeal to a single rule that takes into account all of these facts about one's perspectives, or to some sort of hierarchy of rules.<sup>46</sup>

---

<sup>43</sup> REFERENCES Smith (2010) also discusses problems for theories of subjective oughts (or "subjective rightness") raised by the fact that we don't always have access to our beliefs.

<sup>44</sup> See Srinivasan's (2015) discussion of the normative consequences of what she calls Anti-Cartesianism. For discussions of problems in this ballpark for theories of subjective oughts in particular, see Smith (1988, 2010), Sepielli (2014), Kagan (2018), Spencer and Wells (*forthcoming*)...

<sup>45</sup> Though I haven't discussed normative uncertainty, Sepielli (2014) argues that there are incommensurable 'orders' of rationality that might disagree about what one ought to do when faced with different kinds of uncertainty.

<sup>46</sup> See Lasonen-Aarnio (2014), for a critical discussion of so-called Über-rules, Kagan (2018, footnote 5) for a brief discussion of a "multilayered subjectivised principle" and a worry that no such principle would solve the problem, Smith's (1988) critical discussion of "rules of thumb", and Spencer and Wells (*forthcoming*) and Spencer (*manuscript*) for an attempt to work out how such a view would work. Smith (2010) discusses problems raised by cases in which a subject lacks access to her own beliefs. She sees these as problematic because of reasons having to do with guidance. As a remedy, she (2010, 1988) suggests a

I want to motivate a completely different diagnosis of the original problem and correspondingly, a different account of the kinds of “subjective” evaluations needed to solve it. Because of this aim, I won’t here discuss views that aim to respond to at least some of the problems raised. I take what has been said above to be enough to show that perspectivist views of subjective oughts face a serious challenge. At the very least, it should be clear that things are getting very messy and complicated. And when things start getting very messy and complicated, I think it is good methodology to re-evaluate the framework we are operating in.

Here is where we are. In some situations conforming to more objective norms like *Do what is best!* is difficult in a specific kind of way. These are paradigmatically situations like *Miners* in which one has limited access to relevant facts about the world. In these cases the need for more “subjective” kinds of evaluations arises. According to the access diagnosis, the fundamental source of the problem is that we sometimes lack epistemic access to the very facts that the recommendations of norms like *Do what is best!* depend on. Given this very diagnosis, however, we should not expect perspectivist norms to solve the problem, for we sometimes have limited access to our perspectives. In fact, if the access diagnosis is right, and there are very few facts we have trivial epistemic access to, then the situation might look desperate: we should expect new versions of the problem to keep cropping up! But perhaps the situation is not that desperate. Let me now turn again to dispositionalism.

#### 4. Dispositionalism to the rescue

I will first formulate what I take to be the core reason why versions of the original puzzle keep cropping up for perspectivist norms. This will involve offering an alternative diagnosis of the original puzzle, one already mentioned above. I will then apply the dispositional view to cases like *Another Mining Disaster*. Finally, I discuss whether the norm *Manifest good dispositions!* is susceptible to the same kind of problem. I put forth tentative considerations for thinking that it isn’t: there are no cases in which a subject conforms to it, but is criticisable in the same way as a subject in our original case *Miners* would be criticisable were she to choose to completely block shaft I.

##### *A dispositional diagnosis*

Why are perspectivist norms susceptible to yet new iterations of miner-type cases?

Consider any norm of the form  $\varphi$  just in case C! Given a very wide range of candidate norms, C can obtain even if one’s perspective is limited, and even misleading, regarding facts about whether or not C obtains. This is clearly true of “objective” norms such as *Do what is best!*, but it is, as we have seen, also true of perspectivist norms such as *Maximizes expected value by the lights of your evidence!*. When one’s perspective has limited access to the application conditions of a norm, conforming to the norm often becomes more difficult in a specific kind of way. If I know where the miners are, it is no problem for me to do what is best – to block the shaft they are in – but not so if I have no knowledge of the matter. In the very same way, when I have limited access to facts about the ranking of actions with respect to expected value, or to facts about evidential support, following perspectivist norms becomes more difficult. And in those cases, we might want to revert to a further subjectivized norm to evaluate my choices.

When conforming to a norm is difficult in the relevant sense, I can only conform to it in an accidental, seemingly haphazard way. Here is a dispositional account of the nature of the difficulty, which draws on points made above. In some situations the connection between the dispositions I manifest and my conformity to a norm is highly dependent on idiosyncratic features of my situation. Were we to vary these features, I would no longer conform to the norm: my conformity is not

---

hierarchy of principles, and the idea that what is subjectively right is given by the most highly ranked principle a subject is able to use.

robust across a relevant range of counterfactual cases. As a result, my success of conforming to the norm seems highly accidental. For instance, just which action would conform to the norm *Do what is best!* in *Miners* depends on facts about exactly where the miners are. Conforming to this norm in a robust way across a range of counterfactual cases would require manifesting dispositions that discriminate between cases depending just on facts about where the miners are in. But given your uncertainty about which shaft the miners are in, it is not feasible to manifest a disposition that is thus sensitive to these facts. If you choose to block the shaft the miners are in, it seems that you only *happened* to choose what is objectively best. While we can in turn explain the fact that in *Miners* one cannot manifest dispositions that robustly manifest themselves as choosing the best course of action by appeal to one's limited epistemic access, the source of the problem *ultimately* has to do with feasible dispositions, not epistemic access.

For pretty much any norm, there will be cases where one conforms to the norm, but one's conformity to the norm seems accidental. Perspectivist norms are no panacea, which is why appealing to them does not solve the problem raised by cases such as *Miners*. It is plausible that in *Miners* a subject can conform to the rule *Maximize expected value!* in a kind of non-accidental way: she can manifest dispositions that are suitably sensitive to facts about just which action maximizes expected value across a range of relevant cases. Nevertheless, in *Another Mining Disaster*, it is not feasible for you to manifest such dispositions. It is part of the description of the case that the relevant mathematical facts are complicated, requiring you to make non-trivial calculations, the results of which you have forgotten. Given these facts, whatever dispositions you manifest, it is plausible to assume that they cannot be sensitive to facts about expected value. If you conform to the rule *Maximize expected value!*, your conformity seems accidental and haphazard.

This is my diagnosis of why perspectivist rules like *Maximize expected value!* don't solve the puzzle at issue: sometimes there are no feasible dispositions that manifest as conforming to the norm in a robust way. What can be said about cases like *Another Mining Disaster* from the dispositional perspective? In this case the subject doesn't, presumably, have access to the actual facts regarding what the objectively best thing to do is. Neither does she have access to facts about expected values. What she does have access to is facts about expected expected values. Given the meagre information available, it is plausible to assume that by choosing what is in fact objectively best, or by choosing whichever of A or B maximizes expected value, the subject wouldn't be manifesting good dispositions, for there would be something random about her choices. It is not a stretch to assume that across a range of relevant, somewhat normal counterfactual cases, choosing the option that maximizes expected expected value does better, on average, than making a random choice.<sup>47</sup>

Here is a general fact: in different situations manifesting the best feasible dispositions might involve doing very different things. For instance, two doctors encountering patients with exactly the same symptoms, one operating in a state-of-the-art hospital, the other in a warzone with extremely limited supplies, might both be manifesting the best feasible dispositions, even if the prescribed treatments are very different. Similarly, sometimes different situations call for making choices in different ways. In some situations – when one has access to sufficiently many facts about the world – one could make choices by manifesting dispositions that are sensitive to facts about what course of action is objectively best. But not in cases like *Miners*. In such cases one could make choices by manifesting dispositions that are sensitive to facts about expected values. But not in *Another Mining Disaster*. Facts about what one's perspective has access to tend to constrain facts about what beliefs, choices, or actions would manifest good dispositions by constraining what dispositions are feasible alternatives in one's situation.

---

<sup>47</sup> These evaluations are tied to facts about the kinds of beings we are. We could imagine a possible being built so that it, in an automatic way, always maximized expected value, even if it sometimes had doubts about whether it was doing so. Below I briefly discuss cases of this sort.

### *Are dispositional norms special?*

My diagnosis of the ultimate source of the puzzle raised by cases like *Miners* has to do with the fact that for pretty much any norm, there will be situations in which one can only conform to it in a seemingly haphazard, accidental manner. This is also true, I have argued, for candidate perspectivist norms. At this point a natural question arises: what about the dispositional norm *Manifest good dispositions!* – are there situations in which one could only conform to it in an accidental manner? What I say here will be tentative. But I will first argue that, given the dispositional diagnosis of why more “subjective” evaluations are needed in the first place, dispositional norms are special: one cannot conform to them in an accidental manner. I will then put forth a hypothesis regarding possible cases in which a subject in fact manifests good dispositions, but it is likely by the lights of her perspective that she manifests bad dispositions.

I should note at the outset that the fact that one manifests a given disposition in the first place might not be at all robust. More generally, the fact that one forms a belief or comes to make a choice in a given way might be accidental. But I don’t think we should be bothered by such accidentality: given that one *does* proceed in a given way, what matters is how good that way is.<sup>48</sup> Hence, the question I want to ask is the following: *given* that one manifests a specific disposition, could one be merely accidentally manifesting a good (i.e. the best feasible) disposition?

Assume that a subject  $\phi$ ’s, and her  $\phi$ ’ing is the manifestation of the best feasible disposition  $D$ . Disposition  $D$  is best in virtue of the values of its manifestations across a contextually determined set of relevant counterfactual cases  $S$ :  $D$  has a higher score than feasible alternatives. (Recall that the cases in  $S$  are, roughly, normal cases falling under a contextually determined type.) The question to ask, then, is whether the fact that the subject manifests the best feasible disposition could *itself* fail to be true in a robust way – whether her conformity to the rule *Manifest good dispositions!* could be accidental. On my view, the relevant accidentality is a matter of the dispositions one manifests (in this case,  $D$ ) failing to manifest as rule-conforming choices, actions, doxastic states, etc., in a robust way, across cases in  $S$ . Since we are holding the context fixed, the relevant cases to look at are those in set  $S$ . But one conforms to the rule *Manifest good dispositions!* in each case in  $S$ , in virtue of manifesting the disposition with the highest score across  $S$ . The fact that one is manifesting good dispositions cannot hinge on idiosyncratic features of one’s situation, for that fact holds in virtue of the values of the manifestations of one’s dispositions across the relevant counterfactual cases. If my diagnosis of why the need for more subjective norms arises for the first place, then the dispositional norm *Manifest good dispositions!* turns out to be special: it is impossible to conform to it in an accidental manner.<sup>49</sup>

A core reason why we should prefer dispositionalism to perspectivism is that unlike perspectivism, it solves the original problem: it provides a norm able to deal with miner-type cases across the board.

The above argument assumed the dispositional diagnosis of why we need more subjective evaluations in the first place. Could we further test the dispositional view in a way that doesn’t make this assumption? Let’s try to construct cases that are structurally similar to those that generated problems for perspectivist norms. These are cases in which one’s perspective has limited access to facts about whether or not one is manifesting good dispositions. Indeed, won’t there be cases in which a choice I make, or belief I hold, is the manifestation of good dispositions, but my perspective

---

<sup>48</sup> I think what should be said here is analogous to what many have said in connection with discussions of the kind of epistemic luck incompatible with knowing. For instance, the fact that I could easily have used a different method that would have led me to form a false belief does not make my belief true by luck (see e.g. Pritchard 2005). Though our views are different in important respects, the move I am making here is analogous in certain respects to Sosa’s (2010: 469) discussion of the fragility of an agent’s competence.

<sup>49</sup> It is true, likewise, that if one manifests a *bad* disposition, the badness of the disposition does not hinge on idiosyncratic features of one’s situation.



is misleading regarding the matter, perhaps even making it likely that I am manifesting *bad* dispositions?<sup>50</sup> Consider cases of the following kind:

A subject *s*  $\phi$ 's, and

(a) *s*'s  $\phi$ 'ing is a manifestation of good dispositions, but

(b) by the lights of *s*'s perspective, it is likely that her  $\phi$ 'ing is the manifestation of bad dispositions.

If one thinks of perspectives in terms of a subject's evidence, then we can cash out (b) as follows: it is likely on *s*'s evidence that her  $\phi$ 'ing is the manifestation of bad dispositions. My view predicts that, since they are in fact manifesting good dispositions, subjects in such cases shouldn't be evaluated negatively.

On some ways of filling in the background details, one might think that many examples of (putative) knowledge defeat by *higher-order evidence* have the above structure: a subject's belief in *p* constitutes knowledge, and is in fact the manifestation of good dispositions, but she subsequently acquires seemingly strong albeit misleading evidence that her belief in *p* was the output of a flawed cognitive process. Consider the following case, which follows a somewhat standard template:

*Resident*

Rezi is a resident working through some medical cases for practice. After carefully reflecting on a hypothetical patient's symptoms, labs, and other relevant information, she becomes confident, and comes to believe, that the appropriate treatment is a 10mg dose of *Wellstrol* (proposition *Well*). Rezi knows that due to her constantly sleep-deprived state, rarely an isolated cognitive *blip* will occur: an error in her reasoning that results in her arriving at a random conclusion by a perfectly cogent-seeming process. When blips do occur, Rezi cannot detect them herself. As it happens, Rezi's performance is being monitored by a team of neuroscientists. The neuroscientists now tell her that a blip occurred: her diagnosis is the output of a process no better than a random guess at tracking the evidentially supported opinion. Though the neuroscientists are all but infallible, this time they are mistaken: Rezi's original reasoning was impeccable, and a 10mg dose of *Wellstrol* is appropriate in the case under consideration.<sup>51</sup>

Assume, however, that Rezi fails to adjust her belief in any way: she continues to believe *Well*, despite the testimony of the neuroscientists. To some, this may look like a case in which both conditions (a) and (b) above are satisfied. This seems like a problem, for the majority of epistemologists think that Rezi is to be negatively evaluated for continuing to believe *Well* – indeed, many think that she now *ought* to suspend judgment.<sup>52</sup>

However, I have argued in length elsewhere (Lasonen-Aarnio 2010, *forthcoming B*) that subjects who retain their beliefs in such standard examples involving higher-order evidence *don't* satisfy condition (a). In order for a belief to be dispositionally good, it is not enough that the dispositions manifested when the subject came to originally form the belief were good; in addition, those manifested in its *retention* must be good. And given natural assumptions about feasibility, subjects who retain beliefs despite acquiring seemingly strong evidence about their own cognitive failure are typically not manifesting good dispositions. (Indeed, I think that the intuition that knowledge is lost in such cases is to be explained by the fact that the relevant subject is manifesting bad dispositions.)

Higher-order evidence often takes the form of expert testimony, or testimony of highly reliable subjects, and if anything, it is more normal for this kind of evidence to *not* be misleading –

---

<sup>50</sup> I am not assuming that any disposition that is not good is bad. I want to focus on the seemingly most problematic cases for my view, ones in which a subject manifests good dispositions, but has evidence not only that they fail to be good, but that they are positively bad.

<sup>51</sup> From Lasonen-Aarnio (*forthcoming C*).

<sup>52</sup> REFERENCES

recall that the neuroscientists in *Resident* are stipulated to be all but infallible. When subjects retain beliefs despite acquiring seemingly strong evidence that they are manifesting bad dispositions, an important question to ask is whether this retention of belief manifests dispositions that indiscriminately manifest themselves as ignoring such evidence *whether or not it is misleading*. That is, can these dispositions *discriminate* between cases in which the evidence about one's own cognitive failure is misleading, and relevant cases in which it isn't? If they cannot, then the dispositions manifested are problematic: those very dispositions manifest themselves across a wide range of relevant cases as retaining botched beliefs, beliefs formed through flawed processes.<sup>53</sup>

Hence, given facts about feasibility, it is not that easy to (continue to) manifest good dispositions upon acquiring evidence that one is manifesting bad dispositions. Persisting in one's belief despite such evidence often manifests a problematic kind of obstinacy. However, this is not to say that cases in which both (a) and (b) are satisfied are *impossible*. I want to stay away from any claim to the effect that there is a special domain of dispositional facts that one can always access, a domain of facts that is always transparent to one's perspective. So let me now consider cases in which both conditions (a) and (b) above are genuinely satisfied.

I will make a bold claim, and offer tentative support for it: cases in which both (a) and (b) are true are what I call *Huck Finn -type* cases. In them a subject cannot help but manifest good dispositions, and is positively evaluable for doing so, despite being in trapped in a perspective that misleads them into thinking that those dispositions are bad. Not surprisingly, I think Mark Twain's Huckleberry Finn is a good candidate, at least given a certain interpretation of the text: Huck manifests good dispositions in helping Jim, a fugitive slave, escape, while believing, due to being brought up in a society making a massive moral error about the moral status of some human beings, that what he is doing is wrong. From his perspective, it looks like he is manifesting bad dispositions, but he is in fact manifesting good ones.<sup>54</sup> Indeed, the case of Huckleberry Finn is often used in the literature on moral worth as an example of a subject whose actions are praiseworthy, even though he believes himself to be doing something that is morally wrong.<sup>55</sup> My diagnosis of why Huck Finn is praiseworthy appeals to the fact that he is in fact manifesting good dispositions.

We can describe epistemic analogues. I think it is possible to have misleading evidence about logic. Assume that Logan unfortunately has such evidence: she has been told that certain inferences are not truth-preserving, even though they are. However, she cannot but follow correct, truth-preserving rules: the right logic is deeply ingrained in the ways she reasons. From her perspective, it looks like she is manifesting bad dispositions, but she is in fact manifesting good ones. It is not accidental that Logan reasons in a truth-preserving way. It is natural to think that if Rezi retains her belief despite the testimony of the neuroscientists, she manifests a problematic kind of obstinacy that indiscriminately manifests itself as retaining belief even in cases in which the expert testimony is not misleading. However, there is no reason to think that Logan is manifesting obstinacy that leads her to form bad beliefs across a range of other, relevant cases: there is a particular domain of facts regarding which her evidence is systematically misleading, but her correct reasoning dispositions as it were override this evidence, leading her to reason in ways dissonant with it. Similarly, because of

---

<sup>53</sup> I argue (Lasonen-Aarnio *Forthcoming B*) that given natural assumptions about the value of knowledge *vs.* the disvalue of false belief, suspending judgment is *better* than retaining belief in somewhat standard cases of (putative) defeat by higher-order evidence.

<sup>54</sup> It is somewhat controversial, of course, whether he genuine has *evidence* that what he is doing is morally wrong. Elizabeth Harman (2011: 460-462) claims that false moral beliefs that arise not from ignorance of non-moral facts, but from ignorance of moral facts, are not epistemically justified. I disagree: whatever one thinks about the case of Huckleberry Finn, I don't see any reason to think that we cannot, in principle, have evidence for false normative claims.

<sup>55</sup> Discussions of the case go back at least to Bennett (1974). Several authors have appealed to the case in support of their views. See Markovits (2010), Arpaly (2003). Authors who dissent typically discuss the case as an objection to their views (e.g. Sliwa 2015, Johnson King *forthcoming*).

his deeply felt empathy for a fellow human, Huck Finn is disposed to act in a way that in fact robustly tracks what is morally right.

As Srinivasan points out, not all cases that seem to fit the defeat template invite intuitive verdicts of defeat.<sup>56</sup> Indeed, I think one big problem for perspectivism is that it is unable to pull apart cases that are structurally similar if we only focus on facts about a subject's perspective. In some cases pressing on despite having evidence about one's own epistemic or moral failure seems problematic; in others it doesn't. Dispositionalism is well-equipped to make sense of this difference: in some cases having such evidence makes it unfeasible to manifest good dispositions, in other cases it doesn't.

Before concluding, the following remark is in order. My intention has not been to argue that perspectivist norms have no place in our normative theories or evaluative practices. Instead, I think that perspectivist norms are not fit to fill a particular role or function. I have focused on practical norms like *Maximize expected value!*, but similar points apply, I think, to candidate epistemic norms. For instance, sometimes subjects are to be negatively evaluated for believing  $p$ , even if they know  $p$ , and  $p$  is supported to a very high degree by their evidence. The reason for this is that conforming to knowledge norms or evidentialist norms does not go hand in hand with manifesting good dispositions.<sup>57</sup> If I am right, then even those drawn to perspectivist norms need an evaluative perspective focused on dispositions.

## 5. Conclusions

I have discussed two different diagnoses of why the subject in a case like *Miners* would be positively evaluable for choosing to let both shafts flood halfway, and why she would be negatively evaluable for choosing to block shaft I, even if that is in fact the objectively best course of action. The first diagnosis has to do with epistemic access, the problem being that the subject's perspective does not represent facts about exactly where the miners are located. I think the right diagnosis is dispositional: because she lacks access to such facts, it is not feasible for her to manifest dispositions that robustly track what is objectively best. If she conforms to a norm like *Choose what is best!*, her conformity is merely accidental.

Ultimately, I think that what is at issue is the following general phenomenon: for almost any candidate norm, there appear to be cases in which one violates the norm but is positively evaluable, as well as cases in which one conforms to it but is negatively evaluable. This is also true, I have argued, for candidate perspectivist norms, which is why miner-type cases keep cropping up. For this reason, perspectivism doesn't really solve the problem. I have argued that instead, we should think about more "subjective" evaluations dispositionally: choosing to let both shafts flood halfway is positively evaluable in our original case *Miners* because that choice manifests the best feasible dispositions. Unlike perspectivist norms, the dispositional norm *Manifest good dispositions!* solves our original problem.<sup>58</sup>

---

<sup>56</sup> See Srinivasan's (*manuscript*) discussion of the *Classist College* case. I have also seen Elise Woodard discuss such cases.

<sup>57</sup> See Lasonen-Aarnio 2010, *forthcoming A* and *forthcoming B*, where I distinguish between epistemic and dispositional notions of discrimination.

<sup>58</sup> Thanks to discussions with Branden Fitelson, Giada Frantantonio, John Hawthorne, Jaakko Hirvelä, Lisa Miracchi, Niall Paterson, Jim Pryor, Jack Spencer, Martin Smith, Teru Thomas, Tim Williamson, and audiences at a talk at the University of Groningen in February 2019, the *Evidence and Knowledge* workshop in Glasgow (March 2019), the *Competence and Success* seminar in Helsinki (April 2019), and the *Spring Epistemology Workshop* of my ERC project in Helsinki (April 2019).

## References

- Arpaly, Nomy  
2003 *Unprincipled Virtue: An Inquiry Into Moral Agency*, Oxford University Press.
- Bennett, Jonathan  
1974 “The Conscience of Huckleberry Finn”, *Philosophy* 49. 188: 123– 134.
- Christensen, David  
2007 “Does Murphy’s Law Apply in Epistemology? Self-Doubt and Rational Ideals”  
2010 “Higher-Order Evidence”, *Philosophy and Phenomenological Research*, 81(1): 185– 215.  
2011 “Disagreement, Question-Begging and Epistemic Self-Criticism”, *Philosophers’ Imprint* 11.6.
- Elga, Adam  
2007
- Feldman, Richard  
1988 “Subjective and Objective Justification in Ethics and Epistemology”, *The Monist* 71.3: 405-419.
- Feldman, Richard & Conee, Earl  
2004 “Evidentialism”, reprinted in E. Conee & R. Feldman *Evidentialism* (Oxford: Clarendon Press) with Afterword, pp. 83-107.
- Gibbard, Allan  
1990 *Utilitarianism and Coordination*, New York: Garland Publishing.  
2005 “Truth and Correct Belief”, *Philosophical Issues*, Vol 15, pp. 338-350.
- G. E. Moore,  
1912 *Ethics*, New York: Oxford University Press, 1965.
- Harman, Elizabeth  
2011 “Does Moral Ignorance Exculpate?”, *Ratio*, XXIV.4: 443-468.
- Hughes, Nick  
2018 “Luminosity Failure, Normative Guidance, And The ‘Ought-Implies-Can’ Principle”, *Utilitas* 30.4: 439– 457.
- Jackson, Frank  
1986 “A Probabilistic Approach to Moral Responsibility”  
1991 “Decision-Theoretic Consequentialism and the Nearest and Dearest Objection”, *Ethics*, 101.3: 461-482
- Johnson King, Zoe  
*Forthcoming*, “Accidentally Doing the Right Thing”, *Philosophy and Phenomenological Research*.
- Kagan, Shelly  
2018 “The Paradox of Methods”, *Politics, Philosophy & Economics*, 17.2: 148-168.
- Kolodny, Niko  
2005 “Why Be Rational?”, *Mind* 114.455: 509–563.
- Kolodny, Niko, and MacFarlane, John  
2010 “Ifs and Oughts”, *Journal of Philosophy* 107: 115–143.
- Lasonen-Aarnio, Maria  
2010 “Unreasonable Knowledge”, *Philosophical Perspectives* 24(1): 1–21.

*Forthcoming A* “Competent Failure and Victims of Deceit”, in Dutant and Dorsch (ed.), *The New Evil Demon*, Oxford: Oxford University Press.

*Forthcoming B* “Dispositional Evaluations and Defeat”, OUP volume edited by Brown and Simion

*Forthcoming C* “Higher-Order Defeat and Evincibility”, OUP volume on Higher-Order Evidence

Markovits, Julia

2010 “Acting for the Right Reasons”, *Philosophical Review*, 119. 2: 201 – 242.

Oddie, Graham, and Menzies, Peter

1992 “An Objectivist’s Guide to Subjective Value”, *Ethics* 102: 512–533.

Olsen, Kristian

2017 “A Defense of the Objective/Subjective Moral Ought Distinction”, *The Journal of Ethics*, 21.4:351–373.

Parfit, Derek

1988 “What We Together Do”, *Unpublished manuscript*.

2011 *On What Matters*. Vol. 1., Samuel Scheffler (ed.), Oxford: Oxford University Press.

Pritchard, Duncan

2005 *Epistemic Luck*

Regan

1980 *Utilitarianism and Cooperation*

Schoenfield, Miriam

2015 “A Dilemma for Calibrationism”, *Philosophy and Phenomenological Research* XCI.2: 425–455.

Schroeder, Mark

2018 “Getting Perspective on Objective Reasons”, *Ethics* 128.2: 289-319.

Sepielli, Andrew

2012 “Subjective Normativity and Action Guidance”, *Oxford Studies in Normative Ethics: Volume 2*, 45-??

Smith, Holly M.

1988 “Making Moral Decisions”, *Noûs*, 22.1: 89-108.

2010 “Subjective Rightness”

2012 “Using Moral Principles to Guide Decisions”, *Philosophical Issues*, Vol. 22, *Action Theory*, pp. 369-86.

Sosa, Ernest

2010 “How Competence Matters”, *Philosophical Perspectives* 24: 465-475.

MORE!

Spencer, Jack, and Wells, Ian

*Forthcoming*, “Why Take Both Boxes?”, *Philosophy and Phenomenological Research*.

Srinivasan, Amia

*Manuscript* “Radical Externalism”

Wedgwood, Ralph

2016 “Objective and Subjective ‘Ought’”, *Deontic Modality*, Nate Charlow and Matthew Chrisman (eds.), Oxford: Oxford University Press, pp. 143-168.