

MIT Open Access Articles

*Knowledge of Objective 'Oughts':
Monotonicity and the New Miners Puzzle*

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation: Muñoz, Daniel and Jack Spencer. "Knowledge of Objective 'Oughts': Monotonicity and the New Miners Puzzle." *Philosophy and Phenomenological Research* (May 2020) © 2020 The Author(s)

As Published: 10.1111/phpr.12702

Publisher: Wiley

Persistent URL: <https://hdl.handle.net/1721.1/130151>

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Terms of use: Creative Commons Attribution-Noncommercial-Share Alike



Knowledge of Objective ‘Oughts’: Monotonicity and the New Miners Puzzle

Daniel Muñoz (Monash University)

Jack Spencer (MIT)

Forthcoming in *Philosophy and Phenomenological Research*.

In the classic Miners case, an agent subjectively ought to do what they know is objectively wrong. This case shows that the subjective and objective ‘oughts’ are somewhat independent. But there remains a powerful intuition that the guidance of objective ‘oughts’ is more authoritative—so long as we know what they tell us. We argue that this intuition must be given up in light of a monotonicity principle, which undercuts the rationale for saying that objective ‘oughts’ are an authoritative guide for agents and advisors.

1. Introduction

When agents are uncertain, we distinguish what they *objectively* ought to do—roughly, what should be done given full knowledge of the situation—from what they *subjectively* ought to do, where this is sensitive to their false and gappy beliefs.¹ Clearly, these ‘oughts’ can come apart. You objectively ought to play the slots if the machine will in fact pay out, but if the odds are low, you subjectively ought to play it safe.

For the gambler and other uncertain agents, the subjective ‘ought’ is the proper guide to action; it is the ‘ought’ of rationality. An agent can be guided only by what is “present to the agent’s

¹ The distinction between objective and subjective ‘oughts’, and the question of which is prior, have received enormous attention from philosophers (Ross, 1939: 146–167; Ewing, 1947: 128; Brandt, 1959: 360–67; Parfit, 1984: 25, 2011: Chapter 7; Jackson, 1986, 1991: 128; Thomson, 1990: 172–73, 2008: 188; Prichard, 2002: Chapter 6; Scanlon, 2008; Zimmerman, 2008; Hurka, 2014: 78–85; Carr, 2015; Wedgwood, 2016; Wodak, 2017). By the subjective ‘ought’, we mean the ‘ought’ that guides deliberation by virtue of being sensitive to agents’ limited knowledge. For simplicity, we will only discuss agents whose beliefs are rational, in the sense of being reasonable given their evidence, and we assume that agents are risk-neutral (see Buchak 2013).

mind” (Jackson, 1991: 467; cf. Gibbons, 2009: 174). Beliefs are present; unknown facts are not. But it is almost irresistible to think that, in some sense, the objective ‘ought’ is more *authoritative*. It is the ideal guide, the ‘ought’ that would guide us if only we knew more, the ‘ought’ whose advice we most desire. It is only when the objective ‘ought’ eludes us that we retreat to the subjective. The rough idea here is that, given knowledge of the objective ‘ought’, we subjectively ought to follow its advice.

This way of seeing the subjective ‘ought’ is extraordinarily appealing. We argue that it is even more problematic than has been realized. The familiar problem is that, in the classic “Miners Case,” an agent subjectively ought to do something that they know they objectively ought not to do. We argue that the case has an even stronger implication: given a plausible monotonicity principle, the agent subjectively ought not to do something that they know they objectively ought to do. This result puts serious pressure on the idea that there is anything normative about objective ‘oughts’, and it challenges some popular claims about the norms of reasoning and advice.

2. The New Miners Puzzle

Tragic news:

Miners

There was a disaster in the quarry, and 100 miners are trapped in Shaft A; the nearby Shaft B is empty. You know that, if you do nothing, the shafts will partly flood and 10 miners will die. You also know that, if you block the shaft where the miners are, you will save all 100; and if you block the empty shaft, the other will totally flood, drowning all 100. But your evidence doesn’t tell you where the miners are; for you, it’s a 50/50 guess.²

² The Miners case is due to Regan, 1980: 265, n.1 and Parfit, ms.; cf. Jackson’s (1991) drug cases. Kolodny and MacFarlane (2010) present the miners as a counterexample to *modus ponens*. In our version of the case, we will suppose that you know what you know with full certainty (and that you are certain that your credences are rational). This is just to keep the case clean and simple. Our arguments would go through even if, instead of knowledge, we were to attribute mere true belief or rational certainty. But we note that the relationship between knowledge of objective ‘oughts’ and subjective ‘oughts’ might be even more tenuous if one accepts a fallibilistic conception of knowledge; cf. Littlejohn (2018). (Thanks to an anonymous referee for suggesting that we clarify this point.)

You are the miners' unique potential savior, and the more you save, the better. What should you do?

Objectively, you clearly ought to block Shaft A, since that saves the most lives. But subjectively you ought to block neither shaft, since a 100% chance of saving 90 lives beats a 50% chance at saving 100.

The Miners case leads to a familiar puzzle, with an important lesson about objective and subjective 'oughts'. It is tempting to think that what we subjectively ought to choose the option with the best chance of being objectively right. At the very minimum, if we know that we objectively ought *not* to do a certain option, then it can't be that we subjectively ought to do it:

Negative Link

If an agent knows that $OW(X)$, then $\sim SO(X)$.

Where ' $OW(X)$ ' means that X is objectively wrong (i.e. that you objectively ought not to do X), and ' $\sim SO(X)$ ' means that it's not true that you subjectively ought to do X .

But this link doesn't hold in the Miners case. For in that case, you know that blocking neither is objectively the wrong choice:

Objectively Wrong to Block Neither
 $OW(\text{Block Neither Shaft})$.

And yet:

Subjectively Ought to Block Neither
 $SO(\text{Block Neither Shaft})$.

You subjectively ought to block neither shaft *even though you know that you objectively shouldn't*. As Parfit (2011: 159–161) puts it, you (subjectively) shouldn't even *try* to do as you objectively ought, because you don't know which shaft you objectively ought to block—and a wrong guess spells disaster.³

³ Dorsey puts it in terms of what's "inappropriate": "it is always morally inappropriate to perform an action one knows is not the right thing, and we *know* that blocking neither shaft is not, as a matter of objective morality, the right thing" (2012: 19). See also Fox, 2019: 226.

This implication is now widely accepted: we must give up NL. But as consolation, it is tempting to think that we can retain *some* link between objective and subjective ‘oughts’. After all, if only you knew which shaft you objectively ought to block, then of course it would follow that you subjectively ought to block it. This suggests a modest:

Positive Link (PL)

If an agent knows that $OO(X)$, then $SO(X)$.

Where ‘ $OO(X)$ ’ means that you objectively ought to do X, and ‘ $SO(X)$ ’ means you subjectively ought to do X. Again, the link seems to spring from the natural idea that objective ‘oughts’ give ideal advice. Sepielli (2018: 789) lays this out very nicely:

suppose I am sure that I objectively ought to do A. Then the question of what I subjectively ought to do need never arise for me. I can simply guide my behavior by my objective “ought”-thought. And if for whatever reason I do ask myself what I subjectively ought to do, my answer cannot, on pain of obvious incoherence, be anything other than “A”.

And so PL, unlike NL, seems hard to resist. If you know what the objective ‘ought’ advises, and its advice is perfect, why refuse to listen?

Well, think about Miners. You, our uncertain agent, know that one of the following is true:

Objectively Ought to Block Shaft A
 $OO(\text{Block Shaft A})$.

Objectively Ought to Block Shaft B
 $OO(\text{Block Shaft B})$.

But you are 50/50 on which is true, and because of your ignorance:

Not: Subjectively Ought to Block a Shaft
 $\sim SO(\text{Block a Shaft})$.

Now, it is widely held, for good reasons (reviewed in §2 below), that ‘ought’ is *upward monotonic*. This means that if you ought to do a certain act X, and X-ing entails Y-ing, then you ought to do Y. So in the case of the objective ought, we have:

Upward Monotonicity for Objective ‘Ought’ (UM)

If X-ing entails Y-ing, then OO(X) entails OO(Y).

This makes intuitive sense. If you objectively ought to call your aunt, then you objectively ought to call someone. If you objectively ought to go to Paris, then you objectively ought to go to France. In the Miners case, upward monotonicity implies that if you objectively ought to block Shaft A (or objectively ought to block Shaft B), you objectively ought to block a shaft. Let’s suppose that you indeed *know* that you objectively ought to block a shaft, since you know that either you objectively ought to block Shaft A or objectively ought to block Shaft B, and you know UM.

Now we have a new puzzle. *You know that you objectively ought to block a shaft, and yet it’s not true that you subjectively ought to do so.* There is a tension between the authority and monotonicity of the objective ‘ought’.

This tension, we should emphasize, is genuinely new. The original Miners Puzzle teaches us that objectively *wrong* acts might be subjectively *permissible*, or indeed subjectively required. Sounds funny the first time you hear it, but it makes perfect sense on reflection. Since you don’t know where the miners are, you don’t know which shaft you objectively ought to block—and since a wrong guess means certain death, it is no wonder that you subjectively should play it safe. But notice that you don’t have to guess whether you objectively ought to block a shaft; given UM, you *know* you ought to, and indeed you know with *certainty* if we suppose that you are antecedently certain that the miners are in Shaft A or Shaft B. This is exactly the kind of knowledge that, by our link principle PL, normally entails a subjective ‘ought’. The objective ought *can* guide you toward blocking a shaft, and it is supposedly the ideal guide, and yet subjectively you shouldn’t listen. That is the new puzzle.

What’s the solution? The two options, apparently, are to reject UM or PL. We think a good solution will reject PL (even if UM happens to be false!), and rejecting PL has surprisingly broad

implications, as it puts sharp limits on what can follow from objective ‘oughts’ and knowledge thereof.⁴ In a nutshell: there are many tantalizing ideas tossed around about how the objective ‘ought’ can offer guidance, both to agents in the know and to savvy advisers. But these ideas can’t be true in general, given UM, and they are hard to qualify.

To give you a flavor of the “tantalizing ideas” we have in mind, consider this passage from Parfit (who uses “fact-relative” and “belief-relative” instead of “objective” and “subjective”):

when we are trying to decide what to do, we can ignore the fact-relative senses of ‘ought’, ‘right’, and ‘wrong’. We cannot try to do what is right in the fact-relative *rather than* the belief-relative sense. Suppose I believe that, to save your life, I must act in a certain way. Though I know that my belief might be false, I cannot try to do what *would in fact* save your life, since what *I now believe* is that acting in this way *would in fact* save your life. We cannot base our decisions on the facts except by basing our decisions on what we now believe to be the facts. (2011: 161)

But this nice idea—that we cannot try to do as we objectively ought rather than as we subjectively ought—is false if PL fails. Given UM, you know that you objectively ought to block a shaft and yet, in defiance of PL, it doesn’t follow that you subjectively ought to. You *can* try to follow the objective rather than the subjective ought, at least when it comes to the question of whether to block a shaft. Even given your imperfect glimpse of the situation, you can see that you objectively ought to block a shaft even though this is not recommended by the subjective ‘ought’, and while you cannot reliably block the particular shaft that you objectively ought to (since you don’t know which one the miners

⁴ Does anyone reject PL? Dorsey comes close: “the best moral decision (i.e., the subjectively right act) in [the Miners] case is to block neither shaft. But equally obvious is that one objectively ought to block either A or B; blocking neither shaft will guarantee a morally suboptimal outcome” (2012: 18). But there is a scope ambiguity here. Is Dorsey saying that one objectively ought to (block A or B)—denying PL—or is he saying that either A or B is such that it objectively ought to be blocked? Fox is similarly ambiguous (2019: 226n.10). By contrast, Sepielli (2018: 790) clearly has in mind the second reading, which is consistent with PL.

are trapped in), you certainly can succeed in *blocking a shaft* if you try to, thus “discharging” *one* of your objective obligations.⁵ If there is a truth in Parfit’s claims, it will require excavation.

But first, we should consider the prospects of solving our puzzle by rejecting UM (§3). After that, we consider what is lost by giving up PL (§4), as well as what we might learn about the proper role of objective ‘oughts’ for agents and advisors (§5).

3. Monotonicity

Recall our puzzle. Your rational credence is split evenly between:

Objectively Ought to Block Shaft A
OO(Block Shaft A).

Objectively Ought to Block Shaft B
OO(Block Shaft B).

But you know:

Upward Monotonicity for Objective ‘Ought’ (UM)
If X-ing entails Y-ing, then OO(X) entails OO(Y).

And so you know that, no matter which shaft the miners are in, you objectively ought to block a shaft. Let’s write this as:

Knowledge of Objective ‘Ought’ (KO)
You know that OO(Block a Shaft).

But it seems plausible that:

Positive Link (PL)
If an agent knows that OO(X), then SO(X).

And BP and KO are inconsistent with a plain fact about your situation:

⁵ Perhaps it is not quite right to say that you “discharge your obligation” to block a shaft when you do so in the wrong way, saving zero miners. Grant and Phillips-Brown (2019) defend a similar point about desire: you don’t “satisfy your desire” to drink milk by drinking spoiled milk.

Not: Subjectively Ought to Block a Shaft
~SO(Block a Shaft).

What gives?

It is natural to think that we need to give up KO, which we arrive at given the setup of the case along with UM. So why not scrap UM? Why can't we just deny that you objectively ought to block a shaft and be done with the puzzle? The answer is that UM is backed up by some formidable arguments, and the objections to it, even if they work, don't apply in the Miners case.

Let's start with a simple argument for UM, which is that it follows from the orthodox semantics of 'ought'.⁶ The orthodoxy: 'ought' is a modal operator, equivalent to a quantificational claim about certain possible worlds. Simplified a bit, the idea is that 'ought(X)' is true just if you do X in *all of the best relevant possible worlds*. To determine how to rank worlds from "worst" to "best" (in line with an "ordering source"), and how to select which worlds are relevant (the "modal base"), speakers often must look to context. Worlds that you can't possibly bring about, for example, might be ruled irrelevant, and whether the ranking of worlds is sensitive to the agent's ignorance will depend on whether the intended reading of 'ought' is objective.

Here is the orthodox take on objective 'oughts' in the Miners case. You objectively ought to block Shaft A just if, in all of the objectively best relevant worlds, you block Shaft A. (The objectively best worlds, we assume, are the ones where you save the most lives.)

Straightaway the orthodoxy entails UM. Why? Because the generalized quantifier 'all worlds' is *itself* upward-monotonic. If all (of the best relevant) worlds are ones at which you do X, and doing X entails doing Y, then all (of the best relevant) worlds are ones at which you do Y. All worlds with

⁶ The view is due to Kratzer (1977, 1981). On Miners and the semantics of deontic modals, see Kolodny & MacFarlane, 2010; Dowell, 2012; von Fintel, 2012; Charlow, 2013; and Carr, 2015.

chairs are worlds with furniture; all worlds where dogs bark loudly are worlds where dogs bark. So, since you block Shaft A in all the objectively best relevant worlds, you also *block a shaft* in those worlds. And that suffices for it to be true that you ought to block a shaft.⁷

Some writers, however, reject the orthodox view precisely because they think that ‘ought’ should not turn out to be monotonic (e.g. Lassiter, 2011; Cariani, 2013). There are two main counterexamples to monotonicity. The first is known as “Ross’s Paradox” (Ross, 1941, 1944; von Fintel, 2012; Cariani, 2013). Suppose you ought to mail an important letter. By monotonicity, it should follow that you ought *to mail the letter or burn it*, since that disjunctive act is implied by your mailing it. But something smells fishy about that inference. The source of the stench, von Fintel (2012: 6) thinks, is that ‘You ought to mail the letter or burn it’ seems to imply that you ought to mail the letter *and* that you ought to burn the letter—this is called a “free choice” inference (Kamp, 1973). Suffice it to say that it is controversial whether this inference really is packed into the meaning of ‘ought’ (as argued e.g. by Cariani, 2013) or merely pragmatic (as argued by von Fintel, 2012). We will not take sides here, but we note that the pragmatic explanation of free choice is available.

The second kind of counterexample to monotonicity is Jackson and Pargetter’s (1986: 235) “Professor Procrastinate” (see also Goldman, 1978; Jackson, 1985; Portmore, 2019).

Professor Procrastinate

Professor Procrastinate is invited to review a book. Ideally, he would accept the invitation and write the review. But he knows that, if he were to accept, he wouldn’t write anything; he would keep putting the task off—which would be even worse than declining the invitation.

⁷ Notice that the point holds even if ‘ought’ means “true in *most* of the best worlds,” a view that Copley (2006) attributes to Horn (1972). “Most Fs ___” is also upward monotonic.

Should Professor Procrastinate accept the offer? Jackson and Pargetter think: *no*. That would lead to the worst outcome, since he would in fact flake if he were to accept. But ought he to accept *and write the review*? They think: *yes*. That would lead to the best outcome.

So we can't just appeal to the authority of the orthodox semantics; there are hard cases to reckon with. Fortunately, there is a second, more direct argument for UM: it passes the negation test. Under the scope of negation, an upward-monotonic environment typically “flips” to a *downward*-monotonic environment, one where substituting *stronger* propositions (predicates, etc.) preserves truth. For example, the blank in ‘All Fs are ___’ is upward-monotone, as we saw earlier (with ‘All worlds’). ‘All cats are very cute’ entails ‘All cats are cute’, though not vice versa. Negated, we find the opposite: ‘Not all cats are cute’ entails ‘Not all cats are very cute’, though not vice versa.

So we have a test. Does the blank in ‘ought(____)’ become a downward-monotone environment when scoped under negation?

Apparently, yes, though there is a complication: it is difficult to get ‘ought’ under the scope of negation in natural language sentences (in ‘you ought not to do it’, the negation wants to scope under ‘ought’; see von Stechow, 2012). But we can run tests for other deontic necessity modals (like ‘have to’, ‘must’, and ‘obliged to’, as opposed to possibility modals like ‘may’). ‘You don’t have to wear a tie or scarf’ clearly entails ‘You don’t have to wear a tie’. It would be “insane” (in von Stechow’s (2012: 13) words) to say ‘You don’t have to wear a tie or scarf, but, of course, you have to wear a tie’—this sounds flatly contradictory. Moreover, so-called “negative polarity items” (NPIs), like ‘any’ or ‘lift a finger’, appear to be licensed in the context of ‘You don’t have to ____’. This is a classic sign of downward-monotonicity. ‘You have to pick up any groceries from the store’ is infelicitous; ‘You *don’t* have to pick up any groceries from the store’ is perfectly fine (maybe even a relief). This

suggests that the latter sentence contains a downward-monotone context, which would mean that the negation-free version is upward-monotone.

We have just seen some considerable arguments for UM.⁸ But in fact, we don't need them. Even if UM can sometimes fail, it clearly *doesn't* fail for the objective 'ought' in the Miners case. Just ask yourself: objectively speaking, should you block a shaft? As a reminder, here are some authors' reports of how they construe the objective 'ought'.

As a heuristic for the objective 'ought', consider what an omniscient being would advise us to do. (Kolodny & MacFarlane, 2010: 117)

We suggest that an omniscient being would advise blocking a shaft, in the course of advising you to block the particular shaft where the miners are. Next:

We ought objectively to do an act just in case, given full knowledge of the facts, we ought to do it in the ordinary sense of 'ought'. (Adapted from Parfit, 2011: Chapter 7.)

Supposing that you know the facts, you obviously ought to block a shaft.

Moreover, the special features of the counterexamples to UM aren't present in the miners case given full knowledge. In Procrastinate, since you will put off your tasks, opting for the best general option (accepting) will lead to the worst specific option (flaking). This same kind of problem might arise if you *don't* know where the miners are; opting for the best general option (blocking a shaft) can lead to the worst (100 doomed). But since we are talking about the objective 'ought', this kind of ignorance is irrelevant; we are concerned with the case where you know the facts.⁹ Given the facts about where the miners are, you ought to block the right shaft, and so ought to block a shaft.

⁸ A third argument, due to Portmore (2019: 115), is that UM validates many everyday instances of good practical reasoning. (Portmore calls UM "deontic inheritance.")

⁹ There are some exceptions to this point. Most notably, the objective 'ought' doesn't line up with what ought to be done in the case where you know the facts if the goal of your action is *itself* to eliminate ignorance. (For example, it might be that you objectively ought to learn things because knowledge is intrinsically good.) This exception does not apply in the miners case, where the goal isn't to learn facts, but to save lives. (Our thanks to Richard Yetter-Chappel for discussion.)

Notice, also, that the miners case with full knowledge does not invite Ross's "paradoxical" free choice inferences. The fact that you objectively ought to block a shaft does not seem to entail that you objectively ought to block Shaft A *and* objectively ought to block Shaft B.

We conclude that rejecting UM isn't easy, and it isn't a satisfying way to solve our puzzle. That leaves only one option: give up BP. But what happens if known objective 'oughts' don't entail subjective 'oughts'? Can we rescue the thought that objective 'oughts' are more authoritative?

4. Relinquishing PL

We are at the crux of the New Miners Puzzle.

In the Miners case, we should accept:

Knowledge of Objective 'Ought' (KO)
You know that OO(Block a Shaft).

But is undeniable that:

Not: Subjectively Ought to Block a Shaft
 \sim SO(Block a Shaft).

And these together are a counterexample to:

Positive Link (PL)
If an agent knows that OO(X), then SO(X).

And so we must reject PL, even though it nicely expresses that objective 'oughts' are authoritative.

We think that PL is indeed false. But it is not enough to wheel out a counterexample. We also need to know why PL is so appealing. Why is *that* the way to express our authority hunches?

Can we recover some of what PL is meant to capture?

We already know that PL gets part of the Miners case right. If you know that you objectively ought to block Shaft A, then you subjectively ought to block it. But here, we are talking about a

specific option—something that cannot be done in relevantly different ways. There is no bad way to block Shaft A. But there are two ways to *block a shaft*—and blocking Shaft A is objectively far better than blocking Shaft B.

So perhaps we can simply restrict PL to choices between fully specific options, so that it doesn't apply when general options are on the menu. We replace PL with:

Restricted Positive Link (RPL)

If an agent knows that $OO(X)$, and X is a fully specific option, then $SO(X)$.

We think RPL is true, and yet still strong enough to be interesting; it is inconsistent with Evidential Decision Theory, for example, while being consistent with Causal Decision Theory. RPL is a plausible constraint on the link between objective and subjective oughts, and it goes some way towards vindicating the hunch that objective oughts are more authoritative.¹⁰

But even if RPL is true and strong, it's not strong enough: it doesn't capture all that we wanted from PL. An initial problem is applicability. Only rarely is it true that we ought objectively to do a fully specific option; typically, there are lots of indifferent ways to discharge obligations. We ought objectively to block Shaft A, but it's not true that we objectively ought to do so while whistling. (Or while not whistling!) Restricting PL to specific options risks undue triviality.

¹⁰ Consider a moral Newcomb problem (cf. Nozick, 1969). There are two boxes, one opaque and one transparent. The agent has two options. They can either save just the people trapped in the opaque box or save the people in both boxes ("one-box" or "two-box"). Ten innocents are in the transparent box. How many are in the opaque box depends on a prediction made yesterday by a reliable predictor. If the predictor predicted that the agent would one-box, they put 1,000 innocent people in the opaque box. If they predicted that the agent would two-box, they left the opaque box empty. The agent knows that two-boxing saves ten more lives than one-boxing, so the agent knows that they objectively ought to two-box. Causal decision theory recommends two-boxing, but evidential decision theory does not, thus violating RP, cf. Ahmed and Spencer (forthcoming) and Spencer and Wells (2019).

The applicability problem can be gotten round. Let A_X be the fully specific option that would be realized if an agent were to do X, and let $OP(A_X)$ be the claim that A_X is among the objectively permissible options, i.e., $\sim OO\sim(A_X)$. We then can state the following principle:

Weak Positive Link (WPL)

If an agent knows both that $OO(X)$ and $OP(A_X)$, then $SO(X)$.

An agent needn't know which fully specific option A_X is. Indeed, in typical cases, an agent won't. But, according to WPL, if the agent knows that they objectively ought to X and knows that the maximally specific option (whichever it is) that would be realized if they did X is objectively permissible, then the agent subjectively ought to X.

But there is a deeper problem, which applies both to WPL and RPL. Sometimes, PL—the unrestricted principle—is a better guide to choosing from general options. Consider another possible disaster:

Climbers

An avalanche traps 100 climbers in Cave C; Cave D is empty. But *you* don't know that. Given your evidence, you think there is a .5 chance that the climbers are all in Cave C and a .5 chance that they are in Cave D. You know that all 100 will freeze to death if you do nothing. But you have options. As the leader of the rescue team, you choose (1) whether the team will bring snowmobiles; and (2) whether the team will first go to Cave C or Cave D. If the team goes to C first, they will arrive before anyone freezes; if they go to D first, 50 miners will die before the team reaches C. Either way, you know that 10 *more* climbers will die on the way down unless the team has snowmobiles to hasten their descent.

In this cousin of the Miners case, you objectively ought to send the team to Cave C with snowmobiles in tow. But this fact can't guide your actions, at least not fully, since you can only guess where the climbers are. It is not true that you subjectively ought to send the team to C. That said, you do know that *one* thing you objectively ought to do is to send snowmobiles. After all, that will save 10 lives regardless of where the team goes first.

Here, PL tells us that you subjectively ought to send in the snowmobiles, which is exactly right. It appears that, in this case, the objective ‘ought’ *is* fit to guide your actions in a limited way, despite your limited knowledge. By contrast, neither RPL nor WPL gives you advice in Climbers. You know that you objectively ought to send snowmobiles, but that option is non-specific; you could send them first to Cave C or first to Cave D. You don’t know which specific option you objectively ought to do (so RPL doesn’t apply), nor do you know whether choosing to send the snowmobiles will result in your picking an objectively permissible specific option (so WPL doesn’t apply)¹¹. Only PL is strong enough to explain why you subjectively ought to send the snowmobiles.

So PL licenses some nice inferences. And these inferences, we think, are part of a seductive picture of rational deliberation. Rational agents take three steps. First, they do their best to find out what objectively ought to be done. Second, if they can do something that *they know they objectively ought to do* (like sending snowmobiles), they resolve to do that. Finally, the leftover choices are settled by the subjective ‘ought’ alone. This picture, in its first two steps, clearly treats the objective ‘ought’ as more authoritative: it gets first say. The final step brings a vital qualifier: when agents cannot *bear* the objective ‘ought’, they shouldn’t bother trying to comply.

On this picture, the objective ‘ought’ is not quite “useless in deliberation” (as suggested by Kolodny & MacFarlane, 2010: 117). It has a limited use, like an instruction manual with missing pages or some wise advice remembered in fragments. If we know for certain what the objective ‘ought’ counsels, we should listen; if we are unsure, attempted obedience isn’t safe.

¹¹ Nor can we restrict PL to the *most specific option that we objectively ought to do* (cf. Carlson, 1995: 102–03, 1999: 258 and Bykvist, 2002: 49 on “invariably optimal” actions). This restriction won’t recommend sending snowmobiles, since there is a more specific option that you objectively ought to do: *send snowmobiles to C*.

This all sounds nice, and it works great in Climbers. But of course we have to reject the three-step picture, along with PL, because of the new Miners puzzle, in which you know that you objectively ought to block a shaft and yet subjectively ought not to.

Now we are at the heart of the matter. Can we tweak PL—or strengthen WPL—to handle both Miners and Climbers? We need a plausible principle that can explain why you subjectively ought to send in the snowmobiles, but subjectively ought not to block a shaft.

Here is the best candidate we know of. Call two options *independent* when both can be chosen, either can be chosen without choosing the other, and both can be omitted. Then:

Value Link (VL)

SO(X) if the agent knows either that:

- (i) A_X has a higher objective value than any alternative to X; or
- (ii) for any Y that is independent of X, the objective value of X & Y is higher than the objective value of X & $\sim Y$.

Assuming that we objectively ought to maximize value, then VL will get the right results in Miners and Climbers. It doesn't say that you ought to block a shaft, since you don't know that blocking a shaft will save more lives (it might save none!). It *does* say that you ought to send the snowmobiles, since the value of sending them is higher no matter where the team is sent.¹²

Is it safe to assume that we objectively ought to maximize value? On some views, we objectively ought *not* to maximize value, when doing so would violate rights (Thomson, 1990; Kamm, 1996). But we can set aside this complication. We are just talking about simple rescue cases where saving lives is all that matters.

¹² Thanks to clause (i), VL also delivers the result that you ought to block Shaft A if you know that you objectively ought to do so (you know that it will save more lives).

We seem to have done it. VL captures the true parts of PL (the verdict about Climbers) without the problematic parts (the verdict about Miners).¹³

But VL, for all its virtues, is a significant retreat from PL: we are giving up the idea that objective ‘oughts’ are authoritative. Instead we seem to be ascribing all of the authority to objective *value*. The best we can say is that, sometimes, what we subjectively ought to do lines up with what we know we objectively ought to do—and sometimes it doesn’t. There is no independent bridge from objective to subjective ‘oughts’, no clear sense in which the objective ‘ought’ is like the ideal guide to which we have non-ideal access.

Is there a satisfying replacement for PL, linking subjective ‘oughts’ to knowledge of objective ‘oughts’? We doubt it.¹⁴ There doesn’t seem to be any interesting way in which objective ‘oughts’, as opposed to the objective values they track, are authoritative guides to rational action even when imperfectly known.

¹³ For the sake of argument, we will grant that VL is true. But we think it faces a counterexample (adapted from Spencer and Wells (2019: 35-6).) Consider a variant of Climbers. The set-up is the same as before, but this time, you know you have a long track record with these sorts of choices, and the record tells you that 90% of the time that you choose to send snowmobiles, you choose the empty cave, whereas you only choose the empty cave 50% of the time when you do *not* send the snowmobiles. You doubt that the track record is a fluke—and you’re right. (You have an unconscious desire to send the snowmobiles to majestic empty caverns; you unconsciously know which cave is empty; and this unconscious desire affects your choice of caves when, but only when, you send snowmobiles.) VL says that you should take the snowmobiles, even in this variant, since you know that the objective value of taking the snowmobiles to a cave exceeds the objective value of going to the same cave without the snow mobiles. But intuitively you subjectively ought *not* to take the snowmobiles. (To get around this, we could amend condition (ii) so that it only applies when X and Y are subjectively probabilistically independent.)

¹⁴ Two more failures, for good measure. First consider *PL**: if you know that you objectively ought to do X, and you know that any way of doing X is better than any way of not doing X, then you subjectively ought to do X. To its credit, *PL** doesn’t say that you subjectively ought to block a shaft in Miners. But *PL** fails to say that you subjectively ought to send the snowmobiles in Climbers. Second, what if we try to do something the contrastive clause (ii) in VL, but with objective ‘oughts’? We might say: *SO(X)* if the agent knows that they ought to do X rather than not X, no matter what else one does. But this fails in the Miners case. It’s false that one ought to block a shaft rather than not.

5. Conclusion: Objectivity Without Authority

We have argued that PL is false, and that there is no good way to capture the idea that objective ‘oughts’ are more authoritative than subjective ones. If this is right, we will have to rethink some familiar claims about what follows from knowledge of objective moral facts. We focus here on two claims: one links objective ‘oughts’ and advice; the other links objective and subjective reasons.

Some philosophers, most famously Thomson (1986: 179), argue that the objective ‘ought’ is the norm of advice (cf. Wodak, 2017: 259; Schroeder, 2018).

On those rare occasions on which someone conceives the idea of asking for my advice on a moral matter, I do not take my field work to be limited to a study of what he believes is the case: I take it to be incumbent on me to find out what is the case. (Thomson, 1986: 179)

If you ask me what you ought to do, I shouldn’t point to the best option by your benighted lights. I should try to tell you what you ought to do given the facts as best I can discover them.

This idea has to be tweaked, of course, in light of the Old Miners Puzzle. If I am just as ignorant as you about the miners’ location, the correct advice for me to give is *not* “Block Shaft A.” I should not even try to give the objectively correct advice; I shouldn’t pick a shaft at random and say “block it!” Surely the right advice is: “block neither.” The advisor should recommend, roughly, what subjectively ought to be done by the advisor’s own lights (cf. the envelope cases in Schroeder, 2018).

So a natural way to amend Thomson’s view—hopefully a bell-ringer—is to say that the objective ‘ought’ is the norm of advice *when the advisor knows what it says*. This is just like our PL, except that it links knowledge of objective ‘oughts’ to correct advice, instead of linking to subjective ‘oughts’. But now we face an analogue of the New Miners Puzzle. Suppose that Ada, your humble advisor, also knows that you objectively ought to block a shaft, but neither of you know which shaft to block. Clearly, ‘Block a shaft!’ is the wrong advice. Picking at random is an ill-advised gamble.

So our puzzle makes it hard to link objective ‘oughts’ to advice.¹⁵ To be sure, we could preserve a link if we restricted attention to fully specific options that the advisor knows ought objectively to be done. But this restriction, like RPL (and VL), represents a retreat. Just because an ‘ought’ concerns a non-specific option, that doesn’t make it a second-class ‘ought’. If non-specific objective ‘oughts’ aren’t advice-guiding, that undercuts the idea that known objective ‘oughts’ are intrinsically fit to guide advice; probably, we are just picking up on the fact that they track objective values, knowledge of which really *does* guide advice.

Finally, what about the link between objective and subjective reasons? We assume that ‘OO(Block Shaft A)’ entails that there is decisive objective reason to block Shaft A. We also believe that in the Miners case, there is decisive objective reason to *block a shaft*. (Why? One reason is monotonicity: the blank in “there is decisive reason to _____” appears to be upward monotone, as in ‘ought to _____’. But just as KO is plausible even if there are counterexamples to UM, it should be independently obvious that there is decisive *objective* reason to block a shaft.)

Here things get interesting. In the Miners case, we can suppose that *you know that there is decisive objective reason to block a shaft, but you still have decisive subjective reason not to* (you subjectively ought not to). You appear to face a bizarre deontic dilemma. The analogue here of PL is false. Knowing that you have decisive objective reason to do X does not entail having decisive subjective reasons to do X, nor does it entail that you subjectively ought to do X.¹⁶

¹⁵ The question also matters. If you ask which shaft to block, I should say: Shaft A. But if you ask *whether you should block a shaft*, it’s bad advice to say “yes” (if I can’t follow up with: “because you ought to block Shaft A.”)

¹⁶ It is an interesting question whether you “have” any decisive reasons to block a shaft. If knowing *that* there is an (objective) reason suffices as having the reason, then you *do* have decisive reasons to block a shaft. This seems false. By contrast, on Lord’s view, having a reason requires more than knowing of its presence:

Possession Enables Rational Routing

If *A* possesses *r* as a sufficient reason to ϕ , then there is a route that *A* can take to *ex post* rational ϕ -ing on the basis of *r*. (2018: 100)

What, in the end, is the significance of losing PL? The point of this principle, we suggested, was to capture the idea that the objective ‘ought’ is an ideal guide to which we have imperfect access—that we should be guided by it, and advise in accordance with it, whenever we know its pronouncements. But this picture just doesn’t make sense given the New Miners Puzzle. You have perfect access to the fact that you ought objectively to block a shaft, but that fact shouldn’t guide you; nor would a good advisor, in light of the facts, advise you to block a shaft (except *en passant* as she suggests blocking Shaft A). Similarly, you know that there is decisive objective reason to block a shaft, and yet there is only weak subjective reason to do so. The moral, we think, is that it is harder than anyone expected to guide oneself by partial knowledge of objective normative facts. The objective ‘ought’ is not so normative after all.¹⁷

Arguably, there is no reason on whose basis you can *ex post* rationally block a shaft. You can’t base your action on the fact that the 100 are in Shaft A, and the fact that the 100 are in *either* Shaft A or B doesn’t rationalize blocking a shaft for someone who has no idea which to block.

¹⁷ We would like to thank Al Hájek, Daniel Wodak, Arif Ahmed, Richard Yetter-Chappell, and Nathaniel Baron-Schmitt for discussion. Our sincere thanks also to the editors at *Philosophy and Phenomenological Research*, as well as to the anonymous referee who gave us such speedy and insightful comments.

References

- Ahmed, Arif, and Spencer, Jack (forthcoming). Objective value is always Newcombizable. *Mind*. doi: 10.1093/mind/fzz070
- Brandt, Richard (1959). *Ethical Theory*. Englewood Cliffs: Prentice Hall.
- Buchak, Lara (2013). *Risk and Rationality*. Oxford: Oxford University Press.
- Bykvist, Krister (2002). Alternative actions and the spirit of consequentialism. *Philosophical Studies*, 107, 45–68. doi: 10.1023/A:1013191909430
- Cariani, Fabrizio (2013). ‘Ought’ and resolution semantics. *Noûs*, 47, 534–558. doi: 10.1111/j.1468-0068.2011.00839.x
- Carlson, Erik (1995). *Consequentialism Reconsidered*. Dordrecht: Kluwer.
- Carlson, Erik (1997). Consequentialism, alternatives, and actualism. *Philosophical Studies*, 96, 253–268. doi: 10.1023/A:1004239306956
- Carr, Jennifer (2015). Subjective *ought*. *Ergo*, 2, 678–710. doi: 10.3998/ergo.12405314.0002.027
- Charlow, Nate (2013). What we know and what to do. *Synthese*, 190, 2291–2323. doi: 10.1007/s11229-011-9974-9
- Copley, B. (2006). What should ‘should’ mean? Ms. of a paper given at the Workshop “Language Under Uncertainty: Modals, Evidentials, and Conditionals,” Kyoto University, January, 2005.
- Dorsey, Dale (2012). Objective morality, subjective morality, and the explanatory question. *Journal of Ethics and Social Philosophy*, 6, 1–24. doi: 10.26556/jesp.v6i3.65
- Dowell, Janice L. (2012). Contextualist solutions to three puzzles about practical conditionals. In R. Shafer-Landau (Ed.), *Oxford Studies in Metaethics, Volume 7* (pp. 271–303). Oxford: Oxford University Press.

- Ewing, A.C. (1947). *The Definition of Good*. New York: The MacMillan Company.
- Goldman, Holly (1978). Doing the best one can. In A. Goldman and J. Kim (Eds.), *Values and Morals* (185–214). D. Reidel.
- Grant, L. and Phillips-Brown, Milo (2019). Getting what you want. *Philosophy and Phenomenological Research*. Early online. doi: 10.1007/s11098-019-01285-1
- Horn, L. R. (1972). *On the Semantic Properties of the Logical Operators in English*. Ph.D. thesis, UCLA, CA.
- Hurka, Thomas (2014). *British Ethical Theorists from Sidgwick to Ewing*. Oxford: Oxford University Press.
- Jackson, Frank (1985). On the semantics and logic of obligation. *Mind*, 94, 177–195.
- Jackson, Frank (1986). A probabilistic approach to moral responsibility. In R. B. Marcus, G. Dorn, and P. Weingartner (Eds.), *Logic, Methodology, and Philosophy of Science VII* (pp. 351–65). Amsterdam: North-Holland.
- Jackson, Frank (1991). Decision-theoretic consequentialism and the nearest and dearest objection. *Ethics*, 101, 461–482.
- Jackson, Frank and Pargetter, Robert (1986). Oughts, pptions, and actualism. *Philosophical Review*, 95, 233–255.
- Kamm, Frances Myrna (1996). *Morality, Mortality, Volume II: Rights, Duties, and Status*. Oxford: Oxford University Press.
- Kamp, Hans (1973). Free choice permission. *Proceedings of the Aristotelian Society*, 74, 57–74.
- Kolodny, Niko and MacFarlane, John (2010). Ifs and oughts. *Journal of Philosophy*, 103, 115–143. doi: 10.5840/jphil2010107310

- Kratzer, Angelika (1977). What ‘must’ and ‘can’ must and can mean. *Linguistics and Philosophy*, 1, 337–355.
- Kratzer, Angelika (1981). The notional category of modality. In H. Eikmeyer & H. Rieser (Eds.), *Words, Worlds, and Contexts: New Approaches in Word Semantics. Research in Text Theory* 6 (pp. 38–74). Berlin: de Gruyter.
- Lassiter, Daniel (2011). *Measurement and Modality: The Scalar Basis of Modal Semantics*. New York University PhD thesis.
- Littlejohn, Clayton (2018). Being more realistic about reasons: On rationality and reasons perspectivism. *Philosophy and Phenomenological Research*, 99, 605–27.
- Lord, Errol (2018). *The Importance of Being Rational*. Oxford: Oxford University Press.
- Nozick, Robert (1969). Newcomb’s problem and two principles of choice. In N. Rescher (ed.), *Essays in Honor of Carl G. Hempel* (pp. 114–146). Dordrecht: Reidel.
- Parfit, Derek (ms.). What we together do. Unpublished.
- Parfit, Derek (1984). *Reasons and Persons*. Oxford: Oxford University Press.
- Parfit, Derek (2011). *On What Matters, Volume 1*. Oxford: Oxford University Press.
- Portmore, Douglas W. (2019). *Opting for the Best: Oughts and Options*. Oxford: Oxford University Press.
- Prichard, H.A. (2002). Duty and ignorance of fact. J. MacAdam (ed.), *Moral Writings*. Originally given in 1932 as the Annual Philosophical Lecture, Henriette Hertz Trust, British Academy.
- Regan, Donald (1980). *Utilitarianism and Cooperation*. Oxford: Oxford University Press.
- Ross, Alf (1941). Imperatives and logic. *Theoria*, 7, 53–71. Reprinted with only minor editorial changes as Ross 1944.
- Ross, Alf (1944). Imperatives and logic. *Philosophy of Science*, 11, 30–46.

- Ross, W.D. (1939). *The Foundations of Ethics*. Oxford: Oxford University Press.
- Scanlon, Thomas (2008). *Moral Dimensions: Permissibility, Meaning, and Blame*. Cambridge: Harvard University Press.
- Schroeder, Mark (2008). Having reasons. *Philosophical Studies*, 139, 57–71. doi: 10.1007/s11098-007-9102-3
- Schroeder, Mark (2018). Getting perspective on objective reasons. *Ethics*, 128, 289–319. doi: 10.1086/694270
- Sepielli, Andrew (2018). Subjective and objective reasons. In D. Star (Ed.), *The Oxford Handbook of Reasons and Normativity* (pp. 784–99). Oxford: Oxford University Press.
- Spencer, Jack and Wells, Ian (2019). Why take both boxes? *Philosophy and Phenomenological Research* 99, 27–48. doi: 10.1111/phpr.12466
- Thomson, Judith Jarvis (1986). Imposing risks. In W. Parent (Ed.), *Rights, Restitution, and Risk* (pp. 173–191). Cambridge: Harvard University Press.
- Thomson, Judith Jarvis (1990). *The Realm of Rights*. Cambridge: Harvard University Press.
- Thomson, Judith Jarvis (2008). *Normativity*. Chicago: Open Court.
- von Fintel, Kai (2012). The best we can (expect to) get? challenges to the classic semantics for deontic modals. Paper for a session on Deontic Modals at the Central Division Meeting of the American Philosophical Association, February 17, 2012.
- Wedgwood, Ralph (2016). Objective and subjective ‘ought’. In N. Charlow and M. Chrisman (Eds.), *Deontic Modality* (pp. 143–168). Oxford: Oxford University Press.
- Wodak, Daniel (2017). Can objectivists account for subjective reasons? *Journal of Ethics and Social Philosophy*, 12, 259–279. doi: 10.26556/jesp.v12i3.246

Zimmerman, Michael (2008). *Living With Uncertainty*. Cambridge: Cambridge University Press.