Blameworthiness Implies 'Ought not'

Simon-Pierre Chevarie-Cossette[1]

Abstract

Here is a crucial principle for debates about moral luck, responsibility, and free will: a subject is blameworthy for an act only if, in acting, she did what she ought not to have done. That is, 'blameworthiness' implies 'ought not' (BION). There are some good reasons to accept BION, but whether we accept it mainly depends on complex questions about the objectivity of ought and the subjectivity of blameworthiness. This paper offers an exploratory defence of BION: it gives three *prima facie* reasons to accept it, provides a plausible interpretation of it, and shows how holding out against objections can yield fruitful lessons. Five objections to BION are considered: the objection from conscience, from reasons, from suberogation, from objectivity, and from excuses. Their main problem is to either over-subjectify blameworthiness or to over-objectify obligations. To accept BION, we must occupy a desirable middle ground.

Keywords Blameworthiness · Ought · Defences · Excuses · Justifications · Objectivism · Perspectivism · 'ought' implies 'can'

1 Introduction

Pierre-Luc works in a bar. He tends to be late, which indisposes his co-workers. One day, he arrives at 19:00, convinced that he is late again. Yet he had no shift that day—which he could have realised if he had thought about it for a minute. What should we say about Pierre-Luc? Maybe he did something wrong—such as showing carelessness about being late—and is (mildly) blameworthy for it. Or maybe he did nothing wrong, really, and so is not blameworthy. Depending on how we refine the case and what best moral theory is on offer, we might go one way or the other. Yet, some phi-

✉ Simon-Pierre Chevarie-Cossette
   spccossette@gmail.com; simon-pierre.chevarie-cossette@unine.ch

1 University of Neuchâtel, Neuchâtel, Switzerland

Springer

losophers have argued that there is a third way: say that Pierre-Luc is blameworthy in spite of having done everything he had to do. I disagree.

To put things positively, and abstractly, I accept the principle that 'blameworthiness' implies 'ought not' or 'obligation to refrain' (I shall equate the two):

BION   If S is blameworthy for having Φed, S ought not to have Φed (where Φ is an act or an omission).

BION is a tricky principle to defend. On the one hand, the arguments which support it do not seem decisive. On the other hand, to rebut counterexamples to it, we must take a stance on several complex issues, issues which are much bigger than BION itself. For instance, we need to assess whether what one ought to do depends on one's evidence and whether ignorance justifies or merely excuses. This is why this paper merely offers an *exploratory defence*: it presents the most plausible interpretation of the principle, and flips objections to yield theoretical fruits. It shows that we can keep obligations and blameworthiness together if we neither over-objectify the first (that is, make obligations entirely independent of the subject's evidence) nor over-subjectify the second (that is, make blameworthiness too dependent on the subject's reasons or beliefs).

Before we turn to discussing the truth of BION, we should have a clear idea of its importance. I believe it is threefold. First, it is relevant to the free will and responsibility debate. It is in fact crucial to one of its most debated questions: does blameworthiness require leeway? Differently put, if I am blameworthy for an act, does this imply that I could have refrained from so acting? Apart from fifty years of literature on Harry Frankfurt's counterexamples to a positive answer (which I shall not discuss here), we have one argument[1]: blameworthiness implies obligation, but obligation implies leeway, so blameworthiness implies leeway. Or, more formally:

BION   If S is blameworthy for having Φed, S ought not to have Φed (where Φ is an act or an omission).
OIC   If S ought not to have Φed, S could have refrained from Φing.
BIC   Is S is blameworthy for having Φed, S could have refrained from Φing.

So, if BION is true, we have a strong reason to believe that blameworthiness requires leeway, given the appeal of 'ought' implies 'can' (OIC) and of its above iteration—'ought not' implies 'can refrain'.[2]

---

[1] See Widerker (1991), Copp (2003, 2008), and Nelkin (2011, Chap. 5). BIC is an interpretation of PAP, the principle of alternate possibilities.

[2] Several philosophers (see e.g. Haji, 1993; Vranas, 2007, 192–96) have argued that the argument equivocates on 'ought': whereas BION is only plausible with a subjective 'ought', OIC—or so the critics alleged—is only plausible with an objective 'ought'. I reject this. These criticisms operate with a dubious notion of subjective 'ought', which is just a belief about an objective 'ought'. Once we employ a *perspectival* 'ought', OIC can be read with it. Note also that 'can refrain from Φing' can be substitued with the less elegant 'can not-Φ' to avoid problems with people fulfilling obligations while they sleep (while asleep, one might fulfill an obligation not to lie while not being strictly able to refrain from lying because refraining requires being awake).

Second, BION touches on moral luck. Suppose that Audrey and Barry are each telling a compromising story about a friend. They are both convinced that they have promised to their respective friend to be discrete about it. But contrary to Barry, Audrey misremembers: she did not in fact promise anything of the sort. Are they equally blameworthy? Those who dislike moral luck—and so who think that relevantly similar agents cannot differ in blameworthiness—will want to say yes.[3] But defenders of BION should say no. Although the principle is not explicit on the matter, its defender are likely to say that what we are blameworthy *for* is always the infringement of an obligation. Now, clearly, Audrey and Barry did not infringe the same obligations: while they perhaps were both indiscrete, only Barry broke a promise. The defender of BION should then accept that Barry is blameworthy for things for which Audrey is not. Given the internal similarity of the two, this amounts to admitting the existence of moral luck.

Third, BION matters for *the conceptual unity of moral responsibility*, the idea that the different senses of the word are linked. This oft-cited passage of *Punishment and Responsibility* reminds us of this polysemy:

> As captain of the ship, X was responsible$_1$ for the safety of his passengers and crew. But on his last voyage he got drunk every night and was responsible$_2$ for the loss of the ship with all aboard. It was rumoured that he was insane, but the doctors considered that he was responsible$_3$ for his actions. Throughout the voyage he behaved quite irresponsibly, and various incidents in his career showed that he was not a responsible$_4$ person. (Hart, 1968, p. 211)

Even if we only focus on *moral* responsibility, this paragraph presents us with four senses of the term. The captain had responsibilities$_1$, that is prospective obligations (future-oriented).[4] The captain was not responsible$_4$, that is he lacked the *virtue* of responsibility: the tendency to discharge responsibilities in the first sense. Had the captain been insane, he would have lacked responsibility$_3$, which is a *capacity* to bear responsibilities. He was also responsible$_2$ for his conduct, which is a retrospective obligation (past-oriented) or liability to *answer for it*, i.e. to explain his conduct, and possibly be liable to apologise, be blamed, punished or pay compensation.[5] But what links retrospective responsibility with the other senses? If BION is true, we may give a reasonable answer: to hold one retrospectively responsible is in part to claim that one has failed to discharge a prospective responsibility. If BION were false, though, it would be difficult to preserve a conceptual link between retrospective responsibility and the other senses of responsibility. *A fortiori*, the unity of moral responsibility would be jeopardised, contra Ronald Dworkin (2011, 461–63), who argued that we could not excise one sense from our moral system without endangering the others.

---

[3] See e.g. Wolf (2001).

[4] Hart calls them 'role obligations' because we'd only use the word 'obligations' for responsibilities which were assigned in virtue of a role.

[5] Hart here speaks of *liability*, but there is a risk of confusing the liability to answer for an act with the liability to pay for it. I follow Anthony Duff (2007, Chap. 1) and others in separating the two.

Now that we have some idea of the importance of the putative truth of BION, we can turn to whether it is in fact true.

2 Three cheers for 'blameworthiness' implies 'ought not'

Why accept BION? First, it explains the paradoxical character of a variety of utterances. Consider the following:

> Carl did not help Dorothea when she was in need and now she blames him. 'It was not my responsibility.' He rejoins. 'So what?' says Dorothea, 'I blame you for it, even though you did everything that you had to do'.

Dorothea's response ('so what?') and following utterance is utterly paradoxical. A simple explanation for this is that she is at once casting blame on Carl and committing to his blamelessness (she says that he did what he had to do, which, by BION, implies that he is not blameworthy).[6] Her assertion is self-undermining, which is hard to explain without BION. BION equally explains the paradoxical character of the following assertions:

> 'You did not have to do the dishes, but I blame you for not having done it.'
> 'It was not your responsibility to help me, but I hold you responsible for it.'
> 'Oh, you did everything you had to do, but you're blameworthy all the same.'

The evidence in favour of BION here is not merely linguistic. The idea is not just that our sentences sound strange when so strung. Rather, this strangeness is a symptom of a misunderstanding from the speaker about our moral practice (and its concept of blame). If one recognises that I have done all I had to do, blaming me is off the table according to our practice.[7]

Is our practice faulty? Well, there is good reason why BION is embedded in it, both interpersonal and legal. That's my second point for BION. It has to do with justifications and excuses. Presenting either a justification or an excuse is a standard defence, either a moral or a legal one. That is, justifications and excuses are ways to avert some adverse reaction such as blame or punishment. In presenting a justification, I explain that I didn't do the wrong thing (that I did what I ought to have done) and so that I am blameless. In presenting an excuse, I explain that even though I did the wrong thing, I did not have an ill will or I lacked full control over my behaviour, and so I am blameless.[8] Now, if BION were false, a justification would not be a good defence. In fact, if BION were false, justifications could be entirely beside the point

---

[6] Might there be other objects for blame than obligation infringements? Kiesewetter (2017, 31–39) gives good reason to think that the answer is negative. His defence of the principle that 'criticisability' implies 'reasons' complements my defence of BION.

[7] Darwall (2006, 26–27) Skorupski (1999, 42–43), Williams (1995, 40–41), and Kiesewetter (2017, 29) have made similar remarks about reasons: we cannot blame or criticise someone for having Φed and yet maintain that there was no reason against Φing.

[8] This is a gloss on Strawson's (1974) and Austin's (1956) seminal presentations.

in a trial, or in a conversation whose aim is to determine who is to be blamed for a harm. In both cases, one should skip justifications and get down to excuses already, given that only *they* can rule out blameworthiness. But it sounds like a terrible idea.

This raises a third but connected point. If BION is false, then most of the time, when people have a justification, they *also* have an excuse. Why? Imagine that an agent was justified and blameless in acting. Since, hypothetically, BION is false, the justification itself does not render this individual blameless, nor does it explain her blamelessness. Something else is needed: a factor that explains why the action was blameless. But, presumably, that's just an excuse. This is strange: excuses are things we don't particularly like having; there is even something derogatory about being excused (see Baron, 2005). Excuses are literally *ex-causa*: they take us away from a charge. They are not the sort of thing that we associate with the fully justified.

Perhaps in part to avoid this undesirable consequence, a denier of BION like Michael Zimmerman (1997, 2008) puts things differently. Instead of assigning both excuses and justifications to the blameless justified, he assigns something new, *accuses*, only to those who are justified and yet not excused. So at first it seems that he can, like the friends of BION, maintain that excuses only exist where there is wrongdoing, that is where the behaviour is unjustified. The problem is that this solution amounts to a verbal stipulation. Zimmerman must stipulate that excuses don't apply to justified behaviour, but he can't explain why. By contrast, friends of BION can explain why excuses don't apply to justified behaviour: justifications are exculpatory (they block blameworthiness), and so if an action is justified, there is no exculpatory work left for excuses. This answer is unavailable to Zimmerman, since he cannot recognise that justifications exculpate. He must merely insist on the inexplicable fact that excuses only apply to wrongdoing.

3 What 'ought'?

With these three prima facie reasons to accept BION, I turn to interpreting it, before I address objections in § 4 and § 5. For BION to be plausible, 'ought' cannot be read objectively. What does this mean? Take:

> HOSPITAL: Eliot's child has just started to show symptoms of what he recognises is a dreadful disease. He believes—and this is what his evidence supports—that the only hospital in the area equipped to cure it is Traditional Hospital, which is north of town. In fact, however, the only hospital in the area equipped to cure this disease is Modern Hospital, which is south of town. There isn't time for Eliot to get to both hospitals before his child starts suffering horribly. What ought he to do? (inspired from Thomson, 2008, 187–88)

Objectivists traditionally judge that Eliot ought to go to Modern. At any rate, *objectivism* (e.g. Thomson, 1990; Graham, 2010; Fassio, 2022) says that all facts may be relevant to determining what an agent ought to do—not just facts accessible to the agent. For instance, the fact that *Modern, though not Traditional, has the cure* may determine what Eliot ought to do.

If objectivists are right, and if there is no other relevant sense of 'ought', then BION is false. Say Eliot goes to Modern on a whim. He'd have satisfied his obligation to go there and yet would be blameworthy for that. From this sort of cases featuring ignorance (see e.g. Capes, 2012, 428–31 for another such case), some conclude that BION is false. (We will discuss a more nuanced objectivist treatment of this case in § 5.)

Yet, BION is defensible if we read 'ought' differently, i.e. as relativised to the agent's evidence:

BION$_r$ If S is blameworthy for having Φed, S ought not to have Φed, *given S's evidence.*

This 'ought' is mostly determined by the agent's evidence. I say 'mostly' because we might accept that external constraints, such as what the agent *can*'t do, matter of oughts (since 'ought' implies 'can'). In HOSPITAL, Eliot clearly ought to go to Traditional given his evidence. So, if he goes to Modern, he is blameworthy for having infringed his obligation. BION$_r$ stands.

Perspectivists (e.g. Zimmerman, 2008; Kiesewetter, 2017) think that this relativised 'ought' is the main moral one. But we could instead admit that there are several equally morally important 'oughts', corresponding to different relativisation bases (see e.g. Olsen, 2017), and still accept BION$_r$. Which one do we use when we are not specific? For *contextualists* (e.g. Henning, 2014; Björnsson and Finlay 2010), the meaning of 'ought' varies depending on the asserter and standards (norms, ends, etc.). This is also compatible with BION$_r$ since the task of determining blameworthiness calls for a perspectivist reading of 'ought' and 'ought' sentences. When we blame someone for having done something, it is what they ought to have done relative to their evidence, not relative to ours, that counts.[9] For *relativists* (Kolodny and MacFarlane, n.d.), the meaning of 'ought' (and ought-sentences) does but a single 'ought' sentence might be true relative to one and false relative to another. Relativists can also accept that blameworthiness implies 'ought not' relative to the agent's evidence (perhaps by saying that the truth that matters is this one). For our purpose, there is no need to commit to one of *perspectivism, contextualism, or relativism.* Accepting the disjunction will do, although assuming perspectivism will simplify our discussion.

Rejecting objectivism does not yet give us BION$_r$. In fact, some perspectivists have explicitly attacked it. Let's start with these objections, before we move to objections which a perspectivist could not accept.

---

[9] Björnsson and Finaly claim that 'contextualists need not construe normative judgements as inflexibly relativized always to the speaker's information' (2010, 12).

4 Perspectivist-friendly objections

4.1 The objection from conscience

Sometimes we have a false belief about what we ought to do. In some of these cases, we think that we are infringing our obligations, but in fact we are not. Differently put, we are acting against our conscience, yet we are acting rightly. In such cases, says Zimmerman (2008, 193–200), we do what we ought and yet we are blameworthy for it; we have an *accuse*. So BION$_r$ is false.

In the form of a deductive argument:

(1) It is possible that S Φs while falsely believing that she ought not to Φ.
(2) But if S Φs while (falsely) believing that she ought not to Φ, she is blameworthy for Φing.
(3) *Therefore*, it is possible that S is blameworthy for Φing even though it's false that S ought not to Φ.

Zimmerman spends some time defending (1), by arguing, rightly in my opinion, that mere belief that an option is wrong or impermissible is insufficient to make this option wrong or impermissible. This is compatible with perspectivism, for to say that what we ought to do depends mostly on our evidence leaves it open that we are mistaken about what we ought to do when we disrespect our evidence. As to (2), Zimmerman explains how he can respond to putative counterexamples.

Capes (2012, 424–28) has shown that (2) is false. If someone believes falsely that she is doing something wrong in lying to the Nazi to protect a fugitive, but sees no other way around it, she is not blameworthy for it. In a word, premise (2) deals poorly with putative cases of moral dilemma. But this response is unlikely to put Zimmerman's worries to rest. For his premise does not need to be so strong for there to be cases where a subject's false beliefs about her obligation not to act makes her blameworthy for so acting. And so (2) can be considerably refined to avoid Capes' counterexample. For instance, Zimmerman's argument could be adjusted as follows:

(1*) It is possible that S Φs while falsely believing that she ought not to Φ, *and while truly believing that there is another permissible option available.*
(2*) But if S Φs while (falsely) believing that she ought not Φ *and while truly believing that there is another permissible option available*, she is blameworthy for Φing.
(3) *Therefore*, it is possible that S is blameworthy for Φing even though it's false that S ought not to Φ.

(1*) still seems true, and Capes' case is not a counterexample to (2*).

What we need then, if we want to respond to Zimmerman's argument, is a way to systematise our treatment of putatively problematic cases where a false belief about what one ought to do may ground one's blameworthiness. My rough approach is to distinguish between two ways of having a false belief regarding our conduct. Take the famous case of Huckleberry Finn, who believes that he is wrong in not return-

ing Jim to the slave master (see Arpaly, 2002). Compare this case with the equally famous case of the husband who feeds to his wife what he believes to be poison but is in fact a cure (see Thomson, 1990). The first case mainly features moral ignorance whereas the second mainly features factual ignorance. If we are interested in cases of moral ignorance, I see no reason to grant that they feature blameworthiness. Huck Finn is not blameworthy at all for not returning Jim in spite of acting against his conscience.[10] If we are interested in cases of factual ignorance, I see no reason to grant that they might not feature an infringed 'ought'. The husband ought to refrain from giving what he falsely believes to be the poison. (Remember that we are working with the perspectival 'ought'.) That's only a rough diagnostic—I shall nuance it soon.

Of course, the distinction between moral and factual ignorance is tricky to use, since almost no question is purely moral and since non-moral ignorance is often the cause of moral ignorance. But a coarse distinction will do. By 'factual ignorance', I mean ignorance about properties of actions or of their consequences expressed in non-moral terms. This includes ignorance about whether we will cause pain, whether we will poison or cure someone or whether our killing will be intentional or not. By 'moral ignorance', I mean ignorance about fundamental axiological truths (related to value), deontic truths (related to duty, the right, etc.), aretaic truths (related to virtues and vices), and hypological truths (related to moral responsibility, culpability, etc.). This includes ignorance about whether late-term abortion is permissible, cunningness is a virtue, or sadistic pain can be good.

Often, it will be difficult to say which kind of ignorance is at play. Suppose that X falsely believes that it is wrong to hit Y—this has to do with pain—but does it anyway. Without more information, we might have two different stories in mind. First, we might think that X's belief mainly stems from a false belief that Y would feel pain when hit. This would make X's ignorance factual. Second, we might think that X's belief depends on a false belief that Y has a special moral status (perhaps because Y can feel pain). This would make X's ignorance moral. This difficulty may explain why the defender of BION$_r$ may hesitate about some putative counterexamples.

But the hardest counterexamples follow a different recipe. In cases of factual ignorance, blameworthiness seems undeniable: for instance, there is no doubt that the murderous husband is blameworthy. Here, the difficulty to mount a counterexample to BION$_r$ is to make sure that there is no corresponding obligation. The best way to secure this is for the factual ignorance to be unreasonable. (By contrast, if someone has good reason to think that there is poison in the glass, it is clear that they ought not to give it to others, given their evidence.) And in cases of moral ignorance, the absence of obligation seems undeniable: there is little doubt that Huck Finn did what he ought to have done. Here, the difficulty to mount a counterexample to BION$_r$ is to ensure that the agent is blameworthy. The best way to secure this is for the moral ignorance to be reasonable. (By contrast, if someone believes that scratching their head is impermissible, it's hard to find them blameworthy for scratching their head).

---

[10] This defence of BION$_r$ might remind the reader of the *asymmetry thesis*, according to which moral ignorance never excuses. Just like it is plausible that the morally false belief that Φing was right cannot *excuse*, the morally false belief that Φing was wrong cannot *accuse* (cannot make blameworthy). See Alvarez & Littlejohn (2017). The two theses are independent of course.

So we have two kinds of counterexamples to BION to consider: cases of unreasonable factual ignorance and cases of reasonable moral ignorance.

Let's start with unreasonable factual ignorance. François is hiking in an abandoned area, and he falsely and unreasonably comes to believe that he ought not to push a pebble off a cliff. If he thought about it for an instant, he would come to know that (1) there is no way that someone could be standing in the wrong place since there is no possible access to that area; and that (2) the pebble is so small that even if there were someone, it could not even scratch them. Yet François pushes the pebble. In such a case, should we say that he ought not to push the pebble, or, alternatively, should we say that he is not blameworthy? Since François' evidence does not at all support the conclusion that it would be dangerous to push the pebble, I'd say he is not blameworthy for pushing it. I'm even tempted to say that he is simply not blameworthy for anything in this situation.[11] But if we insist on saying that he is, we can easily identify things which he ought not to have done such as forming beliefs irrationally or having acquired and cultivating an imprudent and careless character. We can perhaps also blame him for what he is (irrational, careless) rather than for what he does, which does not conflict with $BION_r$ (since it concerns conduct).

Second, consider reasonable moral ignorance. For the sake of argument, assume that we have no obligation to give our working organs to strangers. Now, Gustav and Heidi have carefully considered all available arguments in favour of organ donation; Gustav concludes (correctly) that it is supererogatory; Heidi concludes (incorrectly) that it is obligatory. But neither gives their kidney. I see no good reason for thinking that whereas Gustav is not blameworthy for not giving his kidney, Heidi is.

On the contrary, I see strong reasons for opposing this asymmetry. When we blame someone for Φing, we expect them (ceteris paribus) to alter their course of action so as not to Φ. So if we blame Heidi for not giving her kidney, we expect her to donate it. But this expectation would be unfitting since, as we hypothesised, donating one's kidney is supererogatory. In fact, it would be remarkably unfair to try to influence Heidi, but not Gustav, in this manner. We would be trying to take advantage of someone's honest error of judgement. So, it would be unfitting for us (or anyone) to blame Heidi. Given the circumstances, we should conclude that Heidi is blameless for not donating her kidney.[12] She might be blameworthy for being akratic (something she ought not to be); and *this* is something that it might be fitting to expect her to change.[13] Alternatively, Heidi might simply *not* be blameworthy for anything in this situation, although we might insist on the disvalue of her incoherence.

---

[11] Importantly for the case, François has no intention of hurting anyone with the rock: if he did, he would have *tried to hurt someone*, which he ought not do—see Kiesewetter (2017, 33).

[12] Could we say that blaming Heidi would be unjustified but fitting? No. What makes it problematic to blame Heidi has nothing to do with factors which pull justifiability and fittingness apart, such as complicity and hypocrisy.

[13] See Kolodny (2005) for a defence of the view that akrasia necessarily involves a violation of an 'ought'.

4.2 The objection from reasons

A second perspectivist-friendly objection, due to Tim Scanlon (e.g. 2008, 125; see also Graham, 2010, p. 94), can be answered similarly. Scanlon thinks that one might be blameworthy for Φing in spite of being obliged to Φ if one Φes for the wrong reasons. For instance, if Isidore ought to save someone by amputating a gangrenous leg but does it for his own sadistic pleasure—he would not have helped otherwise—he is blameworthy *for amputating the leg*.

That's implausible. Isidore would be much more blameworthy if he had amputated a leg out of sadism, without knowing that it was gangrenous. This indicates that what he is blameworthy *for* is not simply *amputating the leg* (see e.g. Nelkin, 2011, 104–5; Kiesewetter, 2017, 30–31). Now, Scanlon insists on the fact that we don't have the obligation to act for specific reasons. The reasons for which we act, he says, may be relevant to questions of blameworthiness but not to questions of permissibility and obligation. Even if we grant this (which we don't have to—see McMahan (2009, 354–56), we nevertheless have the obligation not to cultivate and manifest certain vices, sadism being one of them. So we can insist on the fact that Isidore is blameworthy for something he ought not to have done, i.e. manifest sadism (by taking pleasure in what he was doing). But Isidore will certainly deserve much less blame than if he had been blameworthy *for amputating a leg.*

How can we be so sure that Isidore is not also blameworthy *for amputating the leg*? In accordance with our previous argument (about organ donation), blameworthy conduct is conduct that it would be ceteris paribus fitting to expect the agent to change. Whereas it would be unfitting to expect Isidore not to perform the amputation (since he ought to perform it), it would be fitting to expect him not to manifest sadism.

I have shown that we can defend BION$_r$ against Zimmerman's and Scanlon's objections. This serves an important lesson about blameworthiness: we must not over-subjectify it in comparison with 'ought'. True, one does not become obliged not to Φ just by believing that Φing is wrong; but, equally, one does not become blameworthy for Φing just by believing that Φing is wrong while Φing. In a similar vein, one is not prohibited from Φing just because of the reasons for which one Φes. But, equally, one is not blameworthy for Φing just because of the reasons for which one Φes. What one is blameworthy for is conduct that is fitting ceteris paribus to expect the agent to change.

4.3 The objection from suberogation

The last perspectivism-friendly objection involves *suberogatory* acts, acts which respect duty but which are still somehow morally bad. Take the following:

> [I]n boarding a train the person who is first gets first choice of seats. But suppose that the train is almost full, and a couple wish to sit together, and there is only one place where there are two seats together. If the person ahead of them takes one of those seats, when he could have taken another less convenient seat,

and knowing that the two behind him wanted to sit together, then he has done something blameworthy. [...] (Driver, 1992, 286–87).

The idea is that agents who perform suberogatory acts have done everything they morally ought to do and yet are blameworthy. Some take the existence of suberogatory acts to refute $BION_r$ or similar principles (Haji, 2012, 181–85).

I agree with the view that the man's behaviour is bad. But why are Haji and Driver inclined to say that it is both permissible and blameworthy? To simplify, let's say that the commuter's act is a one-off (if not, we can say that what is blameworthy about the case is that he is *yet again doing the bare minimum*). To take this case as a counterexample to $BION_r$ seems to either excessively restrict 'ought' or to excessively expand 'blameworthy'.

First, why can't we say that the commuter ought to give his place? Perhaps because we think that 'oughts' are correlated with rights. I agree that the couple doesn't have a *right* to the seat: if they ask the man to move and he refuses, there is not much that they can appropriately do to make him move apart perhaps from guilt-tripping him (such as asking for a train officer or pushing him). But the 'ought' at issue in $BION_r$ is not correlated with rights, since rights are not dependant on evidence.

Second, why do we have to say that the commuter is *blameworthy* in this case? Depending on the details (the longer the train journey, the more problematic it will be), the act may simply be *criticisable*, which might not imply 'ought' (it might imply 'should'). I am not saying that we *must* distinguish criticisability and blameworthiness. But the impetus which leads the moral philosopher to distinguish actions which are *not prohibited yet still morally bad* (*suberogatory*) might also lead her to distinguish actions which don't deserve blame *yet still deserve some bad moral assessment*. If we want to be refined about deontic notions, we should be refined about hypological notions too.

5 Perspectivist-unfriendly objections

We have tied blameworthiness to obligations with a plausible interpretation of BION, where 'ought' is somehow relative to the agent's evidence. The next two objections attack this interpretation, one directly and one indirectly.

5.1 The objection from objectivity

Take again Eliot from HOSPITAL: if he goes to Modern (thus defying his evidence), he would do what he ought to do, but would be blameworthy for it. So 'blameworthiness' does not imply 'ought not'.

How can objectivists claim that Eliot ought to go to Modern in HOSPITAL? Well, he has a duty to protect his child, and going to Modern is the only way to protect her. It is also what an informed benevolent advisor would tell him to do. If Eliot asked, 'what ought I do?', the informed advisor should not ask Eliot for his evidence—he should tell him the truth: that he ought to go to Modern. As Thomson puts it, the 'ought of directives' must then be objective. In fact, whether there is an adviser or not, when an

agent discovers better information relevant to what she ought to do, it seems that she can progress—and say, for instance, 'I see that I was wrong'.

There are different ways to deal with the better information problem. I shall briefly mention three. Suppose Eliot's current evidence favours Traditional, and he says to himself 'I ought to go to Traditional'. We could follow Gibbons (2010, 355–56) and insist on the self-fulfilling character of advice. It would be immoral, though truthful, for an adviser to tell Eliot that he ought to go to Traditional (or for Eliot to have said it in the first place). But it would be moral *and truthful* for an adviser to tell Eliot that he ought to go to Modern because it's a self-fulfilling claim: by communicating the information, the adviser makes it true that Eliot now must go to Modern because he thereby transmits evidence about where the cure is. Note that if the advisor sent Eliot a letter but that it never reached him, it would turn out to be false, though not a lie. Contextualists like Björnson and Finlay (2010) don't have to say that the advice is false until it's received. When an informed adviser tells Eliot that he ought to go to Modern (or when Eliot acquires the new evidence by himself), it does not strictly contradict his previous claim. It's true that Eliot ought to go to Traditional *given his previous body of evidence*. But that's now irrelevant since his body of evidence has been updated. Finally, for a perspectivist like Kiesewetter (2017, 211–23), when an informed adviser tells Eliot that he ought to go to Modern (or when he acquires the new evidence by himself), it contradicts his previous claim. But that's because the relevant evidence to where he ought to go has always included everything he should have acquired by the time of the action.[14]

So, there are reasonable answers to the objectivist challenge. But there's also an argument *against* objectivism, which I want to develop. Consider the well-known:

> MINE: There was a disaster in the quarry, and 100 miners are trapped in Shaft A; the nearby Shaft B is empty. Jane, who is in charge of the miners' safety, knows that, if she does nothing, the shafts will partly flood and 10 miners will be severely injured. She also knows that, if she blocks the shaft where the miners are she will save all 100; and if she blocks the empty shaft, the other will totally flood, drowning all 100. But her evidence doesn't tell her where the miners are; for her, it's a 50/50 guess. The miners are in fact in shaft A. What ought Jane to do?[15]

Perspectivists (but also relativists, contextualists, and pluralists) plausibly insist that Jane ought to take the safe option, which is to refrain from blocking a shaft. And they

---

[14] It might seem like Kiesewetter's solution is the best, since it alone can 'explain why agents in [seeking new evidence] understand themselves as finding out what they ought to do rather than changing the truth about what they ought to do' (2017, 220). But we must be careful. Suppose that on a trip to rural Hungary, and having to cure my friend, a monolingual local tells me something complex, e.g. that I ought to go to Modern, but using a detour that seems to go to Traditional. Given that what she says is inaccessible to me, she speaks falsely. But if, by sheer luck, a bilingual person stops to help, this makes the claim true. If we have to say this here, why not follow Gibbons in the first place?

[15] The case originates in Regan (1980, 265, n.1). A structurally identical case can be found in Kiesewetter (2017, 202), Zimmerman (2008, 17–18), and Lord (2017, 1136–37).

have traditionally taken objectivists to be committed to the false view that Jane ought to block Shaft A.

Just like non-objectivism comes in more or less subjective varieties, objectivism nowadays comes in more or less objective varieties. The most radical objectivists would say that Jane's evidence—and any agent's evidence really—can be ignored when assessing obligations. This used to include Thomson (1990, 229 *ff.*), but she later changed her mind (Thomson, 2008, 191–99), in part because of cases of 'objectively reckless' behaviour. If someone plays Russian Roulette for money, it is objectively reckless since *no human* could have predicted what would happen. One objectively ought not to play Russian Roulette for money, even if (by luck) no one would have been hurt. But even this revised objectivism still looks too radical. For instance, it leaves the objectivist verdict in HOSPITAL untouched, since it would be only *subjectively* risky to go to Modern.

More generally, this kind of objectivism rules out the obligation against reckless-ness (except in Russian roulette kinds of cases), as well as professional duties such as the journalist's duty to double-check their sources (see Kiesewetter, 2018a, p. 416), and the lawyer's duty (in some jurisdictions) to keep a diary (to avoid memory lapses). If we bracket the ignorance of the agent when we determine her obligations then, in the same gesture, we bracket the recklessness of her act, since an act is only reckless because it is done in such ignorance.

In a recent paper of this journal, Fassio (2022) casts objectivism under a more moderate light. He insists on the fact that objectivism is the view that *any fact* can be relevant to what an agent ought to do (see also Graham, 2010, p. 89), *including facts about the agent's limited evidence*. This allows him to claim that agents often ought not to be reckless. For instance, Eliot in fact *ought to go to Traditional*.

Fassio is surely right to insist that evidence-dependent obligations (against reck-lessness) must be accounted for and so that the more radical objectivism is deficient. But is his moderate objectivism plausible? One clear difficulty is that it no longer benefits from the support of typical arguments for objectivism. On Fassio's view, Eliot ought to go to Traditional. But a beneficent adviser would tell him that he ought to go to Modern instead (and he would be saying the truth). Fassio is then in the same boat as the perspectivist regarding the better information problem.

Moderate objectivism also delivers unsatisfactory verdicts about pairs of similar cases. Eliot ought to go to Traditional because it would be reckless otherwise. But suppose that Kathrin is in an identical situation as Eliot, except that, unbeknownst to her, her child is in a worse shape than Eliot's child. In fact, her child will die if she does not reach Modern very soon.[16] A moderate objectivist like Fassio should then accept that this unknown fact could in principle oblige Kathrin to go to Modern. After all, he says that all the facts are on the proverbial table when determining what one ought to do. But it is implausible to give an asymmetrical diagnostic of Eliot's and Kathrin's cases: if he ought to go to Traditional, so does she; and if she ought to go to Modern, so does he.

---

[16] *Unbeknownst to her* because otherwise the recklessness would be worsened by the higher stakes and Fassio could insist that Kathrin also ought to go to Traditional.

What this brings out is the difficulty of balancing evidential with non-evidential considerations. Fassio grants that this is senseless from the agent's perspective: one can't weigh the fact that a course of action is risky against the fact that it will, in reality, succeed (see also Kiesewetter, 2018b, p. 105). But this balancing also makes little sense from a third-person perspective. What is being compared by the third party? My only guess is that the third party is comparing two incomparable kinds of things, i.e. the axiological (the disvalue of pain or death) and the aretaic (the viciousness of the recklessness of Eliot and Kathrin). To say that Eliot, but not Kathrin, ought to go to Traditional suggests the following farfetched rationale: that, on balance, the vice of recklessness (or its manifestation) is worse than pain (the pain of Eliot's child), but better than death (the death of Kathrin's child). This is confused.

So, objectivism—at least as it stands—must either fail to account for obligations against recklessness or instead both sacrifice its advantages over its rivals (regarding advice) and lead to implausible verdicts about pairs of cases. Either way, objectivism has a recklessness problem. This is by no means a conclusive refutation of objectivism, but it suffices, for our purpose, to sustain a healthy confidence in $BION_r$.

Before I move on to the next objection, I want to offer an olive branch to the objectivist. She can, if she wants, accept $BION_r$. That's because even if she insists on saying that the 'main ought' or the 'ought of directives' is objective, she can recognise that there are statements where 'ought' works differently. In fact, we use the locution 'given' with regards to 'ought' and 'should' claims in our day-to-day lives. Instead of asking 'what should I do?', a beginner playing a complicated card game might ask 'what should I do, given what I know', to indicate that she does not want information about complex strategies that she has not been taught yet. The conscious akratic might say: 'I know that I ought to stop smoking, but I won't. So what ought I to do, given that: take up e-cigarette or not?' And a person who was blamed by an objectivist might rejoin: 'what should have I done, given my evidence?' Note that the 'given…' clause can have at least two roles in English. It can first serve to emphasise a fact to an interlocutor (either remind her of it or insist on it): '$given_{emphasising}$ what I told you yesterday, do you think I should apply to the job?' Alternatively, 'given' can serve to fix a fact: '$given_{fixing}$ that I will not help you tonight, do you still want to host twelve guests?' Here we are interested in the fixing sense.

So the objectivist could grant that the phrase 'ought given one's evidence' makes sense. How should it be defined? As Gibbons (2013, Chap. 2) noticed years ago, there is a difficulty in doing so neutrally. *First attempt*: to say that S ought to Φ relative to a set of facts just means that, if we add the corresponding propositions to the set of premises that S can use in practical deliberation, S ought to Φ (simpliciter). So to say that *Jane ought, given all the facts of the situation, to block shaft A* is equivalent to saying that, *if we suppose that propositions like 'the miners are trapped in A' are part of Jane's set of premises—her evidence—then she ought (simpliciter) to block shaft A*. This sort of definition is clearly perspectivist-friendly, since it defines 'ought given all the facts' with a counterfactual: what would the agent have to do *if* things were different, i.e. if she had access to all the (relevant) facts. *Second attempt*: to say that S ought to Φ relative to a set of facts just means that S ought to Φ, supposing that these facts were not misleading. This sort of definition is clearly objectivist-friendly, since

it defines 'ought given one's evidence' with a counterfactual: what the agent would have to do *if* things were different, i.e. if her evidence were not misleading.

Even if these definitions are not neutral and that we must ultimately choose between perspectivism and objectivism, as Gibbons thinks we must, the objectivist can't say that 'what we ought to do *given our evidence*' is a confused notion. Otherwise, she needs to give us an error theory of our ordinary usage. At worse for the defender of BION$_r$, the objectivist can maintain that 'ought given one's evidence' is a theoretically secondary notion, a notion which is nevertheless not stipulative since it is used in natural language (with the help of the 'given' clause) to communicate. That was the olive branch.

5.2 The objection from excuses

We have tied blameworthiness to the perspectival ought, or 'ought' given one's evidence. And we have just seen some reason to be sceptical of objectivism about 'ought': it mishandles the obligation against recklessness. But there remains an appealing objection to BION$_r$ which leads to objectivism but does not invoke it directly. It appeals not to the nature of obligations but to that of justifications and excuses: ignorance can only excuse, not justify, and so it cannot affect what we ought to do, contra BION$_r$.[17]

What are excuses and justifications? As we saw in § 2, they are defences; they rule out, or qualify, blameworthiness. But they do it differently: justifications rule out blameworthiness by explaining why the act was right, or not wrong, or corresponding to the subject's obligations; whereas excuses rule out blameworthiness by explaining why the agent did not manifest an ill will or a lack of control in doing the wrong thing. Call this the *simple model*. (We can leave aside exemptions for our purposes.)

Wherein does ignorance fall? Typically, excuses manifest something like a human frailty or limitation. But ignorance certainly counts as such. The same goes, it seems, for *inabilities*: that I cannot do what I ought is a limitation, and so seems like a prime candidate for an excuse.

However, this first consideration cannot settle the matter because justified actions often manifest human limitation. Take self-defence: if I legitimately kill you in self-defence, this manifests my inability to disable your threat without injuring you. But self-defence is widely considered a justification, not an excuse.[18] So we cannot infer from the fact that *a human manifests their limitation in Φing* that *Φing was unjustified and so at best merely excused*.

Still, there is something right about this first consideration. Acting in self-defence may manifest a weakness, but the weakness itself is not the justification. *That I was too weak to immobilise my aggressor* was not my justification—not quite. My justification was that I was under attack, and that killing was a legitimate means of resisting

---

[17] Nowadays, the Austinian idea that the confusion between excuses and justifications is at the source of many mistakes is present in epistemology, in ethics, and in philosophy of law. See e.g. Littlejohn (forthcoming), Williamson (2015), and Duff (2007).

[18] See Baron (2005, 388–89). The more controversial cases are those where the subject was not actually under threat but who made a reasonable mistake.

this attack. Of course, this second fact is in turn explained by my weakness. Similarly, *strictly speaking*, my ignorance itself cannot justify my act, even if my misleading evidence can.

A stronger consideration in favour of the view that ignorance (or misleading evidence) could not justify comes from the idea that even when we act out of ignorance, we still infringe some important standards. But a justified action does not, according to this objection, infringe important standards. What's more, these standards are quite stable, as is for instance the standard against infringing others' rights. But our evidence fluctuates a lot. So, ignorance (or misleading evidence) can only be an excuse, not a justification. The same goes for (all-in) inabilities: they fluctuate more (since opportunities fluctuate) than our obligations.

The underlying idea is that justification in a normative domain just is conformity to a simple and fundamental norm of that domain.[19] For instance, some might claim that one is epistemically justified just if one respects the knowledge norm: believe *p only if* you thereby come to know that *p*. Or, one is morally justified just if one respects the rights norm: Φ *only if* Φing respects others' rights. Or perhaps, one is legally justified just if one respects the lawful norm: Φ *only if* Φing is lawful. Everything that falls short of these fundamental norms is, in its respective normative sphere, at best excused. So, ignorance (or misleading evidence) cannot justify, since it neither changes these norms nor helps satisfy them.[20] The same goes for inability: it does not change the norms or help satisfy them. In fact, as I have now noted thrice, there is a striking similarity between ignorance and inability. Although this is a discussion for another day, the fate of BION and OIC seem tied together.

The view of justification as norm conformity is sometimes said to fit well with mundane obligations, professional or legal. The Dean's responsibility is to be present during the Faculty meeting, alas she is stuck abroad. She does not lose her responsibility and so cannot be justified: she is merely excused. Similarly: I don't know that this boat here is yours: during the night, mine, which looks the same, has been stolen and, early this morning you moored yours where I did mine previously. When I paint your boat in a distasteful shade of yellow, I am not justified: I still damaged your property. But I have an excuse.

Yet we don't need to accept the view of justification in ethics as norm conformity. In ethics, it seems that we deal with a domain where there are several norms in conflict, or differently put, where there is no primary one, unless a trivial one like 'be moral'. The idea that justification is just conformity to a norm is then unhelpful.[21] On

---

[19] Again, see for instance Williamson (2015).

[20] Might the fundamental ethical norm say 'Φ only if your evidence indicates that Φing is respectful of others' rights'? Apart from lacking the simplicity that the model promises, it poses a problem for right-making properties. If the norm is about conforming one's act to one's evidence, then what makes the act right is no longer properties of the act itself or of its consequences, but evidence of such properties to the subject. That seems false. Is that now to side with the objectivists? No. To say that *what makes an act right are its properties and its consequences, but only those which are somehow accessible to the subject* is not to say that *what makes an act right are pieces of evidence about its properties of its consequences*. For similar reason, what makes a cake good to a taster is not evidence about its taste; it's properties of the cake, but only those which are somehow accessible to the taster.

[21] By contrast, epistemology might have a main non-trivial norm, e.g. the knowledge norm.

the contrary, it is hard to deny that justified behaviour might involve at times breaking important ethical norms or rules.[22]

Likewise, the simple model of justifications and excuses is elegant, but its simplicity is ultimately misleading. It incites us to idealise justification and to excessively broaden excuses. We want to be able to recognise the fact that there is something imperfect in a person stealing food, even if she does it to save someone's life. She stole after all, and so infringed an important moral norm. But to recognise this fact, since we have two categories of non-blameworthy people, it is natural to put her in the second, in the excused. The same applies more generally to the ignorant or to the unable.

To resist this temptation, I want to recommend a different model of defences, which comes from legal philosophy,[23] and on which ignorance and inability can justify. This model does not explicitly contradict the simple model; we can see it as its *refinement*.

On this model, we must distinguish not two, but three defences (again, leaving aside exemptions). When one gives a *denial*, one argues that one did not infringe a relevant norm or rule.[24] Such norm might be the legal 'do not trespass' or the moral 'do not lie'. This gives us a new model: denials rule out blameworthiness by explaining why one has not infringed a norm; justifications rule out blameworthiness by explaining why the act was not wrong even if the agent may have infringed a norm[25]; whereas excuses rule out blameworthiness by explaining why the agent did not manifest a lack of control or an ill will even if she may have done the wrong thing. Call it the *refined model*.

The refined model explicitly recognises that the blame game typically proceeds in three steps (see Duff, 2007, Chap. 1): we check, first, whether one has infringed a norm and so ought to answer for it, or to explain it; second, whether the answerable agent fails to justify their action; finally, whether the unjustified agent lacks an excuse. If, at any of these steps, we can respond 'no', the subject is not liable to blame (nor, in the law, to punishment or imprisonment).

The superiority of the refined model can be illustrated by the following case:

NECESSITY: During his customary evening walk, Larry notices that an apartment building has caught fire. Larry tries to ring people in the building, but there is no response. He quickly spots through a window that there is a fire extinguisher in the corridor of the first floor. He breaks the door, enters, and puts the fire out.

Say Larry is prosecuted for having damaged his neighbour's property. He does not merely have an excuse—that would be a falsely derogatory description of his situation (see Baron, 2005, 2007). On the contrary, Larry has a *justification*: he did the

---

[22] It would, at any rate, require accepting the existence of true dilemmas, as Williamson does in epistemology.

[23] See Baron (2005), Greenberg (manuscript) and Duff (2007, Chap. 11).

[24] This is not to be confused with denials of responsibility involved in e.g. insanity defences. See Gardner (1998).

[25] We probably want to accept that someone may be justified in cases where no norm was infringed.

right thing. The Common Law expresses this idea by allowing that Larry be, in principle, asked to explain his behaviour in the form of a 'necessity' defence—he has, after all, broken a door. But a necessity defence gives a justification, not an excuse—an excuse would not be applicable since Larry did the right thing. Nor does it give a denial—which would be inaccurate since Larry did damage his neighbour's door.

The refined model of defences naturally can treat misleading evidence as justificatory. Jane does not have a mere *excuse* for not blocking shaft A; it was the right decision given what she knew: she was justified. Of course, she infringed a moral norm or rule: she did not save the miners working under her supervision from injury. She lacks a *denial*, for sure, and so she should be ready to justify her action to the injured miners. What's more, she will have to answer for something for which she is blamelesss (see Kiener 2024 for a similar treatment of William's lorry driver). But she is justified.[26]

Both Larry and Jane then infringe a norm while being justified. True, their cases are quite different: whereas his reason to extinguish the fire *outweighs* his reason to leave his neighbour's property intact, Jane's reason to refrain from blocking a shaft does not outweigh her reason to save everyone. But it's important to note in each case that a norm was infringed, and that each protagonist can be asked to answer for their conduct.

To be clear, the simple model does not strictly imply that ignorance is an excuse and so that BION$_r$ is false. But, contrary to the refined model, it misleads us into thinking just that. Again, if we divide all defences into justifications and excuses, it is tempting to see justified actions (or beliefs) as flawless. For we want to distinguish actions which infringe a norm from those which do not. Having only two baskets, we put all the imperfect cases into the excused basket. We then reject BION$_r$, on the grounds that ignorance does not change norms. If, by contrast, we divide defences into denials, justifications, and excuses, it is less tempting to see justified actions as flawless. Having now three baskets, we can easily separate actions which infringe a norm from those which do not while, at the same time, separating them from the merely excusable actions. BION$_r$ can then be saved from the intuitive view that ignorance can at best excuse and not justify.

6 Conclusion

If someone is blameworthy for having acted, she had an obligation against so acting, given her evidence (BION$_r$). Blameworthiness implies 'ought not'. So, if Pierre-Luc is not on the work shift today, yet shows up at 19:00 believing he was late, we cannot maintain that he did everything he had to do and yet is blameworthy.

This principle is prima facie plausible: without it, we can't make sense of the idea that a justification is a defence. If we show that we did everything right, we don't

---

[26] That is of course not to say that ignorance cannot be an excuse. Some considerations might be both; see Bruno (2022, 2). Bruno ultimately argues that, given that objectivism about 'ought' is false, full excuses have a bearing not just on culpability but also on permissibility. I'm sceptical about this but if he is right, the defence game is somewhat complexified.

need *then* to prove that we are blameless. Differently put, we don't need *both* a justification and an excuse to avoid blame.

But there are strong objections to BION$_r$. We tackled five.

The objection from conscience maintains, roughly, that people who do everything they ought in Φing, but do it against their conscience, are blameworthy; so 'blameworthiness' does not imply 'ought not'. This over-subjectifies blameworthiness. Even when we consider the difficult cases, those of unreasonable factual ignorance and of reasonable moral ignorance, we see that the objection from conscience is unsustainable. Amongst other things, it fails to respect the idea that if a person is blameworthy for an act, it is *ceteris paribus* fitting to expect her to change course of action.

The objection from reasons says that people can be blameworthy for doing the right thing when they act for wrong reasons. Yet again, this runs afoul of the principle according to which blaming a behaviour is only fitting where expecting a change of behaviour is fitting *ceteris paribus*. This objection also oversubjectifies blameworthiness since it deems it sufficient to be blameworthy for Φing that Φing was done for the wrong reason. If people who do the right thing for the wrong reason are blameworthy, it is for something else, something much less serious such as manifesting certain bad attitudes.

The objection from suberogation runs on the idea that sometimes we do what we ought and still act in a morally bad way. Some have wanted to classify these acts as blameworthy. But this turns on an 'ought' that is correlated with rights, which is not the 'ought' that is relativised to our evidence. It is again over-objectified. Or, alternatively, this objection uses a notion of blameworthiness which is too weak.

The objection from objectivity maintains that 'oughts' are objective and so can depend on facts that are inaccessible to the agent. But then, at times, one recklessly does what one (objectively) ought to do by chance, and so is blameworthy—since this was reckless. I argued that this objective conception of 'oughts', whether it takes a radical form (e.g. Thomson's) or a moderate form (e.g. Fassio's), deals poorly with the unavoidable idea that we must not be reckless. We must not over-objectify oughts.

The objection from excuses claims that ignorance is an excuse, not a justification. If that's right, ignorance can't determine what we ought to do—which takes us back to the previous objection. The view that ignorance is at best an excuse relies either on the idea that what manifests human limitation, as ignorance does, must at best excuse or on the idea that ignorance could not contribute to a norm being respected. But this is to idealise justification. Ignorance can justify since justifications may also manifest human limitations. What's more, justifications admit of norm infringements (such as when I am justified in breaking a window to save a cat), as the refined model of defences makes so clear.

By admitting the principle that 'blameworthiness' implies 'ought not', we are led to a desirable middle-ground: no over-subjectification of 'blameworthiness', no over-objectification of 'ought'.[27] By responding to objections to this principle, we can also make progress with regards to difficult ethical notions: suberogation, conscience, rea-

---

[27] This does not mean that blameworthiness and infringements of obligations are coextensive: it leaves open the possibility for fully excused infringements of obligations. For a critical discussion, see e.g. Rivera-López (2006) and Bruno (2022).

sons, objectivity, justifications and excuses. It remains to be seen whether, as I have pointed out at the outset, it has desirable consequences about free will, moral luck, and responsibility.

References

Alvarez, M., & Littlejohn, K. (2017). When ignorance is no excuse. In P. Robichaud, & J. W. Wieland (Eds.), *Responsibility - The epistemic condition* (pp. 64–81). Oxford: Oxford University Press.

Arpaly, N. (2002). Moral worth. *The Journal of Philosophy*, *99*(5), 223–245. https://doi.org/10.2307/3655647.

Austin, J. L. (1956). A Plea for excuses: The presidential address. *Proceedings of the Aristotelian Society*, *57*, 1–30.

Baron, M. (2005). Justifications and excuses. *Ohio State Journal of Criminal Law*.

Baron, M. (2007). Excuses, excuses. *Criminal Law and Philosophy*, *1*(1), 21–39. https://doi.org/10.1007/s11572-006-9001-2.

Björnsson, G., & Finlay, S. (2010). Metaethical contextualism defended. *Ethics*, *121*(1), 7–36.

Bruno, D. (2022) Being fully excused for wrongdoing. *Pacific Philosophical Quarterly*, 104: 324–47. https://doi.org/10.1111/papq.12425.

Capes, J. A. (2012). Blameworthiness without wrongdoing. *Pacific Philosophical Quarterly*, *93*(3), 417–437. https://doi.org/10.1111/j.1468-0114.2012.01433.x.

Copp, D. (2003). Ought implies can, blameworthiness, and alternative possibilities. In D. Widerker, & M. McKenna (Eds.), *Moral Responsibility and Alternative Possibilities: Essays on the Importance of Alternative Possibilities* (pp. 265–99). Ashgate.

Copp, D. (2008). Ought implies can and the derivation of the principle of alternate possibilities. *Analysis*, *68*(1), 67–75. https://doi.org/10.1093/analys/68.1.67.

Darwall, S. L. (2006). *The second-person standpoint: Morality, respect, and accountability*. Harvard University Press.

Driver, J. (1992). The Suberogatory. *Australasian Journal of Philosophy*, *70*(3), 286–295. https://doi.org/10.1080/00048409212345181.

Duff, A. (2007). *Answering for crime: Responsibility and liability in the criminal law*. Legal Theory Today. Oxford.

Dworkin, R. (2011). *Justice for hedgehogs*. Harvard University Press.

Fassio, D. (2022). What the doctor should do: Perspectivist duties for objectivists about ought. *Philosophical Studies*, *179*(5), 1523–1544. https://doi.org/10.1007/s11098-021-01717-x.

Gardner, J. (1998). The gist of excuses. *Buffalo Criminal Law Review*, *1*(2), 575–598. https://doi.org/10.1525/nclr.1998.1.2.575.

Gibbons, J. (2010). Things that make things reasonable. *Philosophy and Phenomenological Research*, *81*(2), 335–361.

Gibbons, J. (2013). *The norm of belief*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199673391.001.0001.

Graham, P. A. (2010). In defense of objectivism about moral obligation. *Ethics*, *121*(1), 88–115. https://doi.org/10.1086/656328.

Haji, I. (1993). Alternative possibilities, moral obligation, and moral responsibility. *Philosophical Papers*, *22*(1), 41–50. https://doi.org/10.1080/05568649309506392.

Haji, I. (2012). *Reason's debt to freedom: Normative appraisals, reasons, and free will*. Oxford University Press.

Hart, H. L. A. (1968). *Punishment and responsibility*. Clarendon.

Henning, T. (2014). Normative reasons contextualism. *Philosophy and phenomenological research*, *88*(3), 593–624. https://doi.org/10.1111/j.1933-1592.2012.00645.x.

Kiener, M. (2024). Strict Moral Answerability. *Ethics, 134*(3), 360–386.

Kiesewetter, B. (2017). *The normativity of rationality*. Oxford University Press.

Kiesewetter, B. (2018a). What kind of perspectivism? *Journal of Moral Philosophy*, *15*(4), 415–443. https://doi.org/10.1163/17455243-20170007.

Kiesewetter, B. (2018b). How reasons are sensitive to available evidence. In C. McHugh, J. Way, & D. Whiting (Eds.), *Normativity: Epistemic and Practical* (p. 0). Oxford University Press. https://doi.org/10.1093/oso/9780198758709.003.0006.

Kolodny, N. (2005). Why be rational? *Mind*, *114*(455), 509–563. https://doi.org/10.1093/mind/fzi509.

Kolodny, N., & MacFarlane, J. (n.d.) Ought: Between objective and subjective. *Unpublished Manuscript*.

Littlejohn, C. (forthcoming). A plea for epistemic excuses. In F. Dorsch, & J. Dutant (Eds.), *The new evil demon problem*. Oxford University Press.

Lord, E. (2017). What you're rationally required to do and what you ought to Do (are the same thing!). *Mind*, *126*(504), 1109–1154. https://doi.org/10.1093/mind/fzw023.

McMahan, J. (2009). Intention, permissibility, terrorism, and war. *Philosophical Perspectives*, *23*, 345–372.

Nelkin, D. K. (2011). *Making sense of freedom and responsibility*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199608560.001.0001.

Olsen, K. (2017). A defense of the objective/subjective moral ought distinction. *The Journal of Ethics*, *21*(4), 351–373. https://doi.org/10.1007/s10892-017-9258-9.

Regan, D. H. (1980). *Utilitarianism and co-operation*. Oxford University Press.

Rivera-López, E. (2006). Can there be full excuses for morally wrong actions?* *Philosophy and phenomenological research*, *73*(1), 124–142. https://doi.org/10.1111/j.1933-1592.2006.tb00607.x.

Scanlon, T. M. (2008). *Moral dimensions: Permissibility, meaning, blame*. London: Belknap.

Skorupski, J. (1999). *Ethical explorations*. Oxford University Press.

Strawson, P. F. (1974). *Freedom and resentment and other essays*. Routledge.

Thomson, J. J. (1990). *The realm of rights*. Harvard University Press.

Thomson, J. J. (2008). *Normativity*. Open Court.

Vranas, P. B. M. (2007). I ought, therefore I can. *Philosophical Studies*, *136*(2), 167–216. https://doi.org/10.1007/s11098-007-9071-6.

Widerker, D. (1991). Frankfurt on ought implies can and alternative possibilities. *Analysis*, *51*(4), 222. https://doi.org/10.2307/3328762.

Williams, B. (1995). *Making sense of humanity: And other philosophical papers*. Cambridge University Press.

Williamson, T. (2015). Justifications, excuses, and sceptical scenarios. *The New Evil Demon*. OUP.

Wolf, S. (2001). The moral of moral luck. *Philosophic Exchange* 31 (1).

Zimmerman, M. J. (1997). A plea for accuses. *American Philosophical Quarterly*, *34*(2), 229–243.

Zimmerman, M. J. (2008). *Living with uncertainty: The Moral significance of ignorance*. Cambridge University Press. Cambridge Studies in Philosophy.