Ethical Theory and Moral Practice (2022) 25:21–36
https://doi.org/10.1007/s10677-021-10213-4

Moral Phenomenology and the Value-Laden World

William J. FitzPatrick [1]

Accepted: 11 June 2021 / Published online: 21 June 2021

Abstract
Do the introspectively ascertainable aspects of our moral experiences carry *ontological objective purport*—portraying reality as containing worldly moral properties and facts, thus supporting moral realism? Horgan and Timmons (2008, 2018) answer this question in the negative, arguing that their non-realist view, cognitivist expressivism, can accommodate the introspectively ascertainable moral phenomenology (including categorical authoritativeness) just as well as realism can—where *accommodating* the phenomenology means accounting for it without construing it as misleading or erroneous. If sound, this constitutes an important defense of cognitivist expressivism, undermining a central attraction of realism. They thus pose a challenge to realists to identify any aspects of moral phenomenology that cannot be accommodated by expressivism and instead favor realism. I here take up that challenge, in two stages. First, I argue that cognitivist expressivism does not after all capture certain important aspects of the phenomenology of the sort of moral experience on which they focus, while realism does. This argument does not depend on claiming that the phenomenology has ontological objective purport. The claim so far is just that there is more to categorical authoritativeness than the expressivist account captures, though this leaves the door open to Kantian rationalism (and perhaps other non-realist accounts) as well as realism. Second, I will go on to argue that although some aspects of moral phenomenology may only point to this broader range of views, others do specifically carry ontological objective purport and thus directly support realism insofar as we take the phenomenology seriously.

Keywords Moral phenomenology · Moral realism · Expressivism · Moral objectivity · Ontology

1 Introduction

Metaethics takes up semantic, metaphysical and epistemological questions about the ethical domain, setting aside first-order ethical controversies to examine these more abstract questions

✉ William J. FitzPatrick
william.fitzpatrick@rochester.edu

[1] University of Rochester, Rochester, NY, USA

about ethics as such. Still, many of us think that in doing metaethics it is important to take our own engaged ethical experience seriously, allowing our reflective interpretation of it to guide our metaphysical, epistemological and semantic inquiry concerning value and normativity.

This does not mean, of course, that one could not still be led by compelling metaethical arguments to conclude that one's ethical experience is misleading and thus to adopt a deflationary interpretation of it in the end—just as many have done with respect to certain religious experience. The starting point in first-personal appearances is a defeasible one. But it is important nonetheless in guiding metaethical inquiry. Ethical experience has more power to shape our metaethics—potentially even our ontology—when considered from a first-person perspective than when viewed merely from a third-person perspective, as by scientists in the course of seeking parsimonious causal explanations of people's moral beliefs, feelings, and motivation. One is less likely to discount ethical appearances, such as the sense of being categorically bound by objective moral requirements, when these appearances are one's own and central to one's practical life, rather than mere reports of others' impressions that might readily be discounted in the name of theoretical parsimony.

Suppose, then, that we take our *moral phenomenology*—the appearances of moral experience—at least provisionally at face-value as we seek to develop a metaethical view that coheres with it, without deflating it or impugning it as erroneous. We might then ask what exactly the moral phenomenology gives us pro tanto reason to believe about the nature of morality. More specifically, to take up a question helpfully framed and explored by Terry Horgan and Mark Timmons: do the introspectively accessible aspects of our moral experiences carry *ontological objective purport*, portraying reality as containing worldly moral properties and facts within its ontology (Horgan and Timmons 2008, 2018)? That is, does moral phenomenology, when given a serious role in metaethics, support an ontologically robust form of moral realism (henceforth just 'realism')?

Horgan and Timmons say no: we could take moral phenomenology as seriously as you like and still not thereby be driven toward realism. Their argument turns on showing that their non-realist view, cognitivist expressivism, can accommodate the introspectively ascertainable moral phenomenology just as well as realism can—where *accommodating* the phenomenology means accounting for it without construing it as misleading or erroneous (Horgan and Timmons 2008, 2018). If sound, this would constitute a powerful defense of cognitivist expressivism, undermining a central attraction of realism: for realism loses much of its appeal if any moral objectivity seemingly revealed by moral experience can in fact be captured more economically within an expressivist framework, without need of instantiated moral properties and facts in the world. They thus pose a challenge to realists to identify any aspects of moral phenomenology that cannot be fully captured by cognitivist expressivism and might instead favor realism.

I here take up that challenge. The project proceeds in two parts. First, I will argue that cognitivist expressivism does not after all capture certain important aspects of the phenomenology of the sort of moral experience on which Horgan and Timmons focus, while realism does. This argument does not, however, depend on claiming that the phenomenology has ontological objective purport such that it specifically supports realism. The claim is instead just that there is more to categorical authoritativeness than the expressivist account captures, though this still leaves the door open to other non-realist accounts such as Kantian rationalism (Korsgaard 1996, 2003) or non-realist cognitivism (Parfit 2011, 2015), which might so far do as well as realism but without robust ontological commitments. If this is correct, then while it does not yet specifically favor realism it does identify an advantage over expressivism, given

realism's clear handling of all the phenomenology, which also provides a basis for a larger argumentative strategy for realists, as described in the concluding section.

My second task is to go on to argue that although some aspects of moral phenomenology may only point to a range of views that include but are not limited to realism, others do seem to point specifically to realism. This second line of argument will require broadening the scope of the phenomenology under investigation. If successful, it suggests that certain aspects of moral phenomenology do carry ontological objective purport and thus directly support realism insofar as we take the phenomenology seriously.

2 Cognitivist Expressivism and the Phenomenology of Direct Moral Experience

Horgan and Timmons focus their discussion on the introspectively ascertainable moral phenomenology associated specifically with *direct* and *intuitive* experiences of moral *obligation*—experiences involving spontaneous, intuitive moral judgment and the appearance that one ought to perform some action in response to the circumstances confronting one. Their claim is that cognitivist expressivism "can smoothly accommodate the introspectively accessible aspects of the phenomenological features" of such cases (Horgan and Timmons 2008, 2018).

Let's begin by considering a relevant example and the associated phenomenology according to the model they adopt from Maurice Mandelbaum (1955).

> **Hit-and Run:** While out walking during a break between meetings, you observe a car make a careless turn and sideswipe a child on her bike, knocking her to the ground at the side of the road. The car speeds off, and though the child does not appear to be seriously injured, she may have some injuries and is visibly upset. Immediately, you feel a *demand* to act—a sense that you *ought* to rush to her aid, make sure she is out of harm's way, call for help, and stay with her to provide comfort. You experience this demand as coming from *outside* of yourself—from the *objective circumstances* you are confronting and in a way that is *independent of your desires* and important enough to override any concerns about being late to your next meeting. It seems to you that these are things you must do, based simply on the situation and the obvious *unfittingness* of just walking away.

The concepts emphasized in this description bring out the central features of the moral phenomenology. There is a felt demand involving a spontaneous, intuitive moral *ought* judgment, which Mandelbaum takes to be grounded in an immediate apprehension of the unfittingness of failing to help the child: such a failure, whether out of indifference or due to prioritizing a meeting over aiding a child in need, would be inherently unfitting to the situation, which "calls for" action. Moreover, the demand is experienced as being *objective* in two senses that, as we will see later, it is important to distinguish:

(A) it comes across as a *categorical* demand, binding you independently of what you happen to desire, with "categorical authority" (Horgan and Timmons 2018), and
(B) it seems to be "emanating from 'outside' [you]", having its *origin* or *source* in something external to yourself (Mandelbaum 1955, 54).

Our question, then, is whether cognitivist expressivism can really accommodate this phenomenology, as Horgan and Timmons claim. As they put it, the crucial element to capture here is the experience of "a concrete action being called for by desire-independent features of one's present circumstances" (Horgan and Timmons 2008, 293)—in our example, the helping behavior being called for by the objective circumstances of the child's need for care, safety and comfort. How does cognitivist expressivism, despite its rejection of ontological moral objectivity, accommodate this? The idea goes as follows.

First, belief comes in two varieties: is-belief and ought-belief. Is-beliefs involve an 'is-commitment' to some descriptive propositional content, taking it to be true: your believing *that an accident has occurred*, for example, consists in your having an is-commitment with regard to the content *that an accident has occurred*. Ought-beliefs instead involve an 'ought-commitment' with respect to some equally descriptive propositional content: your believing that you ought to help the child, for example, consists in your having an ought-commitment with respect to the content *that you help the child*. (This contrasts with the realist's construal of the ought belief as an is-belief with respect to a *moral* propositional content, i.e., having an is-commitment with regard to the content *that you ought to help the child*.) So the moral judgment here is construed as an intuitive ought-belief formed spontaneously in response to the circumstances you confront (Horgan and Timmons 2008, 290–92).

Next, they maintain that a moral ought-belief is characterized by taking a "non-self-privileging stance" with respect to one's circumstances and action, which involves "being open to being affected by [certain] desire-independent considerations" (Horgan and Timmons 2008, 294). To be morally responsive to our circumstances thus involves being disposed to take up this stance, open to being affected by certain kinds of desire-independent considerations in certain ways. So in our example involving direct experience of moral obligation, what is happening is that you encounter this situation with such a disposition and you spontaneously form an ought-belief (i.e., you become ought-committed) concerning helping the child, *based simply on desire-independent considerations such as the fact that the child is potentially injured and in danger, is scared, and needs help*.

Moreover, and crucially, this is how you experience the situation: you experience yourself forming this ought-belief (or ought-commitment) *because* of those desire-independent considerations about the child and her needs. That is a central part of the phenomenology: you feel yourself becoming ought-committed simply by those objective considerations—as opposed to feeling yourself being ought-committed by anything about your own desires (Horgan and Timmons 2008, 294–95). This is reflected in the fact that if asked why you are helping the child, you will simply point to those objective considerations—the child's plight—rather than bringing in your desires or anything else about you other than your being in a position to help. You will say that you are helping simply *because* the child is in danger, is scared, and needs help: those are your *reasons* for helping, not something about your own desires. The same holds for your reasons why you think you are obligated to help.

This, then, is their account of the moral phenomenology we have sketched for this case—specifically, the experience of "a concrete action being called for by desire-independent features of one's present circumstances" (Horgan and Timmons 2008, 293). Nowhere in this picture is there any ontological, worldly normative property objectively calling for the helping behavior or making it unfitting to fail to help; there is no "in-the-world unfittingness fact" (295). There are just the ordinary factual properties of the circumstances we have highlighted, along with you, who have certain contingent psychological attitudes or commitments, including a disposition to care about children's welfare as such—an ought-commitment with respect

to the treatment of children, which you might express with ought-beliefs framed in terms of the 'unfittingness' of ignoring their needs. When these come together, and you encounter the situation described, you spontaneously form the ought-belief in question, in direct response to the objective features of the circumstances, and you experience this connection. We have therefore accounted for the appearance of a categorical moral demand, binding you independently of your desires, coming from outside of yourself, i.e., from the objective features of your circumstances.

3 Critique of the Cognitivist Expressivist Account of the Phenomenology

There are reasons, however, to doubt that this account succeeds in accommodating the phenomenology. First, let's introduce some terminology. Moral properties such as wrongness are "resultant properties" in the sense that they are attributable *by virtue of* the attribution of certain other properties, the "resultance-base properties" (Dancy 2004, 2006). For example, the wrongness of walking away and failing to help the child is a resultant moral property of such behavior, attributable by virtue of the behavior's having such features as allowing a child in crisis to suffer and remain at risk simply in order to avoid missing a meeting—these being the *wrong-making*, resultance-base properties of the act. Of course, to avoid begging questions against expressivism we should here understand talk of 'properties' broadly and minimalistically, such that as long as we can speak of an act's being wrong, for example, we can speak of its 'having the property' of wrongness, leaving open whether this carries any metaphysical weight. We are here just concerned with common structural points.

When you make a moral judgment, attributing a resultant wrongness to an act such as failing to help the child, you do so *on the basis* of attributing resultance-base properties to the act, as such—properties you take to be wrong-making features. And taking certain features to be wrong-making amounts to viewing them as making the act violate relevant *standards* of moral rightness: for it is only against a background conception of the appropriate moral standards that particular features of an act can be wrong-making, insofar as they make the act fail to 'stack up' against those standards. So in judging that failing to help would be wrong, you are attributing a resultant moral property or status to the act on the basis of its having certain features you take to be wrong-making in light of your background conception of the moral standards for right action (again understanding these structural claims in a broad way compatible with expressivist construals of 'property attribution' or 'conception of the moral standards').[1]

Now the way in which cognitivist expressivism seeks to capture the dimension of *objectivity* within the phenomenology is again by emphasizing the *desire-independence* (i) of the features of the external circumstances you confront, i.e., the facts about the child's plight, and (ii) of their bearing on your moral judgment that it would be wrong not to help, such that you are obligated to help. The idea is that these objective factors form the basis of the moral judgment without reference to anything about your own desires. They have to do simply with relevant features of the circumstances: the child's situation and the fact that walking away

[1] It doesn't matter for present purposes how exactly moral standards are to be construed—whether as principles or as something more like virtue-theoretic criteria with limited codifiability, for example. The point is just that attribution of the resultant property of rightness or wrongness to an act based on certain of its ordinary properties involves some conception of the ways in which such properties (individually and in various combinations) bear on rightness or wrongness, amounting to a conception of what the correct moral standards are.

would be allowing a child in crisis to suffer unnecessarily. Given relevant background standards, the basis for the moral judgment that you must help consists simply in certain objective features of the circumstances and therefore of the possible actions open to you. These are the objective resultance-base features you cite as *your reasons* for thinking you must act, and then for acting, *given* certain background standards you are employing.

This is all true as far as it goes, but the problem is that it fails to capture an important part of the phenomenology. Your moral experience is not simply that of being required to do something by objective features of the circumstances you face *given* a certain set of standards *to which you happen to be committed*. A crucial part of the phenomenology is the powerful sense that the standards you are employing are *themselves* imposed upon you independently of your choices or contingent commitments or causal psychological shaping by your society. Not only are the wrong-making features of walking away from the child objective, along with their relation to a given set of standards, but *their wrong-makingness* itself seems to be objective, which is to say that the associated *standards* themselves have an objective status. The moral experience is that of being confronted by a moral demand that is backed up by categorically authoritative standards to which you are committed because they *objectively merit* that commitment—not because you have simply been raised to be so committed or made the choice to be.

This appearance is not itself explained by the facts that:

(1) the resultance-base properties involve objective, desire-independent factors (features of the circumstances involving the child's needs),
(2) the standards in question have desire-independent *content* (about how properly to respond to such needs, regardless of one's desires), and
(3) the resultance-base features interact directly with the standards (without reference to your own desires) to rule out walking away from the child.

The moral experience is not just a matter of becoming ought-committed to some action by certain desire-independent features of the circumstances, but more specifically of becoming ought-committed or obligated by those features because of how they interact with a set of moral standards to which one experiences oneself as committed *not* due to any psychological or sociological contingencies but because they are *categorically prescribed* to one as the objectively correct standards.[2] This involves the experience of being categorically bound by the background standards themselves, as part of the moral experience of the felt moral demand: it is part of what we experience as issuing (as Mandelbaum emphasized) from outside of ourselves, and it goes beyond anything addressed by the three factors above, to which Horgan and Timmons (2008) appeal in explaining the phenomenology.

Let me emphasize that I am not at this stage claiming that the phenomenology in question carries *ontological* objective purport and thus specifically supports realism. My claim so far is only that there is a part of the phenomenology that is missed by the features of desire-independence and objectivity on which Horgan and Timmons focus in their expressivist accommodation project. If this is correct, then their accommodation project remains

[2] Cf. Jean Hampton's (1998, 94–96) characterization of moral norms as "culture-independent": they are "not a matter of social and psychological contingencies," and possess an independent, objective authority, rather than being optional—an authority that "is not the invention of the agent, nor of human communities, but something to which agents and human communities respond" (98).

incomplete, missing what is arguably the most crucial aspect of the phenomenology. I'll come back to this when considering possible responses. First let me elaborate upon the present point by examining a contrasting *non*-moral case, which I'll argue contains the same aspects of objectivity highlighted by Horgan and Timmons (the three factors laid out above) but lacks the deeper objectivity I'm claiming characterizes moral experience:

> **Struggling Cafe** Suppose you have voluntarily chosen to take on a commitment to the success of a café you've come to frequent. You took on this commitment not out of any sense of moral obligation but just because you happen to value it and like its owners, and so care about its success. But *given* that commitment, you come to feel that you must make a donation this month, based simply on the fact that without that donation they will not be able to pay their rent and will have to close temporarily, which in turn will get them even further into debt and likely lead to the loss of the café.

Note that in characterizing this as a non-moral case I'm assuming that you do not construe your commitment and choice as the fulfilment of an 'imperfect' moral duty to be charitable—a duty that allows you leeway with respect to which particular charitable activity to engage in. So to keep it simple let's suppose that you take yourself already to have fulfilled any moral duty of charity through more than enough other charitable work. Your commitment to the café stems simply from the personal concern you happen to have for it and its owners, and is not seen in light of any sort of moral duty.[3]

Now what is it that makes you feel compelled to donate—which you'll cite as your reason for donating if asked why you did so? It is the objective fact that the café is in crisis and needs the donation to survive—not something about your own psychology. Naturally, the crisis leads you to feel compelled to act only because of your background commitment to the success of the café: those lacking such a commitment will not similarly feel the need to act. But *given* your background commitment to the success of the café, your reason for feeling the need to act is simply the desire-independent need you've recognized in the circumstances: it's about the café's needs, not your own desires—just as in the moral case.

Of course, you could always indicate your background commitment here by adding "…and I don't want the café to fail," if someone were puzzled by your taking the café's needs as a reason to donate. But that's equally true in the moral case: if someone were puzzled by why you are led by the child's situation to rush over to help, you could always explain by adding "…and I don't want her to be further injured or to continue suffering like that," indicating your commitment to moral standards that preclude allowing such things. But in neither case is it necessary to cite one's own desires *given* the background commitment in question: what moves you, given that commitment, are simply the objective features of the circumstances you confront.

To this extent, then, the two cases are parallel, both exhibiting the sort of objectivity emphasized by Horgan and Timmons and captured by cognitivist expressivism. But there remains a crucial difference in the phenomenology of the two cases, which their account misses. Despite what we've just said about the way in which you respond directly to the objective features of the circumstances in Struggling Café, *given* your background commitment, you are still aware that your background commitment itself has a contingent basis in your earlier choice to support this business. Yes, you respond directly to the café's needs

[3] Thanks to an anonymous referee for pointing out potential complications here, which I think are avoided with this stipulation. (See also the next footnote.)

without bringing your desires into it, but the only reason you *do* so is that you have this background commitment, which you understand to stem entirely from your earlier contingent concerns and choices. If you are not now questioning that commitment or the underlying concerns that led to it, then you needn't have any thought of your own desires here as you perceive the situation and write the check. But you still experience this whole situation very differently from the way you experience a moral obligation, as in Hit-and-Run.

In Struggling Café, it does not appear to you that the background commitment to the café's success is itself objectively prescribed for you by something outside of your own concerns and choices. You are aware that this underlying commitment was voluntary and optional, and that affects the way you experience the pull toward donating: yes, it's something you must now do to save the café given the circumstances, but you also know that nothing forced you to commit to the café in the first place; and you likewise recognize that others needn't be called to respond in the same way. By contrast, in Hit-and-Run, you experience a demand to act that springs from the circumstances together with a set of moral standards to which you feel yourself to be committed *not* due to any such choices or other psychological or sociological contingencies, but *because they merit that commitment from anyone*, such that anyone who fails to share this commitment is making an objective mistake.[4]

It might be tempting for the expressivist to try to capture this deeper and more categorical objectivity just by *building something more into the content* of the commitments: not only are you committed to behaving in certain ways but you are also committed to criticizing others who don't share those standards as 'mistaken'. Such a move, however, does not help. After all, someone else could likewise commit to very different moral standards and similarly build into them a commitment to criticize those who (like you) don't share *those* standards, calling them (e.g., you) 'mistaken'. This just yields parity from an objective standpoint, with each of you claiming a privileging of your own standards over the other's and criticizing the other's standards as mistaken: each of you will be *trivially correct* relative to your own adopted standards but incorrect relative to the other's. What we are after, however, is the appearance of a deeper, asymmetrical objectivity such that you experience your commitment to certain moral standards as itself being called for by the objective correctness of those standards themselves, which objectively privileges them over others.[5]

If this is correct, then although cognitivist expressivism can accommodate the sort of objectivity common to the phenomenology of the moral and non-moral cases above, it has not thereby captured the deeper objectivity in the moral case. It has not accommodated the sense of being bound by a demand stemming not just from objective resultance-base properties of the circumstances *given* a certain background commitment one has taken on, but from

[4] Specific fulfillments of an imperfect moral duty of charity would also differ from Hit-and-Run in that those acts would not themselves be experienced as required of every relevantly situated agent, as helping is in Hit-and-Run. But one would still experience the more general demand to be charitable as stemming from standards that seem to merit commitment from anyone, rather than being optional. That phenomenology needs to be accounted for as well, posing a further challenge for expressivists. But again, Struggling Café is meant to raise a different challenge. The point there is just that Struggling Café exhibits the forms of objectivity delivered by the expressivist account while lacking a crucial form of objectivity exhibited by Hit-and-Run, thus highlighting what the expressivist account fails to capture in Hit-and-Run.

[5] This is related to criticism of sophisticated expressivist attempts to escape worries about a deep sort of relativism, in FitzPatrick (2011 and Forthcoming-b). These worries, along with those raised in the text, are also related to problems Andy Egan has raised for quasi-realist expressivism in connection with the idea of "fundamental moral error" (Egan, 2007).

objective resultance-base properties of the circumstances in light of a special set of moral standards that objectively and categorically merit one's attention and commitment.

Again, I'm not claiming that this appearance of categorical authoritativeness carries ontological objective purport (though that would be one way to account for it, as discussed below, and I will go on to consider other cases that I'll argue do carry such purport); for all I've said so far, Kantian rationalism may capture categorical authoritativeness just as well as realism, without any ontological objective purport. The claim so far is just that the expressivist accommodation project doesn't fully capture categorical authoritativeness, since it speaks only to what Hit-and-Run and Struggling Café have in common, not to the additional elements I've claimed (along similar lines as Hampton 1998) characterize the former but not the latter.

Now one response to this objection would be just to deny that there is any such further element of the moral phenomenology that needs accommodating. As I read Horgan and Timmons (2008 and 2018), however, they would not rush to take this line. In describing the appearance of "inescapability" (see below), for example, they seem to grant that the phenomenology of categorical authoritativeness is as I've described it, and claim to have recognized and fully accommodated it. What I've argued is that their account has not succeeded in doing that: they may have recognized all the phenomenology and accounted for some of it, but a crucial part has been left out of the account. So another response would be to grant that the above account (from Horgan and Timmons 2008) misses something but to claim that this further element can still be accommodated within expressivism.

In their more recent work, Horgan and Timmons (2018) focus on identifying elements of the phenomenology such as "pull", "source", "grip", "independence", and "inescapability," arguing that these capture the full range of the phenomenology of categorical authoritativeness and can be fully accommodated by expressivism—where again *accommodation* means accounting for them *without* impugning moral experience as being misleading or erroneous (116). For the most part their account of these elements points to the same three factors already highlighted above: you feel yourself strongly *pulled* toward a certain moral judgment concerning a need to act, where this pull has its *source* in an *external* feature of the circumstances, in a strong way that *grips* you and seems to do so *independently* of your contingent desires or social conventions—insofar as the *content* of the background standards being employed makes no reference to such desires or social conventions in determining that you must act given the circumstances (2018, 135). As before, this accounts for *something*, and does so without impugning the experience as having non-veridical purport, but it still misses the further and more fundamental aspect of categorical authoritativeness we've identified.

This further aspect is best captured by what they refer to as "inescapability"—the experience that "becoming thus ought-committed is both involuntary and independent of [one's] pre-existing desires" (135). Or, as they put it: in a case where one violates the felt moral demand, "the authority one has contravened is experienced as operative upon oneself independently of any medium-term or long-term intentions one might have voluntarily formed, and independently of one's pre-existing desires" (137). This gets at the difference between Hit-and-Run and Struggling Café: in the latter you are aware that the reasons you're perceiving stem ultimately from background commitments you've voluntarily undertaken but could instead have declined without error; in the former it instead appears that the reasons you're perceiving stem from standards to which you are committed because they are objectively prescribed for you as inherently meriting your commitment. The question, then, is how, having recognized this appearance of inescapability, expressivism *accommodates* it—accounts for it *without* making it out to be misleading or erroneous.

Nowhere in the expressivist model are there any such facts about objective prescriptivity or meriting of commitment with respect to the basic moral standards themselves, whether in an ontologically involved realist sense or in a non-ontological Kantian rationalist sense. Instead, the story of our coming to be committed to a given set of standards will be a matter of contingent psychological and sociological influences on personal development, values and choices. It needn't, of course, *feel* that way, which is precisely what Horgan and Timmons acknowledge in recognizing that moral phenomenology involves a sense of *inescapability* as opposed to a transparent appreciation of the *actual* processes underlying the formation of one's moral commitments, which latter in fact yield "normative governance that obtains only contingently" (128). The problem, however, is that this now commits the expressivist to seeing the phenomenology as being *in error*.

The error here is not the one Horgan and Timmons explicitly disavow, i.e., an error rooted in the phenomenology's purporting to represent ontological, in-the-world, objective moral properties with categorical authority: again, I have not claimed that the ontological purport is built into the phenomenology surrounding the sort of moral case we've considered so far. The crucial phenomenology I've pointed to is a more general sense of inescapability, equally compatible with a non-ontological Kantian rationalism, according to which we are (somehow) subject to a normative governance that does *not* obtain only contingently. Since expressivism claims that in fact all normative governance ultimately obtains contingently (in the sense described above, concerning the true nature of our commitments and how we come by them), it therefore must take this aspect of the phenomenology to be *misleading*: the appearance is that the basic standards are (somehow) imposed on us independently of our contingent psychologies, and expressivism insists that this is not actually what is happening, which is instead purely sociological and psychological.

In other words, expressivism has not in fact *accommodated* inescapability: it seems instead to have *recognized* it and *explained it away*—much as an error theorist might explain it away as a non-veridical, misleading appearance (Horgan and Timmons 2018, 125–26); but that is to abandon the intended accommodation project. Of course, nothing I've said shows that the appearance of inescapability is in fact accurate, and neither a realist nor a Kantian rationalist can expect someone who rejects its accuracy to be swayed by an appeal to it. Such an appeal will obviously get traction only with those who are more convinced of the veridicality of the phenomenology than they are by arguments for being skeptical of it. What I think I have shown, though, is that if the expressivist does recognize the appearance of inescapability to belong to the moral phenomenology, and if I am right that this goes beyond the sort of desire-independence highlighted in the three factors laid out earlier, then the expressivist must renounce the project of fully accounting for the moral phenomenology: given their metaethical view, a central part of that phenomenology will have to be construed as misleading.[6]

[6] Horgan and Timmons might claim that the error in question isn't strictly one of "misrepresenting the world" (116), since it is not a matter of erroneous *ontological* objective purport, but only the more general error I've identified, which needn't be ontology-involving. But although it isn't a metaphysical misrepresentation it would still be non-veridical, misrepresenting what is in fact that case—in just the way a Kantian view would be in error by representing categorical authority as stemming from practical reason if in fact it does not do so, even though this error doesn't involve false metaphysical claims about "the world" and its properties. I take it the project of accommodation should not welcome either sort of error in the phenomenology.

By contrast, the realist captures the deeper categorical authority without deflation or construing the phenomenology as misleading. According to the realist model, an important part of our moral experience is that of apprehending the objective correctness and categorical bindingness of the basic standards themselves, by grasping their inherent fittingness to the world. This inherent fittingness of a certain set of standards to the world is not a psychological illusion or something we feel due to a contingent commitment we've made, but an objective fact about those standards consisting in the fact that they articulate inherently fitting responses to worldly values, such as the value of a child and her well-being, by calling for helpful action when confronted by a child in crisis (FitzPatrick 2008 and Forthcoming-a, Forthcoming-b). Whatever reservations one might have about the metaphysical commitments involved, realism at least captures the deeper layers of categorical authoritativeness here.

4 What Follows?

It would, of course, be too quick to conclude that the moral phenomenology points us directly to realism: for even if realism accounts for aspects of the phenomenology missed by expressivism, there may be other views that do so as well, without similar ontological purport, as Horgan and Timmons (2008) point out. Recall the two central phenomenological aspects of the demand here:

(A) it comes across as a *categorical* demand, binding you independently of what you happen to desire, with "categorical authority" (Horgan and Timmons 2018), and
(B) it seems to be "emanating from 'outside' [you]", having its *origin* or *source* in something external to yourself (Mandelbaum 1955, 54).

As we've noted, Kantian rationalism might capture A by appeal to an invariant set of moral standards springing from the rational will itself independently of our choices or concerns or other psychological and sociological contingencies. And while strictly speaking it thereby rejects the accuracy of B, since the standards are not external to yourself but have their source in your rational will, it might still be able to explain the appearance of external sourcing. It could do this by emphasizing the independence of the standards from your desires, attitudes and actual choices, along with the fact that the true sourcing in our own rational will is far from introspectively transparent to us, requiring complex philosophical work to illuminate. The externality claim would thus be somewhat deflated, but it might be thought that capturing categoricity and independence at least from contingent elements of one's psychology is enough—all the externality the appearances really give us if we're being careful.

Kantian rationalism would therefore remain on the table so far along with realism (and perhaps other alternatives, such as non-realist cognitivism, with its categoricity and independence even from the will, but without ontological commitment), when it comes to cohering with the introspectively ascertainable phenomenology. Let us turn, then, to the phenomenological case for realism in particular, with its more robust embrace of B above.

5 Realist Moral Phenomenology

Consider a different stretch of the phenomenological landscape, involving emotionally-laden moral responses to violations of core moral values.[7] I have in mind here moral distress and indignation upon observing someone's being egregiously abused, for example, and moral distress, guilt, shame and remorse upon having wronged and hurt someone oneself. These moral experiences, while obviously regrettable, often play an important role in moral formation and the development of concrete moral sensitivities. Here are two examples:

> **Bullying:** From a distance, you observe three teenagers stop a smaller, shy classmate walking home from school. They taunt him aggressively, and as he becomes upset they mock him for it, humiliating him. When he tries to get away they shove him repeatedly to the ground and finish by beating him and scattering his books in the mud, after which they run away laughing.
>
> **Lashing out:** You are having an argument with a family member with whom you are normally close. As it gets heated, there is an escalating exchange of harsh words. You find yourself shouting out something particularly hurtful (drawing on something they had confided to you in a moment of vulnerability), cruelly betraying their trust, and you further express your anger by tearing up a drawing they had made for you. No sooner do you put down the pieces than you see the deep hurt your words and actions have caused.

In Bullying, it is not merely that you make an immediate moral judgment that a moral norm has been violated—for example, that the bullies have acted on a maxim that cannot consistently be willed to be a universal law—and then have a powerful emotional response to that abstract judgment (non-universalizable maxim!). The phenomenology is far more concrete and victim-centered with respect to the apparent source of the moral distress and indignation. Your experience is likely to be one of witnessing—almost literally *seeing*—a violation of this person's dignity in a grossly unfitting encounter with what seems to be a real and present value abiding in him, treating him as if his personhood had no intrinsic significance and could appropriately be met with such behavior, as if he were a mere object to be used for amusement.

It is not just that the circumstances of this person's suffering and abuse are external to you, and they combine with moral standards you happen to accept to tell you that a wrong has been committed: you instead experience the *source* of that wrong as lying in the objective moral significance of that person, in the dignity you are seeing being violated, which is likewise the source (in the more general form of the dignity of persons) of the moral standard prohibiting such behavior as inherently unfitting. The emotionally-laden distress and indignation are the form taken by the apparent recognition of such wildly unfitting responses to real value encountered in the world. This, at any rate, is how things will seem to many of us.

Similarly, in Lashing Out the phenomenological content of the guilt, shame and remorse you feel at having so hurt this person presents itself as a recognition of having *encountered* and *violated* something precious through behavior that disrespected this person and their feelings by being inherently unfitting to their nature and significance. It's not just about basing moral judgments on certain objective features of the situation, but about experiencing the moral

[7] Horgan and Timmons (2008, 282) refer to these as "second-order" moral experiences, since they presuppose other moral judgments about the thing in question—as the experience of guilt piggybacks on the more basic judgment that one has acted wrongly. I take such experiences to be just as important to this debate as "first-order" ones.

reality here—through the vehicle of your guilt, shame and remorse—*as having its concrete source in real values you have encountered and failed to respect*. And this is precisely what the realist picture gives us: not merely categorical moral norms violated, but categorical moral norms violated through inherently inappropriate responses to the worldly values you have encountered. The moral emotions are a manifestation of the apparent recognition of having encountered important values and failed to respond appropriately to them—or worse, having responded to them in a way grossly unfitting to their nature.

Now it might seem ironic that the view I've chosen to contrast with the realist view is a Kantian rationalist one, which I've suggested—with a flip reference to the Formula of Universal Law—fails to capture the phenomenology of these cases. After all, Kantianism is the *locus classicus* of the idea that persons are ends-in-themselves, with a dignity that lies at the core of the moral law. Indeed, much of the most illuminating writing on dignity comes from neo-Kantian philosophers. But while Kantian treatments of dignity in the domain of normative ethics are insightful, what I want to target here is the anti-realist constructivist construals of them given in some (though not all) influential neo-Kantian metaethical accounts.

Christine Korsgaard is admirably clear in laying this out. While it may be tempting to read Kant as attributing an objective, realist value to persons (or to the rational nature in persons) as ends-in-themselves or possessors of dignity, she explicitly rejects such a reading. As she understands the Kantian picture, the world contains no such objective, real value to be discovered or encountered—not even in connection with persons. In this sense the world is as inherently value-free as it is for Hume. Nor is this an incidental feature of the view: it is, Korsgaard believes, essential to answering "the normative question." On her view, any authoritative values or normative principles must have their source *in your own will* rather than coming from outside of yourself; if they imposed themselves on you from outside your own will, as on a realist picture, they could never bind your will, she thinks—always leaving open the question why you should heed them (Korsgaard 1996, 2003).

Neo-Kantian constructivism seeks to explain the normativity of moral principles by doing two things. First, it reduces the normative bindingness of a principle to the fact of one's being *committed* to it, on pain of inconsistency if it is flouted, by practically necessary procedures constitutive of one's exercise of agency as such—i.e., by things one allegedly *must* do to function as an agent at all, such as *regarding* oneself, *qua* rational agent, as an inherently valuable source of value. Second, it then tries to show that rich moral principles can be derived precisely as such commitments undertaken through the unavoidable exercise of agency as such.[8]

The details of this anti-realist constructivist account needn't concern us here: the bottom line is that on this approach, the source of another person's value or dignity, insofar as it normatively binds your will and calls for action from you, as in Hit-and-Run, is *not* the person herself or anything outside of yourself: it is your own will, via the commitments you take on in the way you must 'regard' things as being simply in order to exercise your will at all in general. And despite the prima facie concessions made earlier, it's doubtful that such a picture captures even the phenomenology of the direct moral experience of obligation in Hit-and-Run, let alone the phenomenology in the sorts of case we are presently considering.

If what I've said about Bullying is correct, then your experience is one of seeing a gross violation of the victim's dignity—a value present in him that is being desecrated by this encounter. It's hard to see how we could take this phenomenology seriously while proposing

[8] I explore and critique Korsgaard's radically practical approach to normativity in FitzPatrick (2005 and 2013).

that what you are witnessing is ultimately just behavior that is *inconsistent* with a *commitment* the bullies have taken on just by exercising their agency at all. Such a story—according to which any exercise of agency commits the bullies to regarding themselves as valuable sources of value insofar as they are persons, and thus, to be consistent, to regarding others the same way and so not bullying them—may strike neo-Kantian constructivists as an attractive theoretical account of the moral violation involved in the bullying. But as an account of the *phenomenology* associated with the case it seems distinctly unpromising. By contrast, the realist view speaks directly and plausibly to that phenomenology.

Similar points apply to Lashing Out. It is theoretically possible, if not very plausible, that your moral guilt and shame at so hurting this person are ultimately *justified* by the fact that you have acted inconsistently by violating a principle to which you are generally committed by procedures you literally must carry out simply in order to exercise your will at all. But however that may be, such a picture is far removed from the phenomenology of that guilt and shame, which the neo-Kantian would have to regard as simply misleading. For the appearance involves the powerful sense that you have acted in an inherently unfitting way toward this person given the real significance of their personhood and of their feelings, which you have violated.

This phenomenology is structurally parallel to the horror one might experience after desecrating a work of art—encountering something of significant value, given its nature, and treating it in a way grossly unfitting to that nature and value. The realist model straightforwardly captures this character of the phenomenology in a way that rival views that instead locate the source of value and normativity *within oneself* fail to do. Perhaps there are other alternatives—such as non-realist cognitivism (Parfit 2011, 2015) or metaethical conceptualism (Cuneo and Shafer-Landau 2014)—that do so as well, locating the source of value suitably outside of the agent responding to it while avoiding ontological commitments to instantiated moral properties in the world. I've said nothing here to rule out such possibilities.[9] On the face of it, though, if I'm right about the phenomenology of these cases then it seems to point toward sources of value and normativity that not only lie genuinely outside of the agent encountering and gripped by their authority but that are also *worldly*, residing in real elements of the world with which we interact in ethical life—rather than existing abstractly on the model of logical, mathematical or conceptual truths. This suggests an advantage specifically for ontologically committed realism.

6 Conclusion

If the argument of the previous section is right then there are introspectively ascertainable aspects of moral phenomenology that carry ontological objective purport, directly supporting realism. Even setting this aside, however, the success of the earlier argument would still be enough to establish an advantage of realism over expressivism, and it would support realism at least as part of a broader class of theories that capture the deep inescapability that is part of moral phenomenology. While this would not itself support realism specifically, such attention to the phenomenology could still be a crucial part of the argumentative path toward realism.

[9] I've elsewhere raised worries about whether such views can plausibly make sense of categorical authoritativeness (FitzPatrick 2018 and Forthcoming-b).

Indeed, the best prospect for supporting the kind of robust realism we have been discussing may lie precisely in taking the moral phenomenology seriously, seeing which views accommodate it, and then working to eliminate rival views on other grounds (FitzPatrick, Forthcoming-b). For example, even if Kantian rationalism succeeds in capturing the full scope of phenomenology after all (*contra* the argument of the previous section), we might have independent grounds for rejecting it, such as doubts about the proposed derivations of rich and normatively authoritative moral principles from generic facts about the conditions for the exercise of agency as such (FitzPatrick 2005 and FitzPatrick, 2013).

This process would be an ongoing and open-ended one, taking on rivals as they arise and seeing how they fare against realism, all-things-considered, in light of broad metaethical reflection. It would be nothing like a single argument purporting to establish realism once and for all. I doubt, however, that we should expect there to be any such master argument or that realists should accept a demand to produce one. It may be that the only effective way to support the strong claims made by robust realism is the messier sort of process I've described, which is more likely to yield interesting metaethical results in any case. And an important driving force in that process is taking the moral phenomenology seriously, seeing where it takes us in light of ongoing metaethical reflection.[10]

References

Cuneo T, Shafer-Landau R (2014) The moral fixed points: new directions for moral non-naturalism. Philos Stud 171(3):399–443

Dancy J (2004) On the importance of making things right. Ratio 17:229–237

Dancy J (2006) Non-naturalism. In: Copp D (ed) The Oxford handbook of ethical theory. Oxford University Press, Oxford, pp 122–145

Egan A (2007) Quasi-realism and fundamental moral error. Australas J Philos 85(2):205–219

FitzPatrick WJ (2005) The practical turn in ethical theory: Korsgaard's constructivism, realism, and the nature of normativity. Ethics 115(4):651–691

FitzPatrick WJ (2008) Robust ethical realism, non-naturalism and normativity. In: Shafer-Landau R (ed) Oxford studies in Metaethics, vol 3. Oxford University Press, Oxford, pp 159–205

FitzPatrick WJ (2011) Ethical non-naturalism and normative properties. In: Brady M (ed) New waves in Metaethics. Palgrave MacMillan, New York, pp 7–35

FitzPatrick WJ (2013) How not to be an ethical constructivist: a critique of Korsgaard's neo-Kantian Constitutivism. In: Bagnoli C (ed) Constructivism in ethics. Cambridge University Press, Cambridge, pp 41–62

FitzPatrick WJ (2018) Ontology for an uncompromising ethical realism. Topoi 37(4):537–547

FitzPatrick WJ (Forthcoming-a) Ardent moral realism. In: Copp D and Bloomfield P (eds.) The Oxford Handbook of Moral Realism. Oxford University Press, Oxford

FitzPatrick WJ (Forthcoming-b) Ethical realism. Cambridge University Press, Cambridge

Hampton J (1998) The authority of reason. Cambridge University Press, Cambridge

Horgan T, Timmons M (2008) What can moral phenomenology tell us about moral objectivity? Soc Philos Policy 25:267–300

Horgan T, Timmons M (2018) The phenomenology of moral authority. In: Machuca D (ed) Moral skepticism: new essays. Routledge, New York, pp 115–140

Korsgaard C (1996) The sources of normativity. Cambridge University Press, Cambridge

Korsgaard C (2003) Realism and constructivism in Twentieth Century moral philosophy. J Philos Res 28(Supplement):99–122

Mandelbaum M (1955) The phenomenology of moral experience. The New Press, Glencoe\

[10] I am grateful to two anonymous referees for helpful comments.

Parfit D (2011) On what matters, vol 2. Oxford University Press, Oxford
Parfit D (2015) On what matters, vol 3. Oxford University Press, Oxford