

## Constitutivism's plight: inescapability, normativity, and relativism

OLOF LEFFLER 

University of Siena, Italy

*Constitutivists often argue that agency is inescapable. This is supposed to, among other things, explain why norms that are constitutive of agency are forceful. But can some form of inescapability do that? I consider four types of inescapability—psychological, further factor, standpoint, and plight—and evaluate whether they manage to explain four necessary features of normative force: that it does not vary with desire change, that ought-implies-can and can-fail, and that we are criticizable for failing to live up to forceful norms. The former three all fail to explain some of these features, but a version of plight inescapability does, in fact, appear able to explain them all. This is good news for constitutivists. The catch? The plight inescapability explanation generates relativism. This leaves constitutivists at a choice point: accept the plight inescapability explanation of normative force and relativism—or give up the plight inescapability explanation.*

**Keywords:** constitutivism; inescapability; normativity; normative force; relativism.

Constitutivism says, roughly, that some property which is a feature of our subjecthood explains or grounds why norms are normative for us: they apply to us and have normative force.<sup>1</sup> This is among the most promising features of constitutivism. One of the deepest mysteries of metaethics and metanormativity is to explain why norms have such force.<sup>2</sup> Constitutivism promises an answer.

**Correspondence to:** Olof Leffler, [olof.leffler@unisi.it](mailto:olof.leffler@unisi.it)

<sup>1</sup> I use (1) 'explain', 'ground', and 'be the source of' as well as (2) 'have force' and 'bind' interchangeably in this paper.

<sup>2</sup> From here and on, I leave the question of which norms apply to the side.

Constitutivists are, accordingly, often concerned with what recently has been dubbed ‘the Arché question’—the question of the ground of norms (cf. Korsgaard 1996; Bengson, Cuneo and Shafer-Landau 2023). Yet, it has increasingly become recognized that constitutivism seems unable to explain normative force just by saying that some norms are constitutive of our subjecthood. If anything, constitutive norms rather serve to *transmit* normative force from elsewhere than to bind by themselves (Ferrero 2018; 2019; cf. Leffler 2024). But we also want to know *why* these norms are binding: a norm is not forceful just because it is a feature of our constitution. Rather, we need to show why we are subject to some possible norms rather than others.

A number of hypotheses about why constitutive norms are normative have appeared. Perhaps because of their value (Smith 2015). Or perhaps their fittingness (Worsnip 2021). Or perhaps they hold for imperfect creatures with temptations to do otherwise (Schafer 2023). But most have tried to argue that the norms are, in various ways, inescapable (e.g. Korsgaard 2009; Velleman 2009; Katsafanas 2013; Walden 2018; Ferrero 2018, 2019).

Putting other possibilities to the side, I aim to discuss whether some inescapability proposal from this literature, charitably construed, can explain why norms are forceful by way of capturing four necessary features of normative force and the consequences of that. I first consider psychological, further factor, and standpoint inescapability. They fail to capture some of the features. But then I take on Christine Korsgaard’s plight inescapability. Charitably reinterpreted, it can make good on the features. However, it entails that normative force is relative to the norms that are implicit in the agency some agent already has, so agents with different constitutions are bound by different norms. Hence, plight inescapability generates relativism. This yields a choice point: accept the plight inescapability explanation plus relativism—or deny the explanation.

To show this, I start in Sections (I) and (II) with some housekeeping: I distil the relevant explanatory concern from the debate surrounding shmagency, suggest that some forms of inescapability from the shmagency debate are unhelpful, and set out some necessary features of normative force as explanatory *desiderata*. These are desire change invariance, ought-implies-can and can-fail, and the criticizability of failure. Then, in Section (III), I discuss psychological inescapability, in Section (IV) further factor inescapability, in Section (V) standpoint inescapability, and in Section (VI) plight inescapability. In Section (VII), however, I consider whether plight inescapability can explain categorical norms and argue that it entails relativism. I conclude in Section (VIII).

But first a disclaimer for clarity. This paper is exploratory: it aims to evaluate the potential success of extant inescapability proposals, charitably construed, at explaining why norms are forceful by way of capturing four necessary features of normative force and the consequences of that. I assume that the forms of inescapability and constitutivism I discuss are plausible enough

to be worth discussing, but I do not ultimately defend them as true. Instead, my conclusion is a choice point. One may accept the plight inescapability explanation of normative force and become committed to relativism. Or one may deny the plight inescapability explanation: perhaps because it entails relativism, or perhaps for some other reason, such as finding even the charitably reinterpreted version of plight inescapability or the types of constitutivism that might benefit from it implausible for independent reasons. Then one no longer has to go with the relativism that plight inescapability entails. Regardless, we learn much about the explanatory prospects of inescapability along the way.

### I. Shmagency, explanations, and inescapability

Again, many constitutivists aim to explain the applicability and normative force of norms implicit in our subjecthood. But constitutivism is a broad church. What is implicit in our subjecthood need not just be norms that are constitutive of being an agent, let alone a practical agent, in some narrow sense. It could also be things constitutive of our actions, (systems of) propositional attitudes, rational capacities, and much else besides. However, for simplicity, I use the label 'agency' in a broad sense below to refer to these varying features of our subjecthood.<sup>3</sup>

Constitutivism also covers many normative domains. Views range from those where moral norms are constitutive of agency, such as on Korsgaard's Kantianism (1996; 2009), to views where dispositions of structural rationality are constitutive of (systems of) propositional attitudes (Brunero 2020; Worsnip 2021) or views where reason as a capacity aims at understanding (Schafer 2023). Moreover, some views feature constitutive principles, such as the categorical imperative or requirements of means-ends coherence, and others aims, such as a desire for self-understanding.<sup>4</sup> I am mostly silent on these distinctions too. What matters is just that these things are implicit in features of our subjecthood—or agency, in the broad stipulated sense. I call the normative features implicit in or explained by our agency, broadly understood, *commitments* of agency. The characterizations of the main types of inescapability below make clear that many, though sometimes varying, accounts and commitments of agency are compatible with them.

<sup>3</sup>Some, such as Schafer, contrast capacities-based versions of constitutivism with those that focus on agency, narrowly construed. The distinction does not matter here: one could try to explain normativity using inescapability on both views. True, for his capacities constitutivism, Schafer adopts a Kantian explanation where a constitutive principle is normative because it is held by imperfect beings (Schafer 2023: 106–12). But capacities constitutivists who do without that explanatory hypothesis may go for some inescapability proposal instead.

<sup>4</sup>For the distinction, see Katsafanas (2013, 2018; cf. Ferrero 2019).

Constitutivism is, however, famously challenged by the shmagency objection. It is charged with showing why we cannot escape norms by being ‘shmagents’ rather than agents: similar to agents, but with other constitutions and commitments (Enoch 2006, 2011; Leffler 2019, 2024). In response, many have argued that agency is inescapable, popularizing the notion of ‘inescapability’ in the literature (Ferrero 2009, 2018; Korsgaard 2009; Velleman 2009; Katsafanas 2013). But what we are after here is inescapability as an explanation of normative force.

To clarify what kind of inescapability matters for that, I start off by showing how the most familiar version of the shmagency objection overlaps with, yet is distinct from, explanatory concerns, and how only some of the types of inescapability discussed in the literature serve our purposes. We may, at least for now, dismiss forms of inescapability that appear irrelevant.

I start with the disentangling. As Luca Ferrero has pointed out, constitutivists appeal to inescapability both to respond to the shmagency objection and to explain normative force. While closely connected, these projects are distinct.

How come? Most interpretations of the shmagency objection take it to concern the possibility of scepticism about constitutive norms (cf. Enoch 2006, 2011; Ferrero 2009, 2018; Leffler 2019, 2024). One may ask the shmagency question if one is an agent or if one is not, yielding *internal* or *external* scepticism, respectively. Internal shmagency is ordinarily not considered to be problematic. Normatively committed agents ask what their normative commitments are when they ask the internal shmagency question, but they are still subject to the commitments. External shmagency is more threatening, for it seems to open up a space outside agency. It challenges the extensional adequacy of constitutivism. Constitutivism does not appear to ground norms for a sufficiently wide range of entities, since one may ask about reasons to be an agent even without having reasons grounded in one’s agency (as one could be a shmagent).

However, this type of shmagency is often thought to be defused by some kinds of inescapability. In particular, dialectical inescapability is often brought up here (Enoch 2011; Ferrero 2009, 2018). The idea is that acting, including by asking the shmagency question, involves committing oneself to the norms of agency. This makes it self-defeating not to live up to them.<sup>5</sup> However, the risk of self-defeat may not appear to help ground anything positive, at least *prima*

<sup>5</sup>This label, and the ‘asking the shmagency question’ terminology, is confusing. While other interpretations are possible, I think dialectical inescapability is best understood as a conceptual or metaphysical necessity of agency (depending on whether constitutivism holds conceptually or metaphysically). For Ferrero argues for it by taking agency to be the enterprise of the largest jurisdiction, so all actions involve it, and closed under reflection, so even reflection counts as acting (Ferrero 2009: 308–10). These points turn on conceptual or metaphysical properties of agency, not on whether anyone *actually* deliberates about or discusses being an agent.

*facie*. It shows that there is no option to agency for agents, but it is unclear how that explains anything normative. One could be a sophisticated shmagent for whom we should explain norms without having been an agent at first, and then one does not seem subject to the norms of agency (Leffler 2019, 2024).

The second type of inescapability is normative inescapability (Leffler 2019, 2024). Here, the idea is that some type of agency is normatively valuable, justifiable, or otherwise such that we have reason to be that type of agent, making agency normatively necessary for us. But taking normative inescapability to explain why we should adhere to the norms of agency just pushes the explanatory question a step backwards. What grounds the force of the norm that makes agency and its commitments normatively inescapable?

Extensional adequacy aside, however, there is a further question implicit in some early formulations of the shmagency worry. This is whether constitutivists show how the constitution of agency grounds norms *at all* (Enoch 2006, 2011; cf. Lavin 2017). That question is independent of the extensional shmagency objection, for it is possible to explain the force of some norm for some agents (or, indeed, shmagents) using properties of agency as long as those agents (or shmagents) have the force-grounding properties, regardless of whether it holds for others. Hence, I like to treat this worry separately from shmagency. Still, due to its background in the shmagency literature, I would not quibble with someone who calls it an explanatory shmagency objection, as long as it is distinguished from the extensional one. That issue is just terminological.

But how is it different? Assume that the commitments of agency are inescapable in some other way than dialectically or normatively. Due to this form of inescapability, the norms constitutive of agency may be forceful for some agents *independently* of whether they are so for others. Maybe humans and animals have wildly different psychologies, and the norms implicit in human psychologies are forceful for us, while animals have different constitutions from us but still qualify as agents in some other sense (Korsgaard 2009). Then maybe something about *our* agency explains why our norms are normative for us, quite independently of whether animals are subject to norms and of the extensional shmagency point. Hence, we can ask the Arché question about these norms without being concerned with how many possible agents there are, or even with whether there are shmagents.

It is to answer this Arché question that constitutivists often have discussed other forms of inescapability than dialectical and normative inescapability. More specifically, they have emphasized psychological (Velleman 2009; Katsafanas 2013), further factor (Walden 2018), standpoint (Ferrero 2019), and plight inescapability (Korsgaard 2009; Katsafanas 2013). These are, therefore, the candidate explanatory inescapability proposals about normative force that I shall focus on.

But before that, a *caveat*. Agency *is* also escapable in many ways on these views (cf. Ferrero 2018). This is because it is not plausibly necessary for the entities who typically instantiate agency to be agents on standard types of necessity. I was not an agent on conception, and I will not be one after death. Hence, it is conceptually, metaphysically, and nomologically possible for the entity which instantiates my agency to not be an agent. It may even be psychologically possible. For example, assume with Velleman (1989 [2007], 2000, 2009) that a desire for self-understanding is partially constitutive of agency. It seems psychologically possible to lose that desire while maintaining other mental states, and hence to have a psychology which escapes agency. However, other things could be, in various ways and with various types of necessity, inescapable *for agents*. I specify them below when assessing the explanatory power of psychological, further factor, standpoint and plight inescapability.

## II. Normative force

What, then, is the normative force that is an *explanandum* for inescapability? This question is notoriously elusive.<sup>6</sup> Hence, rather than giving a full account, I shall present some familiar properties that ordinarily are associated with it. I take them to be necessary features of normative force, and therefore treat them as explanatory *desiderata*. There might be some types of norms that lack them, but if there are, they are not forceful (in, at least, the standard sense). But if some type of inescapability can be shown to capture all the necessary features, it is a promising contender for explaining the force of norms. The features are:

(*INVARIANCE*) To have normative force, a norm cannot shift merely because agents' desires shift.

This is because norms cannot be avoided just because agents would start to desire something else than to follow them. In at least one sense, the norm therefore applies independently of what we desire. *INVARIANCE* holds both for individuals and larger groups, for both may have desires that go wrong.

However, *INVARIANCE* does not rule out desires from being part of the explanation of normative force. Perhaps a desire for self-understanding (Velleman 1989 [2007], 2000, 2009) or desires to help and not hinder agents (Smith 2012) can feature in the explanation. It would be question-begging to rule out such views. Rather, *INVARIANCE* rules out shifts in norms that depend *merely* on shifts in desires, meaning that we lack a story about why the new desires

<sup>6</sup>For surveys of normativity, see Finlay (2019) and Parfit (2006).

would be preferable in the explanation of force. Such shifts open the door for potentially wrongful desires.

(*OIC*) Ought-implies-can.

Ought-implies-can (*OIC*) is often thought platitudinous: if something is a norm, agents can follow it. Hence, I shall assume *OIC*. But while there is much controversy about how it should be interpreted, and different philosophers reasonably may want to treat it differently, I shall leave the exact formulation of *OIC* open. It will be enough to discuss cases that intuitively clearly fit or fail to fit *OIC* below, regardless of how we characterize its details.

(*OICF*) Ought-implies-can-fail.

It also seems like an agent has to be able to *fail* to live up to a norm for it to be a norm (Korsgaard 1997). This is because we are in the business of evaluating, not just of describing: we need to be able to say how an agent who follows a norm does so successfully rather than unsuccessfully. Hence, ought implies can and *can fail*.

(*CRITICIZABILITY*) An agent who fails to live up to a norm is criticizable in virtue of failing to live up to the norm.

The point behind criticizability is that if there is no criticizability inherent in a norm, it seems idle. Much like how the norms of the Swedish language are idle for me when I write this paper in English, a norm that does not explain criticizability seems to have no force for me. Hence, *CRITICIZABILITY* should also be captured.

In sum, *INVARIANCE*, *OIC*, *OICF*, and *CRITICIZABILITY* are the necessary features of normative force that I will use to assess the types of inescapability. They are, however, necessary features of normative force, not necessarily sufficient. It is possible that normative force has other features too. But, for now, showing how inescapability could capture them all will take us far.

Or will it? Here, the reader may be troubled by worries related to explanatory reduction. There is the open question argument, and the naturalistic fallacy, and is/ought-distinction (Hume 1739–40; Moore 1903). Would not explaining these features of normative force by saying that agency is in some sense inescapable run straight into such concerns?<sup>7</sup>

These are big worries—but we can avoid them. There is a happy convergence here between recent work on the grounds of normativity and the relation between constitutivism and other metaethical views. It has increasingly been recognized that the question of the ground of normativity is for

<sup>7</sup>Indeed, some formulations in Enoch's version of the shmagency objection are versions of these worries (cf. Enoch 2011: 211, 216).

everyone, not just constitutivists (Bengson, Cuneo and Shafer-Landau 2023). Simultaneously, there are forms of constitutivism that fit different metaethical frameworks about moral discourse: for example, Smith (2012) is a cognitivist, Ridge (2018) is an expressivist, and Korsgaard (2003) is a constructivist. A full metaethical view should address both grounds and discourse, but explanatory reduction worries concern the latter. This makes them orthogonal to the Arché question.

### III. Psychological inescapability

Let us, instead, evaluate the relevant approaches to inescapability. The first is *psychological inescapability* (Ferrero 2018: 139–44), sometimes also called ‘natural inescapability’ (Velleman 2009: 136).<sup>8</sup> The idea is that some psychological feature stands behind all human actions (broadly construed, including omissions and refrainings). For any first-order aim we may have, satisfying this feature is a higher-order aim. Hence, the psychological feature is also inescapable: as all actions have aims, we higher order aim to satisfy it in all actions. This, one may think, has normative implications. However, the modality of psychological inescapability should *not* be conflated with psychological necessity. It could be interpreted in many ways, but the most natural interpretation is metaphysical, as it usually is defended as partially constitutive of the nature of agency.

The most familiar defences are probably due to David Velleman and Paul Katsafanas. They take certain motives to be psychologically inescapable. I use their views to exemplify psychological inescapability, though *mutatis mutandis*, other things could well be psychologically inescapable too, such as various capacities. For Velleman, however, the motive is self-understanding. It serves as a higher order aim relative to our first-order aims, and it grounds our reasons for action, since a reason for action is a consideration in virtue of which an action makes sense (Velleman 2009: 146). Katsafanas, instead, thinks there are two aims of agency: activity and will to power. The latter is the relevant aim here. The will to power, or a will to overcome resistance, is taken to be implicit in the structure of drive-motivated action, and all actions are motivated by drives. Hence, it is a standard based on which we can assess other values and reasons (Katsafanas 2013).

Can the alleged psychological inescapability of such features of agency capture normative force? Start with *INVARIANCE*. A constitutive aim is different from other aims. One can have many different first-order aims, so the constitutive aim can be a higher order aim relative to many different first-order

<sup>8</sup>Ferrero considers two types of psychological inescapability: motives so entrenched in a psychology that the agent cannot avoid acting on them, making questions of not acting on them moot, and motives that are ‘impossible to remove *short of* ceasing to exist as an agent’ (Ferrero 2018: 143, his emphasis). Both are compatible with the idea.



ones. And it does not go away with changing first-order desires. At first glance, then, psychological inescapability appears able to handle *INVARIANCE*.

However, *INVARIANCE* says that a norm cannot shift merely because agents' desires shift. Yet changing aims clearly alters normative force. If one were to be motivated by desires that lack connections to self-understanding or will to power, self-understanding and will to power would no longer be motives behind those actions. To be sure, one may not count as an agent on Velleman's or Katsafanas' views if one were to act (or 'shmacro') without these motives. But so what? Being a non-agent, or a shmagent, sometimes seems valuable. If an evil demon threatens to torture an agent for all eternity unless she becomes a shmagent for two minutes, shmagency seems the way to go.

Moreover, this point generalizes. Perhaps one no longer would count as having a mind (or otherwise lose something else that is deeper than agency) without the psychologically inescapable motive. But even losing that could be valuable in the right circumstances. If an evil demon were to threaten to torture an agent for all eternity unless she no longer had a mind for two minutes, then so much the worse for the mind. Psychological inescapability does not capture *INVARIANCE* if we are prepared to give up our constitutive motives and be shmagents. But, as the examples indicate, we should be so prepared.

*OIC* is, however, easier to handle. On psychological inescapability, norms are explained by inescapable motives. Trivially, if one inescapably has a motive, one *can* have it. *OICF* is only marginally more complicated. The psychologically inescapable motives only *motivate*, they do not guarantee that actions based on them will be successful. For example, aiming at self-understanding or power need not guarantee success. Hence, *OICF* is captured alongside *OIC*.

We reach *CRITICIZABILITY*. Why would one be criticizable for failing to live up to some motive one has? The idea is presumably that having it sets up a standard of success. We can live up to it or fail to do so. But we may also wonder why any particular aim *matters*, and therefore is such that failing to live up to it grounds criticism. Why satisfy one constitutive motive rather than another? For example, assume that there are two possible psychologies: one that aims at self-understanding and one that aims at power. Why have the former rather than the latter? If we cannot answer this question, it is hard to direct genuine criticism at those who aim to satisfy the other standard. And psychological inescapability does not answer it. It only says that some motive is a higher-order aim for all agents. But why not be a shmagent and have another motive? As per the discussion of *INVARIANCE*, that may sometimes be valuable.

To explain normative force with psychological inescapability, we would have to explain why some constitutive motive is preferable to others. Perhaps we could appeal to some deeper type of inescapability than psychological inescapability to do so. But this means that psychological inescapability falls back into other explanatory factors to show why norms are forceful for us.

Hence, psychological inescapability seems unable to do the relevant explanatory work, at least on its own.<sup>9</sup> I conclude that it fails: while it can capture *OIC* and *OICF*, it cannot capture *INVARIANCE* and *CRITICIZABILITY*.

#### IV. Further factor inescapability

Another possibility is inspired by Kenneth Walden (2018). Walden tries to stake out a middle ground between Kantian and Humean forms of constructivism. To do so, his thought is that there are logically, metaphysically, and nomologically contingent factors that nevertheless have impact on the explanation of our first-order norms, enabling and foreclosing certain possible courses of action. Walden calls these ‘further factors’ from the operation of practical reason as such.

What, then, is further factor inescapability? A further factor is one of the contingent factors just mentioned. They are inescapable in the (weak) sense that an individual agent cannot do away with them on her own, at least at a time. They are not in principle impossible to change for agents, but one single agent cannot do so at the point in time of action. As such, their modality is not heavy duty like metaphysical or logical necessity: they are *practically* necessary, enabling and foreclosing certain courses of action (Walden 2018: 140–1). *Mutatis mutandis*, they could presumably also enable or foreclose certain combinations of mental states, exercises of capacities, and similar. What matters is that they impact the features or commitments of our agency.

The further factors are, however, not typical features or commitments of our agency. Again, it is factors that affect practical reason *not* as such that are further factor inescapable. Walden gives two examples: socio-cultural schemata and embodiment. The former are ‘clusters of culturally shared concepts, beliefs, and other attitudes that enable us to interpret and organize information and coordinate action, thought, and affect. Schemas are public (...) but are also internalized and guide behaviour’ (Haslanger 2016: 126; quoted in Walden 2018: 127–8). They structure what an agent sees as possible, penetrate her perceptual and cognitive architecture, are relatively resistant to change and updates, and cannot be understood individualistically. The latter is our embodiment, or physical capacities. They, too, shape our possibilities.

For now, we may agree with Walden on taking schemata and embodiment to be relevant further factors. And, like him, I shall primarily use social schemata to illustrate my arguments—though I sometimes write ‘inescapabilities’ in the plural to indicate where more than one may be at work. Then, the question is this: can further factor inescapability capture our *desiderata*?

<sup>9</sup>Interestingly, Velleman also appeals to dialectical inescapability (Velleman 2009: 137–45) and Katsafanas to plight inescapability (Katsafanas 2013: 52–7).

Start with *INVARIANCE*. Walden suggests that 'by locating some of the normative significance of morality in agents' background framework of practical thought, we are able to explain why, in certain instances, the normative force of morality is unconditioned by agents' particular evaluative attitudes' (Walden 2018: 149). Hence, further factor inescapability seems like it can capture *INVARIANCE*. An agent could desire *X* or *Y*, but nevertheless be subject to moral norms which have their force based on further factors such as social schemata.

But it is not enough to say that the further factors may have impact on normative force in certain instances. If the desires of sufficiently many people in some social setting change, and the new desires are sufficiently strong, that is also likely to impact the social structures implicit in social schemata. Their desires will make them give some up and take on new ones. This means that schemata are not invariant. Of course, sometimes, not enough desires change for social patterns to change. But it is easy to think of cases where enough desires *do* change to affect schemata.

What about *OIC*? There are puzzling implications here. Consider a society with homophobic social schemata. Homosexual desires are unlikely to disappear for at least some agents in that society, despite their social unacceptability. Assume also that we should think of an agent's reasons for action as grounded in her rationalized desires filtered through the further factor inescapabilities she faces. Then, an agent with homosexual desires need not attain any reason from her desires, since they clash with the social schemata. However, the agent's own desires may then clash with the reasons she ends up with to the extent that she cannot act on the reasons. Perhaps the agent ends up with a reason to find a spouse of the opposite gender while only desiring one of the same gender. As such, adding social schemata as an explanatory factor to a common view of reasons seems to give us *OIC*-style worries.

However, Walden also writes that 'schemata are not deliberative straitjackets (...) the right circumstances and a strong enough desire (...) can cancel [their] effects' (Walden 2018: 149). Perhaps sufficiently strong desires are not subject to *OIC* concerns: Walden does not say this, but he does not seem to think that reasons are grounded in rationalized desires filtered through further factor inescapabilities. Maybe we should think that reasons are grounded in an agent's rationalized desires plus further factor inescapabilities, *ceteris paribus*, where circumstances and desire strength can cancel the impact of the inescapabilities. But why think that? Here, defenders of further factor inescapability are challenged to show how we can explain normative standards without *OIC* concerns. It is an open question whether that will work.

If they can meet that challenge, there is at least an upside. It seems straightforwardly possible to fail to live up to what is suggested by further factor inescapabilities, and therefore to capture *OICF*. The homophobic society case above clearly shows that.

This takes us to *CRITICIZABILITY*. Unfortunately, further factor inescapability does very poorly here. Why would anyone be criticizable for failing to live up to normatively arbitrary or even oppressive social schemata, such as the homophobic one? Walden himself notes that schemata might require social critique (Walden 2018: 151). But that indicates that they are unable to explain *CRITICIZABILITY*. Sometimes the schemata are wrong, not the individuals facing them.

Summarizing, further factor inescapability does not seem to help with normative force. It has clear problems with *INVARIANCE*, *OIC*, and *CRITICIZABILITY*. Perhaps some other further factor could do better than social schemata. But I am not sure which.

## V. Standpoint inescapability

A third type of inescapability has been suggested by Ferrero (2019). We saw in Section (I) above that the shmagency question may be asked internally or externally. The internal question is unproblematic for constitutivists. But if there is no standpoint external to agency from which we can evaluate the external question absent self-defeat, as dialectical inescapability suggests, perhaps the external version is unproblematic too (Ferrero 2009, 2018)?

This idea lays the ground for *standpoint inescapability*. As I interpret it, here it is the practical standpoints from which we act (broadly construed, including omitting or refraining from action) that are inescapable. The idea is that, successfully or not, agents always exercise their agency (by acting, omitting, or refraining) from there to some degree. Hence, our practical standpoints are always operative as a condition or constraint on agents. As it is based on dialectical inescapability, standpoint inescapability plausibly has a similar modality: most likely metaphysical or conceptual necessity (cf. Section I, fn. 5). Moreover, if we construe agency in terms of capacities, the exercise of the capacities plausibly comes from an analogous standpoint of exercising those capacities.

The question, then, is whether standpoint inescapability explains our *desiderata* about normative force for some norms implicit in or explained by agency. I use familiar Kantian norms in my examples below for simplicity, but feel free to assume others. Unfortunately, standpoint inescapability is premised on a reply to shmagency from dialectical inescapability, which may seem to lack positive explanatory force (cf. Section (I) above). In fact, even Ferrero (2019) himself raises two worries for standpoint inescapability. The first is similar to mine: it is unclear how standpoint inescapability could ground normative force positively. Second, the type of agency to which there is no alternative could be very minimal, which rules out ambitious forms of constitutivism.

However, it is still fruitful to evaluate standpoint inescapability using my *desiderata*. Capturing them could *ipso facto* be what allows it to explain normative force. The explanatory worry is solved on the cheap if standpoint inescapability is explanatorily powerful. And the minimal agency worry need not be decisive. Explaining the force of, say, the hypothetical imperative without the categorical imperative is still a significant result. Hence, standpoint inescapability may still be promising.

In fact, standpoint inescapability does *prima facie* remarkably well. It is unexpectedly explanatorily powerful. Start with *INVARIANCE*. Standpoint inescapability does not shift with agents' desires. If there is no standpoint external to agency, then all actions are internal to agency, regardless of what agents may desire. It is also easy to capture *OIC*. If all our actions are internal to our practical standpoints, we can obviously act from the standpoints.

But standpoint inescapability struggles with *OICF*. It should be possible to fail to live up to all norms our constitutions as agents commit us to without stopping to be committed to those norms. But here standpoint inescapability no longer counts us as agents. For example, if the categorical imperative is the moral law and the hypothetical imperative the law of instrumental rationality, these norms are constitutive of agency, and no other norms matter, it should be possible to fail to live up to both at once while still remaining committed to them due to our constitution. Assume that it is wrong to lie, and to lie successfully about why I am late to a meeting I need to say that my car broke down, but I say that I am late because it is raining. The norms still hold for me.

A standpoint inescapability defender cannot say that, for the view entails that we cannot live up to everything we are committed to as a feature of our constitution as agents and still be agents. Doing that would place us outside our standpoints: as our commitments are constitutive of our agency, we would no longer be inside the standpoint of agency when we fail fully. But I am still an agent if I fail to lie, even though I seem to have reached a space where I live up to neither of the two allegedly standpoint inescapable imperatives constitutive of agency.

This problem is an instance of the problem of bad action for constitutivism (e.g. Lavin 2004; Korsgaard 2009; Lindeman 2017; Coleman 2023). If agency entails living up to some norms, at least to some extent, how can there be bad agents: agents to whom norms apply but who entirely fail to live up to them? This problem has been discussed at length by constitutivists, but it is particularly concerning for standpoint inescapability.

Not least, the most familiar solution to the problem, so-called threshold constitutivism, is unavailable to defenders of standpoint inescapability. On threshold constitutivism, agency comes in degrees and can therefore be more or less fully exercised, but all exercises of agency above a minimum threshold of degrees count as actions. This allows agents or actions constituted by norms

to fail: failures are imperfect exercises above the threshold (Korsgaard 2009; cf. Lindeman 2017 for criticism and Coleman 2023 for defence). However, the failed lie case suggests that it is possible to fail to live up to *any* of the constitutive norms of agency to *any* degree and still be subject to constitutive norms. Here, one would not be exercising agency at all, and hence would also end up outside the standpoint of agency. But as the case shows, we can fail to live up to constitutive norms to any degree. Standpoint constitutivists cannot take such failures to be possible.

Finally, we reach *CRITICIZABILITY*. Standpoint inescapability theorists also struggle with the criticizability of non-agents who occupy a standpoint outside agency. This point is related to the problem of bad action. Suppose that it is impossible to perform an action which fails in all its normative dimensions, as one then no longer is an agent. Then one no longer seems criticizable for failing to perform it, since what one does when failing in all normative dimensions is not an action. For example, if the categorical and hypothetical imperatives are constitutive of action, I should be criticizable for lying *and* for failing to lie when I attempt to do so. But standpoint inescapability says that I am not acting, and hence am not criticizable, when I fail to live up to both norms. Therefore, standpoint inescapability fails to show how agents can be criticized when failing fully. Accordingly, alongside *OICF*, standpoint inescapability fails to capture *CRITICIZABILITY*.

## VI. Plight inescapability

I have considered psychological, further fact, and standpoint inescapability. They all fail to capture various *desiderata*. But Korsgaard's plight inescapability remains (Korsgaard 2009: 1–2). I now turn to it.<sup>10</sup> Unfortunately, Korsgaard is not very clear about what it involves. She takes acting, in the broad sense which includes omissions and refrainings, to be something agents necessarily are in a position to do (or are 'faced' with). While it is possible to fail to act, she thinks it is impossible to undertake to fail absent 'acting, in a wonderfully double sense of the word' (Korsgaard 2009: 1). Seemingly, undertaking to fail involves performing an action: faking that one is not acting. Hence, as long as an agent is not, as she puts it, 'derailed' from acting, the agent must act. But acting is not logically, causally, or rationally necessary. It has a kind of practical necessity (or 'plight'). As a necessity is inescapable, action is *plight inescapable*.

A helpful, and more theoretical, interpretation of these brief remarks comes from Ferrero (2018). He thinks that plight inescapability is a two-way power 'to either make an antagonistic intervention in the natural course of events or

<sup>10</sup>See also Leffler (2023). My interpretation here builds on and updates that one.

to refrain from such intervention?—unless, to accommodate the possibility of derailment, the agent is in a ‘a dormant state’ (Ferrero 2018: 135).

However, Ferrero’s interpretation has two problems. First, two-way powers are contentious. What the relation is between powers, capacities, abilities, dispositions, or other potentialities and action is unclear. We should be able to make Korsgaard’s points with less metaphysical commitment. Second, the notion of a ‘dormant state’ seems unhelpful. Etymologically, it suggests being asleep, but there are more ways for agency to be derailed than that. A stone could fall on one’s head, or one may have tied oneself to the mast. We should be more informative.

Developing Ferrero’s interpretation, I propose that we view plight inescapability as an agent’s being positioned to continuously exercise her agency fully, absent external limits at the point in time of the exercise. Here, plight inescapability is a *constraint* that agents face when they attempt to act (or not to act). For some commitment of our agency to be plight inescapable, then, is for us to be continuously positioned to exercise that commitment fully, absent external limits at the point in time of the exercise. However, *what* we are committed to depends on what the nature of agency is: perhaps principles (e.g. the categorical or hypothetical imperatives), aims (e.g. satisfying a desire of self-understanding), or some other end-state (e.g. understanding, if reason is a capacity for that).

To make full sense of this proposal, we need to explicate what it is to (1) be positioned to exercise agency, (2) do so continuously and (3) fully, (4) what an external limit at the point in time of the exercise is, and (5) how plight inescapability is kind of practical necessity. I shall now do that, simultaneously showing how this suggestion improves on Ferrero’s two-way power interpretation.

(1) Start with being positioned to exercise agency. Think of agency as a potentiality on part of the agent: for example, an ability, power, capacity, or tendency, which the agent may exercise (but not necessarily exercises). Then the agent may exercise her agency given the opportunity to do so. The fact that she may exercise it given the opportunity *is* her being positioned to do so. But the agent could also fail to exercise her agency, partially or fully, for whatever reason: perhaps because of some mild hiccup (e.g. mild forgetfulness) or perhaps because of some external limit (e.g. a rock hitting the agent’s head). I return to this point in (4) and when discussing *OICF* below.

The language of ‘exercising agency’ is neutral between acting or refraining from or omitting action. Hence, we can be metaphysically neutral on whether to treat refrainings or omissions as types of action. More generally, it does not commit us to any particular metaphysical interpretation of what kind of potentiality agency is, whether an ability, power, capacity, tendency, etc. Hence, *contra* Ferrero, we are neutral on two-way powers. What matters for

constitutivism is just that agents are positioned to exercise their agency (and its commitments).

(2) What is it to exercise agency ‘continuously’? It means that agents are positioned to exercise agency at each new point in time. New possibilities appear regardless of what agents want, and, as Korsgaard emphasizes, undertaking not to act is still an action. The latter point is similar to Ferrero’s argument for dialectical inescapability (cf. Section I, fn. 5). But here, it rather serves to emphasize that agents always are positioned to exercise agency, regardless of what they try to do.

(3) Agents are in a position to exercise agency fully. This means that they are positioned to exercise agency successfully relative to what ‘the thing to do’ is. ‘The thing to do’ is that which, among all possible alternative options about what to do at some point in time, is supported and determined by (all of) the norms that have a bearing on that exercise of agency. Here, ‘norms’ are construed broadly, including reasons, values, constitutive principles, or whatever else matters. If they support or determine more than one alternative equally, the agent may just pick one.

Importantly, ‘having a bearing’ varies with context. Norms an agent cannot live up to do not do that, as ought implies can. Hence, agents are only positioned to exercise their agency fully relative to the standard of success in ‘the thing to do’ for some particular exercise of agency, which is context-relative. Assume again that the two norms constitutive of my agency are the categorical and hypothetical imperatives, there are no other norms, and nothing stops me from acting on them now. Then they determine what the thing to do is for me, so I am positioned to exercise my agency fully successfully by living up to them. But if I were to lose the ability to act on either one or some requirements they entail, what I lose the ability to do no longer affects the thing to do. Hence, the agency I am continuously faced with enacting is only that which is within my power to enact.

(4) What is an ‘external limit at the point in time of the exercise’? It is a limit external to some token exercise of agency (at the point in time at which it is exercised), which makes it very hard or impossible to realize certain opportunities with that token exercise. Hence, it is an obstacle that gets in the way of a potential exercise of agency.<sup>11</sup> Obstacles can be independent of the agent’s own agency, such as a rock dropping on one’s head, or be self-imposed, such as tying oneself to the mast. This improves on Ferrero’s treatment of derailment from agency in terms of a ‘dormant state.’ External limits can come from many directions, not just being asleep. However, not all failures to exercise agency are due to external limits or derailment. Some are due to causes

<sup>11</sup>The language of obstacles comes from Lewis’ (2020) analysis of abilities.



which do not make it very hard or impossible to realize certain opportunities, such as mild forgetfulness.<sup>12</sup>

(5) Finally, the modality of plight inescapability is a distinct form of practical necessity. It is not logically, metaphysically, nomologically, or psychologically necessary that I exercise agency. Nor is it practically necessary in the sense that I may take some concrete courses of actions to be (im)possible under my circumstances, as Walden construes practical necessity on his further factor account. But *because* I am, *qua* agent, continuously positioned to exercise my agency (albeit context-relatively fully and absent external limits), on pain of leaving agency, I cannot avoid being so positioned. This makes my position practically necessary in the different sense of being a precondition for any exercise of my agency. Hence, it constrains all exercises of my agency.

Summarizing, I have developed Korsgaard's brief remarks into a characterization of plight inescapability with (1)–(5). I have also solved the problems for Ferrero's interpretation (in (1) and (4) above). But can plight inescapability capture the normative force *desiderata*?

Start with *INVARIANCE*. My being positioned to exercise agency fully continuously does not shift with what I (or others) may start to desire. The positioning is independent of what is desired. Hence, *INVARIANCE* is easily captured.

What about *OIC*? Agents are positioned to exercise agency with respect to the features and commitments of their own agency, not other things. An agent is positioned so as to exercise *her* agency. Hence, we can ordinarily exercise our agency. What if things are not ordinary? Plight inescapability says that agents are positioned to exercise their agency absent external limits at the point in time of the exercise. Hence, they are not positioned to exercise agency in ways that conflict with obstacles to its exercise. If I can drive from *A* to *B*, but there is a massive rock on the only road to *B*, I am not positioned to drive to *B*. Hence, external limits to what we are positioned to do pose no problem for *OIC*.

What about *OICF*? What is plight inescapable is being positioned to exercise one's agency fully, and hence successfully relative to what the thing to do is. This does not rule out the possibility of exercising agency less than fully, or not at all. First, one may exercise agency to some extent without doing so fully, such as by drinking half a glass of gin when one intends to drink all of it. This is an unsuccessful exercise of agency, yet nevertheless an exercise. Second, one may not exercise one's agency at all. Before enacting the intention to drink the gin, the agent may fall asleep involuntarily, and hence not reach out for the glass. It is possible to fail to act by exercising one's agency insufficiently and by not doing it at all. And both types of failures could be due to some mild

<sup>12</sup>However, given my characterization of external limits, severe forgetfulness (e.g. from dementia) might be one.

hiccup (e.g. mild tiredness) or some external limit (e.g. a large rock falling on the agent's head).

Plight inescapability allows for all kinds of failure as it only takes agents to be *positioned* to exercise agency fully. It may therefore be combined with a view of agency where the agent may fail to live up to *any* of the norms constitutive of agency. Hence, the failed lie which conflicts with both the categorical and hypothetical imperatives, which I used to argue against standpoint inescapability, is still possible on plight inescapability. While the agent implausibly ends up outside agency on standpoint inescapability, plight inescapability merely takes agents to be positioned to exercise agency. This allows the agent to fail either due to some hiccup (e.g. momentarily forgetting which lie would be successful) or some external limit (e.g. a rock falling in her head, leaving her confused). Hence, plight inescapability can explain *OICF*.

To defend standpoint inescapability, one may want to add an 'absent external limits' clause to the idea that any exercise of agency comes from the standpoint of agency. With that clause, it may seem possible to avoid exercising agency at all on standpoint inescapability too. However, what standpoint inescapability takes to be inescapable is the agent's *actual* continuous exercise of her agency, whereas what is plight inescapable only is being *in a position* to exercise agency. This still leaves plight inescapability with an advantage. It should be possible to not exercise agency at all without being hindered by external limits, such as in the forgetfulness interpretation of the failed lie case. That is possible when an agent merely is positioned to exercise agency—and does not invoke the 'absent external limits' clause.

Finally, there is *CRITICIZABILITY*. Failures are criticizable on plight inescapability because it features a success condition. As I emphasized with property (3), an exercise of agency is successful or not relative to 'the thing to do', which is supported and determined by the norms that have a bearing on that exercise of agency. If agency involves the categorical and hypothetical imperatives, and no other norms are relevant at the moment, one succeeds or fails in virtue of whether one lives up to them. This allows us to criticize failures: agents may fail to do the thing to do.

The criticizability here need not be terribly strong, however. We are not necessarily talking about moral praise or blame. Not all choices are moral: ordinarily, nothing important turns on whether I choose to have a beer or a glass of wine. Rather, the criticizability is the criticizability involved in living up to, or failing to live up to, a standard of success set by the norms that support and determine the thing to do. Sometimes, that is just having the right drink.

However, in the discussion of psychological inescapability above, I charged that view with failing to explain why criticism has any bite, as it is unclear why one should have one higher order motive rather than another. That may look like a problem here too. Why be committed to what the thing to do is for *my*

agency rather than any other possible one? Fortunately, the standard of success in plight inescapability is set by the features and commitments of whatever agency one has. All agents are positioned to exercise their own agency fully (absent external limits) *regardless* of what their respective agency involves. The worry for psychological inescapability turns on the possibility of being set up in different ways. But exercising *some* kind of agency, whatever it is, is the plight of agents. For the same reason, interestingly, exercising their shmagency is the plight of shmagents.

Suitably explicated, then, plight inescapability explains the normative force *desiderata*. It captures *INVARIANCE*, *OIC*, *OICF*, and *CRITICIZABILITY*.

## VII. Whither categoricity?

But now we face a question. Normally, constitutivists do not just want to explain how norms have force for agents who already are committed to them. Rather, they aim to explain why they have force for all, or at least a suitably extensive set, of agents.<sup>13</sup> This point is often made by saying that the norms should have *categorical* force. Does plight inescapability give us that?

We may think of categoricity in several ways: as applicability, invariance, convergence, and constitution independence. Plight inescapability can at best get us the former three, but I shall argue that, together with the plausible assumption that there can be variations in the constitutions between and within kinds of agents, they generate relativism. And it cannot get us the latter.

Start with categoricity as applicability. This is the famous view of Foot (1972). She thinks that both norms of etiquette and morality are categorically binding because they apply equally to all agents, regardless of whether they endorse or have reason to follow them.

There can, however, be different kinds of agents: for example, human, animal, robot, and corporate agents. They are very plausibly constituted differently, leaving room for significant normative variation between kinds. Even worse, not all agents within these kinds seem to have identical internal commitments. Some corporate agents are democratic, others are dictatorial. There is plausibly variation even among human agents: being a member of *homo sapiens* does not guarantee that one's psychology is set up in any particular way. For example, children are often thought not to be able to act on norms in the same way as adults. Hence, there may well be significant variation within kinds of agents, too. So, while constitutivists typically argue that agency constitutively involves various normative commitments, it remains to

<sup>13</sup>This is especially so regarding morality (e.g. Korsgaard 1996, 2009; Smith 2012, 2015; Walden 2018), though see (Street 2008; Velleman 2013) for dissent. But the *desideratum* is also present elsewhere, such as regarding rationality (Brunero 2020; Leffler 2024).

be shown whether these extend across all kinds of agents or even all agents within the same kind.

Hence, different agents have different norms applying to them in virtue of their different constitutions. But plight inescapability can only explain how the norms internal to some agent's agency are forceful for that agent. Hence, agents with different constitutions have different norms applying to them as forceful.

We may, instead, try to restrict categoricity to what I have called *INVARIANCE*. Here, something is categorically binding for agents if it has force independently of what they desire to do. If categoricity is invariance, it seems to follow straightforwardly from plight inescapability, for plight inescapability captures *INVARIANCE*. However, there can be many different commitments on part of different kinds of constitutions that different agents are positioned to live up to. So, categoricity as invariance still leaves us with relativism.

One might, instead, suggest that we can capture categoricity if all agents, individually, have identical commitments (Smith 1994). For if all agents have individual normative commitments the extensions of which converge, the norm would hold for all (regardless of which norm it is). Call this categoricity as convergence.<sup>14</sup> Unfortunately, this view still leaves us with too much relativism. Different agents could have different constitutions, and different norms will therefore be binding for them. We still arrive at relativism.

Maybe all relevant agents *ought* to function identically, then? That might dissolve the worry just raised and vindicate categoricity as convergence. Indeed, I think that is what many constitutivists hope for. But I am not optimistic about the prospects of finding identical commitments for all. There are different kinds of agents, and they differ internally. What makes us think there would be a constitution they *all* ought to approximate?

There could, more plausibly, be universal commitments for theoretically *idealized* agents. Perhaps those idealized agents can somehow explain reasons for actual agents (Smith 1994, 2012). But that does not help actual, non-ideal, agents. They are positioned to exercise their own agency and commitments, and these are plight inescapable at the time of their exercise, regardless of where idealization might take them. Hence, categoricity as convergence still faces relativism due to the possibility of differing constitutions. Different agents could have different *actual* plight inescapable commitments regardless of what their idealized counterparts have.

<sup>14</sup>Lavin (2017) defends a pluralist Aristotelian constitutivism which allows for many kinds of agents with separate norms together with a curious convergence thesis, according to which some type of agency seems necessary for us *qua* humans due to a combination of our nature and upbringing (Lavin 2017: 188–93). I am sympathetic to pluralism between kinds, but why not also allow for within-kind variation among human agents? Consider someone who spends a lot of effort to change their practical dispositions, for example, in therapy. Why could this person not change between *being* different kinds of agents with different features and commitments?

I suspect that a stronger notion of categoricity is needed to capture categorical normative force without turning relativist: a norm must be binding for agents independently of how they are constituted. Call this categoricity as constitution independence. But plight inescapability cannot get us that. It can only say that whatever an agent is constituted by or committed to is forceful for that agent. Hence, if we allow for variation in constitutions between and within kinds, plight inescapability lends force to different norms. It cannot guarantee constitution independence.

### VIII. Conclusion

I have argued that plight inescapability may serve to explain the force of norms for constitutivists. In Section (I), I cleared some ground to emphasize the relevant question of explaining normative force. In Section (II), I introduced some explanatory *desiderata* about it. In Sections (III)–(V), I argued that psychological, further factor, and standpoint inescapability fail to capture some of them. But in Section (VI), I argued that plight inescapability, charitably reinterpreted, does capture them. In Section (VII), however, I then argued that plight inescapability seems to entail relativism.

This leaves us with a choice point. One may go along with the plight inescapability explanation of normative force, but then a kind of relativism follows. Or one may find that unpalatable, perhaps because of worries about relativism or of some assumption in the argument, such as plight inescapability or the types of constitutivism for which it may explain normative force. If one then wants to maintain constitutivism, one has to explain how norms are forceful in some other way. Whether another explanation could work is a question for elsewhere.

### Acknowledgements

Early versions of some ideas in this paper appeared in different forms in my dissertation, *The Constitution of Constitutivism*, at the University of Leeds (2019). I thank everyone I discussed them with at the time. More recently, the material has been redeveloped at the Universities of Vienna and Siena. At the University of Vienna, I received funding in the project ‘The Normative and Moral Foundations of Group Agency’, which was funded by the European Research Council under the European Union’s Horizon 2020 research and innovation programme (grant agreement No. 740922), PI: Prof. Pauer-Studer. As of the time of writing these acknowledgements, I am employed at the University of Siena in the project ‘A theory of retrospective responsibility as part of the system of social and moral action control’, which in turn is financed by

the European Union - Next Generation EU as the PRIN project PRIN2022 CUP B53D23033960006, PI. Prof. Lumer. I am grateful for my funding at both Vienna and Siena. Recent versions of the paper have been presented at the Third Groningen Metaethics Workshop (at Groningen), Possibilities, Impossibilities, and Conflict in Ethics (at Pardubice), and ECAP11 (at Vienna). I thank everyone in the audiences and everyone else who has given me feedback on these versions too - especially Debbie Roberts, Christoph Lumer, Pekka Väyrynen, and two anonymous referees for this journal.

## References

- Bengson, J., Cuneo, T., and Shafer-Landau, R. (2023) 'The Source of Normativity', *Mind*, 132: 706–29.
- Brunero, J. (2020) *Instrumental Rationality: The Normativity of Means-Ends Coherence*. Oxford: OUP.
- Coleman, M. (2023) 'Threshold Constitutivism and Social Kinds', *Journal of Ethics and Social Philosophy*, 25: 642–9.
- Enoch, D. (2006) 'Agency, Shmagency: Why Normativity Won't Come from What Is Constitutive of Action', *The Philosophical Review*, 115: 169–98.
- Enoch, D. (2011) 'Shmagency Revisited'. In: M. Brady (ed.) *New Waves in Metaethics*, pp. 208–33. New York, NY: Palgrave MacMillan.
- Ferrero, L. (2009) 'Constitutivism and the Schmagency Challenge'. In: R. Shafer-Landau (ed.) *Oxford Studies in Metaethics*, Vol. 4, pp. 303–32. Oxford: OUP.
- Ferrero, L. (2018) 'Inescapability Revisited', *Manuscrito*, 41: 113–58.
- Ferrero, L. (2019) 'The Simple Constitutivist Move', *Philosophical Explorations*, 22: 146–62.
- Finlay, S. (2019) 'Defining Normativity'. In: D. Plunkett, S.J. Shapiro and K. Toh (eds.) *Dimensions of Normativity: New Essays on Metaethics and Jurisprudence*, pp. 187–220. Oxford: OUP.
- Foot, P. (1972) 'Morality as a System of Hypothetical Imperatives', *The Philosophical Review*, 81: 305–16.
- Haslanger, S. (2016) 'What Is a (Social) Structural Explanation?' *Philosophical Studies*, 173: 113–30.
- Hume, D. (1739–40) 'A Treatise of Human Nature'. In: L.A. Selby-Bigge and P.H. Nidditch (eds.) (1978) *A Treatise of Human Nature*. 2nd edn. Oxford: OUP.
- Katsafanas, P. (2013) *Agency and the Foundations of Ethics: Nietzschean Constitutivism*. Oxford: OUP.
- Katsafanas, P. (2018) 'Constitutivism about Practical Reasons'. In: D. Star. (ed.) *Oxford Handbook of Reasons and Normativity*, pp. 367–91. Oxford: OUP.
- Korsgaard, C. M. (1996) *The Sources of Normativity*. Cambridge: CUP.
- Korsgaard, C. M. (1997) 'The Normativity of Instrumental Reason'. In: G. Cullity and B. Gaut (eds.) *Ethics and Practical Reason*, pp. 215–54. Oxford: OUP.
- Korsgaard, C. M. (2003) 'Realism and Constructivism in Twentieth Century Moral Philosophy', *Journal of Philosophical Research*, 28: 99–122.
- Korsgaard, C. M. (2009) *Self-Constitution: Agency, Identity, Integrity*. Oxford: OUP.
- Lavin, D. (2004) 'Practical Reason and the Possibility of Error', *Ethics*, 114: 424–57.
- Lavin, D. (2017) 'Forms of Rational Agency', *Royal Institute of Philosophy Supplement*, 80: 171–93.
- Leffler, O. (2019) 'New Shmagency Worries', *Journal of Ethics and Social Philosophy*, 15: 121–45.
- Leffler, O. (2023) 'Agent-Switching, Pliht Inescapability, and Corporate Agency', *Analytic Philosophy*, 00: 1–17. [Online First.]
- Leffler, O. (2024) 'Rationality, Shmatinality: Even Newer Shmagency Worries', *Journal of Ethics and Social Philosophy*, 27: 371–404.
- Lewis, D. K. (2020) 'Outline of "Nihil Obstat: an Analysis of Ability"', *The Monist*, 103: 241–44.
- Lindeman, K. (2017) 'Constitutivism without Normative Thresholds', *Journal of Ethics and Social Philosophy*, 12: 231–58.
- Moore, G. E. (1903) *Principia Ethica*. Amherst, MA: Prometheus Books.
- Parfit, D. (2006) 'Normativity'. In: R. Shafer-Landau (ed.) *Oxford Studies in Metaethics*, Vol. 1, pp. 325–80. Oxford: OUP.

- Ridge, M. (2018) 'Meeting Constitutivists Halfway', *Philosophical Studies*, 175: 2951–68.
- Schafer, K. (2023) *Kant's Reason: The Unity of Reason and the Limits of Comprehension in Kant*. Oxford: OUP.
- Smith, M. (1994) *The Moral Problem*. Oxford: Blackwell.
- Smith, M. (2012) 'Agents and Patients: Or, What We Learn about Reasons for Action by Reflecting on Our Choices in Process-of-Thought Cases', *Proceedings of the Aristotelian Society*, 112: 309–31.
- Smith, M. (2015) 'The Magic of Constitutivism', *American Philosophical Quarterly*, 52: 187–200.
- Street, S. (2008) 'Constructivism about Reasons'. In: R. Shafer-Landau (ed.) *Oxford Studies in Metaethics*, Vol. 3, pp. 207–45. Oxford: OUP.
- Velleman, J. D. (1989 [2007]) *Practical Reflection*. Stanford: CSLI Publications.
- Velleman, J. D. (2000) *The Possibility of Practical Reason*. Oxford: OUP.
- Velleman, J. D. (2009) *How We Get Along*. Oxford: OUP.
- Velleman, J. D. (2013) *Foundations for Moral Relativism*. Cambridge: Open Book Publishing.
- Walden, K. (2018) 'Practical Reason Not as Such', *Journal of Ethics and Social Philosophy*, 13: 125–53.
- Worsnip, A. (2021) *Fitting Things Together: Coherence and the Demands of Structural Rationality*. Oxford: OUP.