1. Introduction

The core claim of epistemic internalism is that epistemic justification is determined by what is internal to the mind, not by facts about the mind-independent world. The implications of this claim depend on additional views about what it is to be internal to the mind. Most contemporary internalists accept one of two views: either (1) to be internal to the mind is to be reflectively accessible, or (2) to be internal to the mind is to be a non-factive mental state.¹ I reject these views, but I think internalism is defensible when founded on a different, neglected conception of internality. My aim is to defend this thought.

The kind of internalism that I prefer is rooted in a rationalist tradition that has been surprisingly overlooked in recent epistemology,² despite inspiring internalist projects in cognitive science like Chomsky’s ‘Cartesian Linguistics’ and the computational theory of mind. Neglect of rationalist insights has, I think, created a distorted impression of the prospects for internalism, by saddling internalists with the burdens of empiricism. Internalists can refuse these burdens by affiliating their view with a better philosophy of mind. Rather than looking to 17th century rationalists, as Chomsky does, I look to Kant, and especially Kant’s idea that the norms of action and thought are explained by the constitution of agency and mind.

I believe that a Kantian version of internalism can avoid stock objections to internalism. It can also address new challenges from Littlejohn (2012, this volume), who prefers externalism because he thinks it fits better with the general theory of normativity. Internalism is, I will show, not excluded by this new methodology: it fits well alongside Kantian ethics and philosophy of practical reason. Internalism hence needn’t rest on conflations of justification and excuse or confusions about reasons and rationality. It follows from a more general non-consequentialist and constitutivist approach to normativity.

With these ideas in mind, here is the plan. In §2, I get back to basics by considering a template for internalism as a view about the metaphysics of epistemic justification. I consider a wider range of ways to fill in this template than has become standard. I note that contemporary internalists neglect rationalist and Kantian insights about how justification could be grounded in the constitution of the mind and hypothesize that if internalism is founded on these insights, some problems disappear. I develop this hypothesis more

¹ For this division, see Littlejohn (2012: 12-16) and Feldman and Conee (2001: 2).
² One exception is Pollock and Cruz (1999), who suggest an analogy with Chomsky’s understanding of linguistic competence. Their view is the closest to mine in recent epistemology.
concretely in §3 with a Kantian version of internalism, explaining its advantages over other internalist and externalist approaches. I further defend this view in §4 on the basis of a non-consequentialist ethics of belief from Sylvan (2020a). I conclude in §5 by explaining why Kantian internalism can answer Littlejohn’s challenges.

2. **Groundwork for the Metaphysics of Epistemic Normativity**

The basic template for internalism is the following thesis:

I: Epistemic justification is determined by what is internal to the mind.

Two claims are extractable from this template if we distinguish *ex ante* justification—there *being justification* for S to φ—and *ex post* justification—S’s φ-ing *justifiedly*:

I_{ea}: Whether there is *ex ante* epistemic justification for S to φ is determined by what is internal to S’s mind.

I_{ep}: Whether S φs with *ex post* epistemic justification is determined by what is internal to S’s mind.

Many contemporary internalists regard *ex ante* justification as prior, and hence see I_{ea} as the core thesis. But some historical forms of internalism prioritize *ex post* normative concepts. Just as Kantian ethics emphasizes the *good will*, so one old form of internalism emphasizes the *quality of thought*, understood as an internal property of one’s thinking determined by whether it is *well-structured*. I will return to this point.

Our template invites four questions:

Q1: What is it to be internal to the mind?

Q2: What is (relevantly) internal to one’s mind?

Q3: What is it for justification to be *determined* by the internal?

Q4: How could *justification* (rather than mere excuse) be determined by the internal?

I think recent internalists entertain an overly limited range of answers to Q1-Q3. I also think Littlejohn (2012) is right that internalists mostly either (a) fail to answer Q4, by resting content with arguments from intuitions about cases like the demon world, or (b) make assumptions undermined by the general theory of normativity (e.g., the assumption that justification = rationality).

---

3 See, e.g., Fumerton (1995) and Conee and Feldman (2002) for representative examples. While this direction of analysis is typical, there are exceptions, like Wedgwood (2002) and especially Pollock and Cruz (1999: 25):

‘The idea behind internalism is that the justifiedness of a belief is determined by whether it was arrived at or is currently sustained by ‘correct cognitive processes’. […] Internalist theories are committed to the principle that the correctness of an epistemic move (or cognitive process) is an inherent feature of it.’

4 Some internalists (e.g., Fumerton (1995)) do better by acknowledging the need for an objective link between justification and truth. But they also are skeptics.
Internalism’s prospects can be improved in several ways. One way is to consider a wider range of answers to Q1-Q3. Another way is to tackle Littlejohn’s challenge and develop an answer to Q4 that fits with the wider literature on normativity. Here epistemic internalists can learn from meta-ethical internalists, and especially Kantians like Smith (1995), Korsgaard (2009) and Markovits (2015). A final way internalists can do better is to learn from non-consequentialists (especially Kant (1785)) who argue that the most fundamental moral demands are demands of respect. I believe that something close to Kant’s own epistemology does it all: it offers better answers to Q1-Q3, mirrors Kantian internalism in meta-ethics, and fits well with an ethics of belief that parallels Kantian ethics.

I’ll now set the stage for this idea with some further reflections on Q1-Q4.

1.1. The Internality Relation

What is it to be internal to the mind? A naïve approach suggested by Littlejohn (this volume)’s title would treat being internal to the mind in spatial terms, so that being internal to the mind = being in the head. Epistemic internalists don’t tend to understand internality in this way, however. It is helpful to remember why.

One reason is that being an epistemic internalist shouldn’t require holding that justification is grounded in brain states. Indeed, for traditional internalists, the brain is just as external as the world. Even if minds are always physically realized, being relevantly internal does not consist in being spatially internal to a physical realizer. Spatial properties on their own needn’t have any epistemic significance for internalists.

This simple point matters, because it suggests that epistemic internalists needn’t automatically reject some ‘externalist’ claims in philosophy of mind, which rest on spatial readings of ‘external’. As I will argue below, they could even allow that justifiers include non-mental items like propositions if they hold that these count as justifiers only if suitably related to minds. Internalists about practical reason since Schroeder (2007) have stressed this point. Epistemic internalists can proceed similarly.

While the spatial answer to Q1 is wrong, it is a special case of a wider idea that has more helpful versions. Note that the spatial view is plausible if one identifies minds with material objects, since what is internal to a material object are its material parts, their non-relational properties, and their organization. There is a more general idea lurking here that doesn’t force a spatial picture: what is internal to any X are X’s constituents, their non-relational properties, and their organization. This generalization doesn’t entail a spatial picture. For there can be constituents of more abstract kinds. Consider the functional and architectural constituents of a computer, the syntactic constituents of a sentence type, or the formal constituents of a work of music. Partly because mental properties are multiply realizable, some mental constituents are plausibly constituents of a more abstract sort. Of course, mental constituents are physically realized. But realizers are realizers, not constituents.

With this idea in mind, we might consider:

Constitutive Internality: To be internal to S’s mind is to be grounded in the constituents of S’s mind, their non-relational properties, and their organization.
Given a notion of constitution that is contingent and temporary, as in Baker (2000), Constitutive Internality doesn’t commit internalists to thinking that justification flows only from essential features of minds. Hence Constitutive Internality only entails that what is internal to minds is what is intrinsic to them on the ‘interiority’ conception, which treats intrinsic facts about X as facts about ‘X itself’.  

Note that internalists who accept Constitutive Internality can allow that there is wide content even if a state’s possession of such content is an extrinsic property. Constitutive Internality only suggests that justification for beliefs with wide content derivates from what is internal to minds. Internalists might hold that justification for wide contents always derives from justification for narrow contents, for example. The derivativeness claim doesn’t require this claim, however. It is compatible with denying that there is a deep narrow/wide distinction: it just says that justification for any content ultimately derives from mental constitution. Either way, internalism as I understand it is compatible with content externalism.

Granting Constitutive Internality, standard non-factive mentalism and accessibilism don’t look like they should be the default views. Standard non-factive mentalists make justification determined by non-factive mental states (e.g., experiences) and their non-relational properties (e.g., phenomenal character). Such states and properties might be temporary mental constituents. But it is unobvious why they are the only epistemically relevant constituents. Similarly for accessibilism: perhaps introspectable states are mental constituents, but it is implausible that only introspectable states are mental constituents. The nature and organization of one’s mind are deeper features that are not straightforwardly accessible. Internalists should consider the epistemic role of such features.

1.2.  What is Internal?

What is internal to a mind? Most contemporary literature focuses on internal states and their properties. Hence the most prominent forms of internalism seek states with intrinsic reason-giving force—seemings or states of direct acquaintance—to serve as epistemic foundations. Externalists and coherentists reasonably argue that the quest for intrinsically reason-giving states is futile. I agree, but the fitting response isn’t to go externalist or coherentist. One can instead be more deeply internalist.

Short-stay occupants of minds like experiences aren’t all that matter. Essential and structural features of minds are also epistemically relevant. Some historical internalists—Kant and some rationalists (e.g., Leibniz)—emphasized such features. Rational capacities, like the capacity of understanding, represent one important historical account of the essence of mind. Structural features, such as Kant’s a priori forms of the Understanding (the Categories), could also play a role. To provide an up-to-date Kantian story, internalists can consider relevant cognitive science—e.g., work on innate structures like ‘core knowledge’ according to Spelke (2022). They might view these structures as constitutive of empirical understanding rather than Kant’s Categories.

5 See Marshall and Weatherson (2018), and cf. Lewis (1983: 197): ‘A thing has its intrinsic properties in virtue of the way that thing itself, and nothing else, is.’

6 This derivativeness isn’t inferential: non-skeptical internalists can allow that we have immediate justification for some beliefs with wide content. A basic/non-basic distinction needs to be drawn orthogonally to the non-inferential/inferential distinction; see Silins (2011: 355ff).
Fixating on states and their intrinsic character needlessly saddles internalists with Humean burdens. Hume was wrong: minds are not bundles of impressions and ideas. There are a priori principles of mental organization. Subjects also actively structure their mental life. Internalists should incorporate these non-Humean points. To help categorize some non-Humean internalist views, I’ll focus on cognitive constitution—in particular:

i. one’s constituents qua cognizer

ii. constituents of one’s cognitive record—i.e., manifestations of one’s cognitive constitution (rather than external factors)

iii. the organization of (i) and (ii)

iv. non-relational mental features of (i) and (ii)

Humeans have a bleak picture of (i-iv): cognitive life is organized by laws of association, one’s cognitive record is a sum of what happens in one’s mind, and one’s cognitive identity is fixed by this record. This picture is mistaken. Cognitive life is sometimes organized by principles of rationality. One’s cognitive record is not a mere sum of what happens in one’s mind. Moreover, one has a cognitive identity apart from one’s cognitive record.

Internalists can develop these ideas in different ways. They may hold that cognitive life is more or less governed by rational laws. They may do more or less to explain non-rational phenomena by external factors. They may have a more or less essentialist picture of cognitive identity. On the extreme ends of this spectrum are traditional rationalists and theorists Callahan (2021) calls epistemic existentialists. They view cognitive constitution as fixed by (i) innate substantive principles of rationality, or (ii) self-determined principles. For some rationalists, it is possible for one’s cognitive record to include voluntary misapplications of these principles. But Panglossian rationalists may see all non-conformity as heteronomy or performance error. On these views, arational and irrational cognitive events are contrary to one’s epistemic principles, which are innate or determined with quasi-libertarian epistemic freedom. On both views, epistemic rationality expresses epistemic freedom—positive freedom (for rationalists), or negative freedom (for existentialists).

These views suggest a very different picture of what is internal than classical foundationalism, phenomenal conservatism, and accessibilism. On these views, which states are properly internal to cognizers is settled by their principles and cognitive identity: these determine which experiences are taken at face value and which are dismissed as misleading. Mere consciousness guarantees states no place in one’s cognitive home.

These views contain part of the truth. To be sure, one’s cognitive identity and principles aren’t all that matter: this risks dogmatism. But cognition is not merely the taking of the given. To avoid dogmatism and skepticism, internalists should allow that cognizers are constituted both by a passive capacity to receive evidence and an active capacity to process evidence. Kantian epistemology rests on this idea.

1.3. The Nature and Possibility of Internal Determination

Besides seeking a richer conception of internality, internalists should develop a subtler account of what it is for justification to be determined by the internal. There is more room
for maneuver here than is generally appreciated. To bring this out, consider a representative approach from Conee and Feldman (2002, 2008). They understand determination in terms of *supervenience*, which they treat as underwritten by *epistemic principles* mapping internal states to justificational statuses. Such principles might include:

(Experience) If S has a perceptual experience with the content that $p$, there is *ex ante* justification for S to believe $p$.

Conee and Feldman (2008) acknowledge four views about such principles. They are either

- broadly analytic *a priori* truths underwritten by logic and *a priori* probability (Strawson (1952), Carnap (1962)),
- synthetic *a priori* truths discoverable through reflective equilibrium like ethical truths (Chisholm (1957, 1977)),
- deeply held beliefs or intuitions (Foley (1987), BonJour (1998)), or
- principles of inference to the best explanation (Conee and Feldman (2008), McCain (2014)).

A fine-grained understanding of determination suggests a better taxonomy. Recent literature influenced by the grounding revolution suggests that determination shouldn’t be understood in terms of supervenience, but rather as an explanatory relation. This idea was anticipated in Chisholm (1957) and Goldman (1979), who viewed the determination relation as like the *right-making* relation introduced by Ross (1930). Call this the *justification-making* relation.

There are two levels at which internalists could theorize about this relation. Internalists could give a first-order account, akin to a substantive ethical theory like Kantian ethics. Internalists could also give a meta-epistemological account, mirroring internalist meta-ethicists. One can simultaneously pursue projects at both levels, with one’s meta-normative and first-order views being mutually supportive. Many internalists in ethics proceed in this way.

These points suggest a better taxonomy. Note that some views on Conee and Feldman’s list aren’t of the same genre. Chisholm has a first-order theory of epistemic principles (an ‘ethics of belief’). By contrast, Carnap and Strawson view epistemic principles as conceptual truths explained by a meta-epistemological story. These are different kinds of views. Once distinguished, we need two (longer) lists of views.

On the meta-epistemological side, we should consider other views that appeal to *a priori* principles, like Wright (2005)’s appeal to *a priori* entitlements. One can imagine other ways to invoke the *a priori*. Suppose we take sufficient reasons to be potential inputs to good reasoning, following Way (2017). We might then understand *good reasoning* as reasoning that accords with the constitutive principles of the faculty of reason. This view provides a more fundamental constitutivist explanation of epistemic principles.

---

7 See Beddor (2015).
8 This is how Fumerton (1995), who introduced the term ‘meta-epistemology’, understands internalism.
The first-order possibilities are even less explored. Chisholm’s particularism is not the only option. Internalists could develop a more systematic first-order view like Kantian ethics. One can, I suggest in Sylvan (2020a), hold that the basic normative explanation of why justified beliefs are justified is that they manifest respect for truth, where this is determined by the structure of one’s thinking and its relation to input states and constitutive principles.

Although first-order and meta-epistemological theories are distinct, they interact. Some internalist meta-epistemologies have first-order implications, and some first-order views exclude meta-epistemological options. This point is central to internalism/externalism debates in practical philosophy. Consider how internalists about practical reason feel that externalists like Parfit (2011), who reject internalism on first-order grounds, beg the question: they think meta-normative questions are prior. Skeptical epistemic internalists like Fumerton (1995) share this methodology.

Internalists can reject these strictures without becoming particularists. Internalists can grant some first-order intuitions as a priori evident (unlike skeptics), but wonder how these intuitions could possibly be a priori truths (unlike particularists). They may then offer a meta-epistemology that explains and justifies the intuitions. Kantian internalists, for example, uphold intuitions that skeptical internalists reject. But they don’t rest content with intuition-mongering. It is, after all, puzzling how justification could be internalist without skepticism ensuing. But this is no problem: it is a puzzle, to be solved transcendentally.

Kant’s Critical philosophy solves a related puzzle, but my approach is different. Kant thought the only way commonsense epistemic principles could be objective and a priori is if transcendental idealism is true. We don’t need transcendental idealism to solve our puzzle. Less radical Kantian ideas help. One is a Kantian view about epistemic value: epistemic value is objectively truth-connected, but the connection is non-instrumental—justified beliefs manifest respect for truth, which is the fundamental epistemic norm. One shouldn’t stop here, since one must explain how epistemic principles could be internally determined, and why respect for truth supervenes on the internal. Hence I recommend a further Kantian view: epistemic principles are constitutive principles of cognition.

In short, I suggest a Kantian account depictable as follows, where the arrows represent support relations:

```
Standard internalist claims and principles are

<table>
<thead>
<tr>
<th>normative truths</th>
<th>that are</th>
<th>a priori</th>
<th>because</th>
</tr>
</thead>
<tbody>
<tr>
<td>↑</td>
<td></td>
<td>↑</td>
<td></td>
</tr>
</tbody>
</table>

Kantian ethics of belief is true ↔ constitutivism is true
```

This approach is consistent with standard internalist intuitions (e.g., demon-worlders are justified) and principles (e.g., experience justifies belief), but treats them as requiring explanation. Kantian ethics of belief gives the normative explanation. Kantian constitutivism gives the metaphysical explanation. These views are independently defensible, but also mutually supportive.
I think internalists need a comprehensive theory of this kind. Littlejohn is right that intuition-based arguments are unpersuasive. The intuitions need support from a more fundamental first-order theory. Typical internalist principles also need support from a more fundamental meta-epistemological theory. These theories work best as a package deal.

I turn to further develop and defend this Kantian approach. In §3, I sketch a Kantian account of epistemic normativity that parallels Kantian accounts in the philosophy of practical reason. I flesh it out by relying on a Kantian account of what is internal to the mind, noting how Kant (1787)’s metaphysics of mind proves illuminating. §4 then gives a normative argument for Kantian constitutivism from my non-consequentialist view in Sylvan (2020a).

3. Epistemic Self-constitution

The first plank of Kantian internalism is an epistemic analogue of Korsgaard's constitutivism:

**Kantian Epistemic Constitutivism (KEC):** All epistemic norms that have authority for S have it only because of their compatibility with the constitutive principles of S’s cognitive capacities and activities,

where these principles include

- **universal** principles constitutive of cognizers *as such,*

- **subject-specific** principles compatible with these that are constitutive of S’s mode of cognitive agency (S’s *epistemic style,* as Flores (2021) says),

and

- **activity-specific** and **concept-specific** principles compatible with these that are constitutive of particular cognitive activities (e.g., deduction, abduction) or full possession of some concepts (e.g., logical and observational concepts).

KEC is internalist, but differs from some popular versions of internalism. The internalist precept KEC honors is familiar in ethics: only rules compatible with your autonomy have normative authority over you. Call this the Autonomy Precept (AP). Goldman (1980: 32) helpfully explained how some internalists invoke AP:

> On the [internalist] perspective […] a [doxastic principle] must not be ‘heteronomous,’ or dictated ‘from without’. It must be ‘autonomous,’ a law we can give to ourselves and which we have grounds for giving to ourselves.

While internalists don’t always invoke AP, some forms of internalism do rest on it. Chisholm (1982: 63) took his accessibilism to follow from AP, citing Mercier (1918: 214):

> [The criterion] should be internal. No reason or rule of truth that is provided by an *external authority* can serve as an ultimate criterion.

Hence KEC is cut from the same cloth as Chisholm’s epistemology. Unlike Chisholm and like Kant, I don’t think AP favors accessibilism. For Kantians, epistemic autonomy is
government by constitutive principles. Such principles are hard to discover. They are not ‘egocentric’, pace Goldman.\(^9\) Autonomy requires obeying categorical principles, because the autos who sets the nomos is the rational agent as such. Like Korsgaard (2009), I think Kantians shouldn’t limit themselves to universal constitutive principles. But they cannot appeal only to subject-specific ones: these must fit with universal requirements.\(^10\)

KEC further diverges from typical internalist views because it is compatible with allowing non-mental factors to have derivative authority. Although what fundamentally matters for KEC is mental constitution, KEC only says that the authority of epistemic norms is grounded in mental constitution. This kind of point is familiar in meta-ethics. After Schroeder (2007), meta-ethical internalists agree that reasons are not mental entities, but rather propositions or states of affairs. They just say that such things have normative authority for S because of facts about S. Hence KEC can allow that reasons are not mental entities.

This point matters, since some of Littlejohn (2012)’s anti-internalist arguments invoke the fact that most meta-ethicists reject psychologism about reasons. But a superficially ‘externalist’ ontology of reasons is compatible with internalism about the reason-relation. Internalists could even agree that normative reason-ascriptions are factive but deny that this implies anything significant about the metaphysics of the reason-relation. Here we can compare internalism with response-dependent theories. Consider a silly but illustrative view: an external object is red iff I experience it to be red. This view allows that external things are red. But it is a subjectivist view, since a thing’s being red is determined by my experiences.

3.1. Applying KEC: Justification and Constitutive Normativity

KEC could be developed in different ways. I will outline a framework for applying KEC to justification and fill it in with a specific Kantian view. I don’t think internalists must accept all my Kantian baggage. But having a specific theory facilitates an existence proof that some internalist views escape Littlejohn’s challenges.

To understand the framework, let’s clarify the target. We can distinguish weaker and stronger concepts of justification. The weakest is a permissive concept: S has minimal justification to φ when S is permitted to φ. φ-ing might be permissible but only barely, by meeting minimum standards. There are more demanding concepts—φ-ing sufficiently above minimum standards (‘adequate justification’), or φ-ing optimally (‘full justification’).

To give a constitutivist account of these properties, I borrow some ideas from Kant’s ethics. Kant distinguished between perfect duties, with which one must comply to meet minimum moral standards, and imperfect duties, which set ideals for which there is some ‘playroom’ for pursuit. For Kant, perfect and imperfect duties are illuminated by two universalizability tests. One is the contradiction in conception test, on which a maxim is flawed if it cannot be universally followed. There are perfect duties against acting on such maxims. Another test is the contradiction in will test, on which a maxim is flawed if one cannot coherently will that it be universally followed. There are imperfect duties against acting on such maxims.

Related tests distinguish two kinds of constitutive norms on φ-ing in general. Say that a norm N is

---

\(^9\) Before the above quote, Goldman (1980: 32) says: ‘traditional epistemology…has been predominantly internalist, or egocentric’.

\(^10\) Hence, as Cohen (Ms) argues, an autonomy-based epistemology needn’t be an epistemic right-libertarianism.
strongly constitutive of \( \varphi \text{-ing} \) iff it is impossible to engage in \( \varphi \text{-ing} \) without being disposed to heed \( N \)

and

weakly constitutive of \( \varphi \text{-ing} \) iff it is impossible to engage in \( \varphi \text{-ing} \) without being committed to \( N \).

Chess rules provide examples: rules about legal moves are strongly constitutive, whereas the ideal of checkmating is only weakly constitutive.

We can also distinguish weakly and strongly constitutive norms of a capacity \( C \):

strongly constitutive norms are ones one must be disposed to heed to have \( C \),

and

weakly constitutive norms are ones one must be committed to heeding to have \( C \).

Linguistic rules are examples: rules about phrase and sentence construction for \( L \) are strongly constitutive of the capacity to understand \( L \), whereas norms of usage for \( L \)’s vocabulary are weakly constitutive.

These concepts facilitate generic constitutivist analyses of justification, to be completed with a metaphysics of cognition. One defensible set is:

(\textbf{Min}) S is minimally justified in \( \varphi \text{-ing} \) iff S’s \( \varphi \text{-ing} \) is permitted by the strongly constitutive norms of S’s cognitive capacities and \( \varphi \text{-ing} \).

(\textbf{Ade}) S is adequately justified in \( \varphi \text{-ing} \) iff S’s \( \varphi \text{-ing} \) is minimally justified and S’s \( \varphi \text{-ing} \) also sufficiently reflects commitment to weakly constitutive norms that any non-conformity is significantly explained by performance error, bad luck, or interference.

(\textbf{Full}) S is fully justified in \( \varphi \text{-ing} \) iff S’s \( \varphi \text{-ing} \) is minimally justified and S’s \( \varphi \text{-ing} \) also fully reflects S’s commitment to weakly constitutive norms, so that non-conformity is wholly explained by performance error, bad luck, or interference.

To understand these claims, consider an example:

\textbf{EXAMPLE}: S makes a valid inference to a working hypothesis from premises that make better sense of S’s current evidence than any alternative premises S has considered. S reasoned to these premises via abduction. S’s set of hypotheses and evidence base are, however, so far slim.

\textbf{EXAMPLE} is a mixed bag. S obeys perfect duties, like \textit{Deduce validly!}, and partially obeys imperfect duties, like \textit{Expand your evidence!}, \textit{Consider alternative hypotheses!}, and \textit{Theorize using abduction!} \textbf{EXAMPLE} is a clear case of minimal justification. It is not a case of full justification. And it is a borderline case of adequate justification. Since it can be fine to
accept working hypotheses on slim grounds if one plans to inquire further, it is rash to deny adequate justification: we need to know more.

What do we need to know? Here epistemology meets philosophy of psychology. I favor the kind of a priori functionalism that invokes a principle of charity.\(^{11}\) As I understand charity, it doesn’t preclude attributing mere non-conformity to norms. It does, however, require minimizing attributions of non-conformity to people’s mental capacities, rather than physical realization, circumstances, or interference. Hence, when people’s beliefs are false, their errors are mostly not ascribable to them, understood as (Lockean) persons. In typical bad cases, error is ascribable to interference from physical realization (poor ‘shape’ in Sosa’s language), circumstances (poor ‘situation’), or rationality being in snooze mode.

These claims may sound externalist, but they fit with internalism if the principle of charity is a priori. Indeed, these claims make good internalist sense. For the same reason why matches between cognition and world are never guaranteed by one’s internal states, mismatches are rarely explained by one’s internal faculties, and owe to disabling conditions. If so, error is no part of one’s record. As Kant put it:

> All errors are, so to speak, crooked lines, which we determine while being driven on the one side by the understanding, from the other side by sensibility. (Blomberg Logic 87)

Rational capacities aren’t what lead us astray. This isn’t to say we’re angels: mismatches are common, and can reflect other aspects of human nature. Kant puts it amusingly:

> Thus it happens that one finds that a village full of farmers has fewer errors by far than an academy of sciences; thoughtlessness protects them…. (Blomberg Logic 83)

But the academicians’ rationality is not the problem: the problem is that they are also realized in crooked timber. Given these points, I suggest that if one’s φ-ing complies with strongly constitutive principles and expresses genuine commitment to weakly constitutive principles, it is at least adequately justified, and non-conformity is better explained by external factors.

3.2. The Constitution of Cognition

KEC needs an internalist-friendly metaphysics of cognitive capacities. It is far beyond the scope of this paper to defend a full picture. But I’ll sketch a theory for concreteness.

3.2.1. Internalist Capacities

Capacities are usually invoked by externalists like Schellenberg (2018), who understand them as aiming at success states. But it is not the mere appeal to capacities that makes a theory externalist, but rather the appeal to their contingent relational properties in explaining epistemic normativity. Internalists can appeal to capacities too, and Kant again helps. On one recent reading,\(^{12}\) Kant has a capacity-based epistemology that appeals to essential properties of fundamental cognitive capacities to ground epistemic normativity.

---

\(^{11}\) See Lewis (1970, 1972, 1974).

One Kantian idea is to understand cognitive capacities in terms of their ‘forms’—hence, Kant analyzes the faculty of understanding by invoking the *Categories*. Modern relatives of this idea include the notion of cognitive ‘architecture’. Another Kantian idea is to understand capacities in terms of principles: ‘Experience therefore has principles of its form which ground it a priori, namely general rules of unity in the synthesis of appearances’ (A156/B195). Cognitive science rooted in Chomsky relies on a more general idea: capacities are grounded in procedural knowledge of rules. Architecture also plays a role: basic person-level capacities are grounded in knowledge of rules written into mental architecture.

What distinguishes Kantian epistemology from other views that invoke procedural knowledge like Pollock and Cruz (1999)’s is that it treats foundational procedural knowledge as *a priori*. It belongs to mind’s ‘architecture’ in Kant’s sense—*its a priori form*. This knowledge is discoverable by philosophical reflection on the possibility of experience, which Kitcher (1994) calls *transcendental psychology*. Cognitive science can help us better understand how the transcendental story is realized. But for Kantians, the most general principles belong to transcendental psychology. This picture is, I think, surprisingly defensible: projects like Spelke’s account of core knowledge of physical objects, for example, are continuous with Kant’s project. I will return to this point.

### 3.2.2. Constitutive Principles of Cognition

The first Critique focuses on three capacities: Sensibility, Understanding, and Reason. I assume a related distinction between a cognizer’s

- *presentational capacity*, which presents her with objects and features on the basis of sensory discriminations (e.g., visual discriminations of visible parts and qualities),

- *recognitional capacities*, which enable her to recognize, of some presented object or feature, that it is such-and-such (kind of) object or (kind of) feature,

- *capacity of theoretical reason*, which enables her to reason from the observed to the unobserved

These capacities constitute a capacity to know a world. But it is only contingently a capacity to know the actual world. For KEC, what matters are the constitutive principles of these capacities and their more specific manifestations.

To identify *a priori* constitutive principles of a capacity C of S, we can ask:

**The Transcendental Question:** What principles of S’s constitution make C possible?

To answer this question, we can apply two tests:

---

13 See Pylyshyn (1986: xvi-xvii) and Newell (1990: 80-88) for helpful general discussion.

14 See Fodor (1968), Cummins (1975), and Chomsky (1986). Pollock and Cruz (1999) get halfway toward an internalist epistemology based on these ideas, but their understanding of procedural knowledge seems Rylean.

15 See Cummins (1975).

16 Spelke et al. (1992) note the parallel with Kant.

17 I allow, like Johnston (2011), that these capacities can be fully manifest in bad cases.
**Strong Constitutiveness:** If we cannot coherently conceive of S as having this capacity unless we represent S as being disposed to conform to a principle R, R is strongly constitutive.

**Weak Constitutiveness:** If we cannot coherently conceive of S as having this capacity unless S is represented as committed to R, R is weakly constitutive.

To have a complete constitutivist view, we must consider the structure of cognition. I will assume that cognition properly so-called begins with the outputs of sensory systems and proceeds from there according to the following flowchart:

```
Outputs of Sensory Systems
↓
Presentations of Ordinary Objects and Features ('Simple Seeing')
↓
Basic Recognitions of Objects and Features (Low-Level Seemings)
↓
Non-Basic Recognitions of Objects and Features (High-Level Seemings)
↓
Theoretical Reasoning
```

We want to understand the constitutive principles of the capacities that underlie the arrows in this flowchart. Kant and 20th century followers like Strawson (1952, 1959) have helpful views. There are analogues in cognitive science developed by Spelke (1988, 1992) for presentational capacities, Carey (1985, 2009) for recognitional capacities, and 'mental logicians' (Rips (1994)) and theory theorists (Gopnik and Meltzoff (1997)) for rationality. For illustration I assume these theorists get different aspects of cognition right.

**Presentation.** How is it possible to be presented with whole, persisting objects like apples? For Kant, I am presented with them by applying a priori principles of understanding and imagination to sensations. In particular, I synthesize the outputs of sensory systems, which only represent parts of apples, with further imagined content, where this imaginative synthesis is guided by rules of the understanding. Only syntheses that obey these rules yield presentations of objective things.

An up-to-date Kantian story is possible. Consider an example from Spelke (1988): an infant’s visual apprehension of the continuous motion of an object passing behind an occluder. How is this possible? Spelke appeals to ‘core knowledge’ of objects—innate principles that resemble Kant’s First and Second Analogies. These allow us to complete what is only partially presented by the visual system, and thereby apprehend ordinary objects.

These principles are strongly constitutive of the capacity to be presented with ordinary objects. Without being disposed to construct presentations of substantial, bounded, and

---

18 See Sellars (1978) for a clear presentation of Kant’s view. See Nanay (2010) and Briscoe (2011) for recent imagination-based accounts of amodal perception.
spatiotemporally continuous objects from sensory input, our capacity to be presented with ordinary things wouldn’t be possible. There are, however, plausibly different specific ways to satisfy these principles. Different styles of apprehension may be constitutive of particular cognizers’ specific capacities, and a cognizer’s style may also change with development.

Some imperfect duties may be weakly constitutive of presentational capacities. Some norms of attention are interesting candidates. Hyposensitivity to input may erode one’s capacity to be presented with ordinary objects. To maintain presentational capacities, one must be disposed to use input systems. Of course, one can be highly inattentive without losing touch with reality. So this is only a weakly constitutive norm.

What can these norms explain? On the positive side, they ground a priori entitlement to evidence about ordinary objects. This entitlement makes perception possible. For Kant, they also explain the possibility of the unity of consciousness (Transcendental Deduction), self-knowledge (Refutation of Idealism), and event-perception (First Analogy). On the negative side, they explain the badness of neglecting sensory input.

Basic Recognition. How is it possible to recognize presented objects and features as objects and features? For Kant, the Categories do the work. There are modern versions of this idea. Spelke’s project explained how we segment the perceived world into units fit for subsumption under categorial concepts like object. A further project is needed to explain how we subsume these units under categorial concepts. Here Carey (2009) is complementary: the concept of an object in Spelke’s sense makes recognition possible. Hence strongly constitutive norms of basic recognitional capacities might include being disposed to conceive of perceptual objects as substantial, bounded, and spatiotemporally continuous. More determinate ways of conceptualizing physical objects may be strongly constitutive of specific recognitional capacities. These may change with development.¹⁹

Some imperfect duties may be weakly constitutive of recognitional capacities. For reidentification to be possible, cognizers must be disposed to remember the information needed for recognition. Kant calls this the capacity for the ‘synthesis of reproduction’ (A102): ‘[I]f I were always to lose the preceding representations…from my thoughts and not reproduce them when I proceed to the following ones, then no whole representation…could ever arise.’ Kant’s formulation allows for lapses due to limited working memory and attention, or other hardware limitations. So the duty is only weakly constitutive.

What can these norms explain? On the positive side, they ground a priori entitlement to belief in the ordinary world, by making it possible to understand perceptual objects as ordinary objects. This entitlement solves the problem of the external world. On the negative side, they explain the badness of wilful inattentiveness.

Non-Basic Recognition. How is it possible to recognize secondary qualities across changes in sensory input (e.g., color), or to recognize garden-variety natural kinds? Strongly constitutive norms of non-basic recognition might include the dispositions associated with psychological essentialism (Gelman (2003)): the disposition to treat observable features and patterns as explained by hidden causes. Kant appeals to reason’s drive toward systematic unity to explain acquisition of special science concepts. This drive, Kant suggests (A652/B680), makes possible a taxonomic understanding of nature.

Such principles are applications of a broader principle of inference to the best explanation. It is hard to see how one could glom onto patterns in the world, and hence acquire high-level concepts, without commitment to this principle. Still, there is room for variation. The principles are only weakly constitutive. The only strongly constitutive norms seem to be requirements on consistent taxonomizing and concept application, in line with whatever introduction and exit rules are strongly constitutive of possessing particular concepts.

**Theoretical Reason.** The principles of high-level recognition are implicit principles of theoretical reason. These principles can be made explicit and used in theoretical reasoning.

Such reasoning can, I think, be helpfully divided into four stages:

- **Wondering:** Wondering what explains some presented state of affairs $E$.  
- **Pondering:** Surveying possible explanations $H_i$ and getting input systems to deliver more presentations $E_i$.  
- **Evaluating:** Evaluating $H_i$ and forming *transitional attitudes* (see Staffel (2019, 2021)).  
- **Concluding:** Drawing conclusions, which can be close-minded (e.g., hard agnosticism) or open-minded (e.g., suspension that disposes one to revisit the question).

Each step introduces perfect and imperfect duties. Pondering should obey the imperfect duty to gather explanations and evidence, evaluating should obey the imperfect duty to seek best explanations and perfect duties to evaluate consistently, and concluding should obey perfect duties constitutive of the relevant type of act (e.g., deducing requires using valid rules).

Explicit thinking should obey these principles. Implicit thinking can be evaluated by them too, but since implicit ponderings and evaluations are less open to rational control, their non-conformity should be partly explained by imperfect realization.

3.3. **Summarizing and Comparing**

The supreme principle of KEC is this:

*Constitutivist Meta-Norm:* For any cognitive activity $\varphi$, $S$ ought to $\varphi$ in such a way that the maxim of $S$’s $\varphi$-ing aligns with strongly and weakly constitutive principles of cognition.

Given our more specific picture, we can see how KEC improves upon standard internalist views and can capture benefits of externalism without making justification determined by contingent relational properties of cognizers.

---

20 This is what is right about viewing perception as inference—though only high-level perception, per Pylyshyn (2003) and Sylvan (2020b). For inferential approaches treat perceptual seemings as outputs of implicit processing of the kind I discuss below, see especially Gregory (1980)’s and Rock (1983).

21 See Lord and Sylvan (forthcoming) for discussion of close-minded and open-minded attitudes.
On KEC, whether a cognitive activity is justified is determined by how well its structure aligns with constitutive principles. All cognitive episodes except direct outputs of sensory systems should be well-structured. Since seemings reflect intermediate cognitive processing, they don’t automatically justify. Hence KEC diverges from phenomenal conservatism.

Since presentations also reflect cognitive processing, there is a deeper epistemic story about justification for basic empirical beliefs than liberal acquaintance foundationalism suggests. But because the relevant principles are strongly constitutive, KEC doesn’t demand more work. Hence empirical justification isn’t as demanding as classical foundationalists suggest.

KEC diverges from stereotypical internalist views in further ways. Phenomenology and accessibility have no fundamental significance on KEC. What matters is the structure of your cognitive activity. Perhaps you necessarily have practical self-knowledge of it, of the implicit sort accompanying all acts that manifest know-how. But this is not knowledge via inner observation, and needn’t be articulable.

KEC also shares some virtues of reliabilism and virtue epistemology while avoiding their vices. Like process reliabilism, KEC treats justification as determined by the quality of one’s cognitive processing. But unlike process reliabilism, KEC denies that this quality is determined by contingent, extrinsic features of this processing: its quality is scrutable a priori from its intrinsic nature. Similarly, like virtue epistemology, KEC sees justification as determined by the quality of one’s cognitive efforts, but denies that this quality is fixed by how well those efforts promote true belief in the actual world. KEC is, however, compatible with a constitutive truth connection: it is impossible for a cognitive activity that follows weakly and strongly constitutive principles to be attributably inaccurate. This is, I think, the strongest truth connection one could reasonably want.

So, KEC promises an extensionally plausible account of justification. While a complete evaluation requires a complete transcendental psychology, KEC clearly avoids stock concerns about internalism. Given plausible views about perception and cognition, KEC avoids skepticism about the external world and ampliative inference, unlike classical foundationalism. It doesn’t overextend justification for empirical beliefs, unlike phenomenal conservatism. It also doesn’t overintellectualize justification: it has more epistemic structure than dogmatist and naïve realist views, but doesn’t require more work. Furthermore, KEC is not a ‘time-slice’ view. Although only the internal structure of a cognitive activity matters on KEC, cognitive activities are not instantaneous. There is no clear reason to place any constraints on their duration in physical time. Facts about what cognitive activity an agent is performing may even be indeterminate at a time, because the activity is not complete. So the relevant window is not necessarily accessible to introspection. Hence KEC doesn’t automatically face objections to time-slice views.

4. Ratifying the Constitution with Epistemic Non-Consequentialism

I’ve argued that KEC is a plausible internalist view. But to show that internalist justification is possible, more is needed. One must answer Littlejohn (2012)’s challenge: one must

22 Wedgwood (2002, 2017), Pollock and Cruz (1999), and Wright (2004)’s views diverge in some similar ways.
23 See Beddor and Pavese (forthcoming) for supportive ideas.
24 Wedgwood (2017)’s internalism shares this feature.
25 For arguments that Kantian views don’t make bad predictions about forgotten evidence cases, see Sylvan (2017: 149-150).
explain why internalist intuitions are tracking justification rather than weaker statuses, like mere excusability. This challenge is pressing given that KEC is a constitutivist theory and invites Enoch (2006)’s ‘shmagency’ challenge: even if certain norms are constitutive of an activity that is inescapable for us, mightn’t this activity still be unjustified?

My response is to defer to an independently defensible first-order account of robust epistemic normativity and show that KEC aligns with it. Elsewhere I defend a non-consequentialist ethics of belief according to which justified beliefs are beliefs that manifest respect for truth.26 This view is compatible with many different metaphysical analyses of justification, but its implications vary given different underlying analyses. It would be good if there were an analysis that made clear why the beliefs we take to be epistemically justified manifest respect for truth. KEC is a good fit.

If we combine KEC with epistemic non-consequentialism, we get the prediction that whether a cognitive activity manifests respect for truth is determined by its constitution. There are good reasons to believe this prediction. Respect plausibly supervenes on intrinsic features of one’s mental activity, at least when respect is understood in a genuinely non-consequentialist way. Respecting values doesn’t require promoting them. If one thinker hits upon truth and another doesn’t, both may manifest respect for truth if the latter was merely unlucky.

There are further reasons to accept this prediction. Manifesting respect consists in manifesting the right kind of motive. Activities can be individuated finely to include motives as constituents (consider murder). Hence one might expect the respectfulness of cognitive activities to depend on their maxims. This seems right: whether an inference was fallacious depends on what kind of inference it was, and what rule it embodied. Purported deductions from non-truth-preserving rules are fallacious, but purported inductions from non-truth-preserving rules are not. So, the link between respect and motivation recommends KEC.27

A further argument rests on the observation that whether you respect a value V depends on how you respond to relevant appearances, including misleading ones.28 If all available information misleadingly suggests that X-ing is harmless to V and you take V into account in your deliberation about whether to X, how could you manifest disrespect for V?

One might balk because of cases of negligence, normative ignorance, and cognitive penetration. What if things appear to you some way for the wrong reasons, or only because you have objectionably limited evidence? Can’t you appear to respect truth but fail to do so? If so, mustn’t the reason you fail to respect truth be external?

But these concerns are too hasty, I think. Consider the case of objectionably limited evidence. Here the problem plausibly depends on the maxim of your overall information-processing, which can taint your thinking even if you don’t realize it. If you have objectionably limited evidence, then plausibly you must have earlier manifested disrespect, by failing to expand your evidence. Persisting with maxims based on earlier disrespect

27 A more direct argument: whether some thinking manifests respect for truth depends on its maxim, but plausibly intrinsic mental duplicates share maxims.

Admittedly, precepts stable in a natural language can have wide content. But your maxims are better understood narrowly. Consider: in a slow-switching case where you begin by acting on the maxim to drink the liquid if it is water, is it plausible that your maxim changes?
28 See Sylvan (2021) for full development of this argument.
transmits disrespect to the current time-slice, since the whole thought process manifests the implicit maxim: *I will proceed to draw conclusions after this [...] too quick, too partial [...]!* survey of evidence. This story does require looking beyond the current time-slice. But this is compatible with internalism, we’ve seen.

The cognitive penetration objection may seem more concerning, since one might doubt that we can tell the same ‘tracing’ story. But Siegel’s (2017) ideas answer her (2013) challenge. Appearances derive from processing. The maxim of some processing may conflict with the processor’s commitment to veridicality. If we can fault your attitudes for being based on appearances generated by bad maxims, we can derivatively attribute those maxims to your cognitive processing. Your perceptual system is integrated with your cognition, after all. So, you go against yourself when you believe, if your perceptual processing was unconstitutional, given the processor’s commitment to veridicality.

These issues illustrate that the full picture must be richer than many internalists admit. But the Kantian story is appropriately richer. What’s necessary is the right relationship between the maxim of your cognitive processing and your cognitive constitution. Kant’s stress-tests for maxims look at more than just the content of your maxim. Moreover, your maxim may be different from your alleged, explicit maxim: it depends on what you are doing, which you may not fully appreciate.

I suggest, then, that respect for truth supervenes on the internal, properly understood. If so, and there are independent arguments for epistemic non-consequentialism, it is possible for *justification* and not mere excuse to be internally determined. This point also answers the Enochian worry. The constitutive norms of cognition matter because one must pay them heed if one is to heed the fundamental robust epistemic norm—a norm of respect for truth.

Hence epistemic non-consequentialism supports KEC. The support is mutual, since KEC also supports epistemic non-consequentialism’s predictions about justification. The two cohere in the way Kantian ethics and Kantian constitutivism cohere.

5. **Epistemic Kantianism and the Epistemology/Ethics Interface**

The coherence of Kantian epistemology and Kantian practical philosophy brings me to a final response to Littlejohn. Littlejohn (this volume) argues that the ‘naive view’ that rights and duties are external conflicts with global internalism, and suggests that since global internalism is false and epistemic and practical normativity are unified, epistemic internalism is false. But this argument and other arguments in Littlejohn (2012) fail given the availability of a global Kantian approach.

To appreciate this point, consider a capsule statement of global Kantianism. On this view, all robust normativity derives from fundamental values that are to be *respected*, not promoted: *accuracy* (in epistemology) and *personhood* (in ethics). *Per* Anderson (1993) and Wood (1999), these values are not *states of affairs*, but rather properties that figure in fundamental constitutive norms. Although these values are not internal to any individual mind, they have *authority* because they are objects of constitutive commitments of agency and cognition. This authority is hence internalist. Moreover, because a person’s manifesting respect for a value supervenes on the quality of her theoretical and practical deliberation, the justification generated by duties of respect is internalist, though not merely subjective.
The existence of global Kantianism undermines Littlejohn (2012)’s central argument for externalism. Littlejohn argued that justification must be objectively linked to truth, and that only externalism permits this. He also argued that because justification requires good reasons and these are (typically non-mental) facts, justification cannot be internal. But the Kantian approach answers these arguments. Epistemic Kantianism forges an objective connection between epistemic justification and the fundamental epistemic value (accuracy). It just understands this connection non-instrumentally, and hence doesn’t require that belief-forming processes that provide justification are instrumental to truth. What matters is being truth-oriented. That is what we should have expected by analogy with ethics.

While cognizers are constitutively committed to a truth norm (believe p only if p!), Kantian internalism implies that one doesn’t flout this norm merely by failing to conform to it. One flouts it only if one’s non-conformity is attributable to one’s cognitive agency. Non-conformity cannot be so attributable if it is impossible for one to comply with this norm given one’s epistemic position. Duties of respect can even require (unwitting) non-conformity to the constitutive standards at which they are directed if the only ways to conform are disrespectful. Hence one may have an objective, respect-based obligation to act on misleading appearances, if failing to respond to them would constitute disrespect. Hence acts and beliefs can be proximately motivated by misleading appearances but remain justified and not merely excused because they obey an objective duty of respect. This duty is the sufficient objective reason in the background.

One might still wonder whether the Kantian approach respects the idea that your duties and others’ rights are not in your head. But it is compatible with the spirit of that view. We must distinguish fundamental and derived rights, and consider how such rights and their correlated duties derive from fundamental values. A person’s fundamental rights on the Kantian view are rights against disrespectful intentional agency. These rights are correlates of a fundamental negative duty not to treat persons as mere means. Since this problematic instrumental orientation can be manifested even in skeptical scenarios in which the relevant ‘persons’ are mere holograms, the fundamental duty can be internalist.29

Fundamental negative rights are rights against intentional agency, not mere causation. A tree that falls on you causes a bad state of affairs, but doesn’t violate your rights. A sleepwalker who steps on your toe causes a bad state of affairs, but doesn’t violate your rights. As the Bystander at the Switch case suggests, it is even possible to cause death without violating negative rights. To be sure, the dignity of personhood is not exhausted by being the basis of negative duties. Persons are ends-in-themselves, and this status confers generic positive rights—e.g., the right to a life of dignity. But what these rights imply for particular agents is complicated. A patient’s derived rights and an agent’s derived duties depend on the agent’s abilities and relationship to the patient. Not all agents have the same derived duties, since limitations on an agent’s abilities constrain what is required of her. This is a good feature of the Kantian view: a diverse kingdom of ends needs agent-relativity.

Now, I may sometimes have duties to pay reparations when I merely cause a bad state of affairs. If I sleepwalk in your house and break something, it may make sense to ask me to help replace it. But this case doesn’t involve rights-violations. Perhaps you have a derived positive right to aid, but it may or may not be a sound convention to expect me to provide this aid. More importantly, there are no clear analogues of these rights-based phenomena in

29 For a full defense, see Sylvan (2021).
epistemology. Facts don’t have rights to be believed. It is incorrect to believe something false. But Kantians can explain this fact: it is incorrect because the truth norm is a constitutive commitment of being a believer.

There are still more analogies than disanalogies between ethics and epistemology, since their fundamental norms are norms of respect. There are pure duties of respect in ethics that behave like the duty of respect that I think is fundamental in epistemology.30 But we should expect some differences given that the objects of respect are different.

I conclude that global Kantianism explains both the fundamental analogies and surface disanalogies between ethics and epistemology. Hence global internalism is a live option.

References

Cohen, A. Ms. ‘A Kantian Account of Epistemic Autonomy.’


Littlejohn, C. 2012. *Justification and the Truth Connection*. CUP.


Spelke, E. 2022. *What Babies Know*. OUP.


Wood, A. 1999. *Kant’s Ethical Thought*. CUP.