Asian Journal of Philosophy (2026) 5:44
https://doi.org/10.1007/s44204-026-00405-z

BOOK SYMPOSIUM

Replies to Hazlett and Volpe

Jakob Ohlhorst[1]

Received: 5 March 2026 / Accepted: 30 March 2026
© The Author(s) 2026

Abstract
This paper responds to Allan Hazlett's critique that hinge epistemology falls victim to the wrong kind of reason challenge. Further, it responds to Giorgio Volpe's concerns regarding hinge epistemology as it is presented in *Trust Responsibly*.

Keywords Hinge epistemology · Virtue epistemology · Transcendental arguments · Wrong kinds of reasons challenge · Trust · Entitlement

I am grateful to both Allan Hazlett and Giorgio Volpe for their careful and critical comments. I will respond to both in turn, beginning with the broad concern that Allan Hazlett (2025) raises for an epistemicist hinge epistemology in general and concluding with the targeted questions for *Trust Responsibly* (Ohlhorst, 2023) by Giorgio Volpe (2024).

Hazlett presents a compelling argument against Wittgenstein-style transcendental arguments that defend the entitlement to believe some propositions to be true absent any evidential reasons in their favour (e.g. Pedersen, 2009; Wright, 2004). He proposes the following reconstruction of these transcendental arguments, going from the premises that we have some legitimate inquiry and that such an inquiry requires accepting certain propositions as true:

1. [Inquiry] E is legitimate.
2. E is legitimate only if S believes that p.
3. Therefore, S is entitled to believe that p. (Hazlett, 2025, p. 3)

To get to the conclusion, 3., this argument requires a further entitlement principle, granting us entitlement to accept necessary conditions of legitimate enquiries.

Hazlett's critique goes as follows: Analogising the situation to Pascal's Wager, he notes that the kind of transcendental argument adduced in *Trust Responsibly* and elsewhere falls prey to a wrong-kinds-of-reasons challenge. Pascal's Wager adduces

✉ Jakob Ohlhorst
jakob.ohlhorst@rwth-aachen.de

[1] Applied Ethics, RWTH Aachen, Theaterplatz 14, 52062 Aachen, Germany

Published online: 13 April 2026

game theoretic considerations in favour of belief in God. But as he notes, we cannot believe in God because that is game-theoretically favoured; that is the wrong kind of reason for belief. Pascal (2005) himself points to the impossibility of doxastic voluntarism, something I will get back to. Hazlett does not go in that direction, instead pointing out that transcendental considerations, just like game-theoretical considerations, are not the kind of reasons for which we can believe something to be true.

The challenge notably is not that this is the wrong kind of reason because it is not evidence, as Hazlett takes care to show. Instead, he points out that 2. cannot possibly be a reason for believing p; it would be a mistake *endorsing p as true* on the basis of 2. as a reason. Hazlett bolsters this by defending a variation on the Ought-Implies-Can Principle: r can only be a reason to φ if it is possible to φ for the reason that r. Otherwise you could have reasons to do certain things while at the same time you could not do these things for those reasons.

As a result, as Hazlett argues, we may end up entitled to believe certain things *without having any reason to believe* them to be true. He analogises this to having a legal right or entitlement to do something without having a reason to do it with the example of being entitled to stand on one leg for five minutes without having any reason to do so.

Many hinge epistemologists sidestep these issues by either denying that hinges are true at all, that they are taken to be true, or that they are subject to any normative considerations (e.g. Coliva, 2015; Moyal-Sharrock, 2004). I do not think that we can avoid these challenges so easily. Even if we are not aware of it, hinges do correspond to propositions that can be true, and in virtue of their possible success or failure, i.e. their truth or falsity, we are subject to normative considerations. A related move is to replace "belief in hinges" with some other doxastic attitude like "acceptance", "trust", or "commitment" (e.g. Pritchard, 2016; Wright, 2004) because of the assumed essential evidence-dependence of "belief". Again, I do not think that this can be made successfully because we need to endorse the truth of hinges wholeheartedly, be certain of it, if they are to play their fundamental role of enabling inquiry and ordinary beliefs. Notably, I take "belief" to simply be the most general doxastic state, i.e. endorsing a proposition to be true, while there are more specific ways of believing something like accepting it or trusting it to be true.[1]

The proposed reconstruction of the Wittgensteinian transcendental argument, I believe, misses a key aspect. Namely, it narrows it all to a single inquiry that has then to bootstrap its own entitlement and reasons. This may be fitting as a first pass at Crispin Wright's (2004, pp. 191–192) "entitlement of cognitive project", but it overlooks a key step for the broader transcendental project. Wittgensteinians point out that without any inquiry, we would remain cognitively paralysed and fail to

[1] Hazlett (2025, p. 7) raises the question which propositional attitudes are fitting here, and how to differentiate them. I would incline towards a multidimensional account to classify different states: one dimension would be direction of fit between attitude and world, another conditions of success, yet another evidential sensitivity. "Belief" is a technical term of epistemology that I would defend to have mind-to-world direction of fit, and whose condition of success is accurate representation—which excludes for example hypotheticals or hopes. Developing such a multidimensional account of attitudes would require its own whole research project.

acquire any beliefs or knowledge at all. We must be cognitively active[2]; thus, we need to undertake at least some inquiry. The argument would then look something like this:

1. Without any inquiry S would remain cognitively inert.
2. S must be cognitively active.
3. E is a legitimate inquiry enabling S to be cognitively active.[3]
4. E is legitimate only if S believes that p.
5. Therefore, S is entitled to believe that p.

Also Wright (2004, p. 191) points to cognitive paralysis as motivating his account of entitlement of cognitive project. The reason to believe that p is then not just 4. that E's legitimacy depends on accepting its preconditions as true, but also on the fact that without doing so we would be unable to believe anything.

Is this the right kind of reason to believe that p? Let us spell it out: Can I believe that p for the reason that if I did not believe p, I would remain cognitively paralysed? We can sharpen it even more: Can I believe that p *is true* for the reason that not believing it would make inquiry and knowledge impossible? This is the crux of Hazlett's argument: we are entitled to believe or trust these things to be true, but that offers us no reason to do so.

As James notes, these considerations are a reason not to become a sceptic who rejects p.

> I cannot do so for this plain reason, that *a rule of thinking which would absolutely prevent me from acknowledging certain kinds of truth if those kinds of truth were really there, would be an irrational rule*. (James, 2014, p. 28)

That is, they are reasons *not* to believe that p where believing p is the legitimate way to avoid scepticism. Double negations can be tricky, notably in intensional contexts, but here it seems fine: I have a reason not to omit something, where the omission is the formation of the belief that p is true. So I would end up having a reason to believe by double negation. In other words, the fact that not believing some p to be true would keep me from knowing *anything at all* in the case where there is something to know is a reason for me to believe that p is true—though it is no evidence.

Note, I do not think that this consideration gives us a reason to believe proposition p to be true in particular. Rather, this transcendental argument gives us a reason to believe some hinge propositions to be true, but which is unclear. I would also argue that we can find reasons to accept some particular p, by finding reasons to undertake some particular legitimate inquiry. If r is a reason to do inquiry E, and believing p to be true is the only way to do E, then r is a reason to believe p to be true. This obviously is the wrong kind of reason for an evidentialist. But is it in general?

As mentioned above, another way of spelling out Pascal-inspired wrong kinds of reasons challenge would be to deny doxastic voluntarism: It is not possible to

[2] This begs the question against a certain type of Pyrrhonian sceptic.

[3] *Trust Responsibly* further complicates this by questioning how we recognise legitimate inquiries.

believe for such transcendental reasons because we cannot decide or choose to believe something in the same way that we can decide or choose to do something like inquiring or cooking an egg. If we can decide or choose to do something, then we can do so for a practical or transcendental reason. This is not the challenge that Hazlett has raised explicitly, but it would block the Jamesian response because we could not believe p voluntarily.

This is where trust comes in. The kind of belief that transcendental arguments support is not just ordinary belief; rather it is the eponymous trust that p. We can distinguish the broad notion of belief as endorsing as true, basically the most general doxastic state, from ordinary belief as a special type of doxastic state. The latter are ubiquitous; they are what we form on the basis of evidence and for which doxastic voluntarism fails. Their frequency does not make them the paradigmatic notion of belief however. The former is more generic, it subsumes many kinds of doxastic states beyond ordinary beliefs. Some of these are insensitive to evidence—certainties. Some are subject to doxastic voluntarism—we can decide to trust or accept certain states of affairs to be true.

Such states can count as beliefs notwithstanding their insensitivity to evidence because they are only rational or permissible in contexts where evidence is not possible. There are contexts where there can be no evidence either way. Still, there is truth to be believed. In such cases, we can trust things to be true, where this counts as an endorsement of the truth. We can then trust p to be true for the reason that without trust in p, we would be unable to inquire or know anything.

Let us now turn to Volpe's critique, which offers a close reading of the book's argument, particularly in light of the debate around Wrightean (2004, 2014) entitlement. The first question he raises is: Why does *Trust Responsibly* commit itself to a controversial semantics—conceptual role inferentialism—if the core of the book's argument is about epistemological issues, namely the problem of demarcation and scepticism?

I agree that the argument about entitlement of cognitive activity could have gone through without any appeal to conceptual roles. However, part of the book's motivation and goal was to integrate the entitlement debate with the Wittgensteinian hinge debate. While Wright's work undoubtedly inspired hinge epistemologists like Coliva (2015) and Pritchard (2016), they have distanced their frameworks from the robustly epistemological concerns of entitlement, in line with the Wittgensteinian *modus operandi*. Part of the book's goal—following arguments I made elsewhere (Ohlhorst, 2022)—is to show that hinge epistemologists cannot ignore epistemic normativity. Following Coliva (2015), hinge epistemology is best interpreted as connecting conceptual roles with epistemological concerns. By adopting this framework, we can see how entitlement and the problem of demarcation are indeed relevant for hinge epistemology, notwithstanding hinge epistemologists' protestations.

Indeed, Volpe considers that my proposal depends on cornerstones or hinges being implicitly believed. In other words that "a proposition plays the cornerstone role within a belief set if and only if [implicitly] *believing it is a necessary condition for believing a large set of other believed propositions*" (Volpe, 2024, fn1). The key idea is that, if you accept a bunch of peripheral propositions, then you would be irrational to reject the entailed cornerstones—you are rationally committed to implicitly

taking it to be at least as certain as the most certain peripheral proposition. Volpe critically notes that this is simply not how *Trust Responsibly* introduces cornerstones, and I could have been clearer in these passages: the entailment relations for cornerstones mean that you implicitly have to accept them lest you risk irrationality. Note that Hazlett's (2025) criticism above nevertheless targets exactly the connection of believing a cornerstone being a necessary condition for inquiries that I was aiming for, so the idea is clearly present.

Relatedly, Volpe raises the challenge if chains of entailment are indeed sufficiently robust to generate the set of hinges to which we are committed. For instance, I am in no way committed to the hinge that "everybody has parents" just because I believe that "Bob opened the door". Indeed, I am not certain if entailment is the only relation that determines hinges and conceptual content. I am open to the idea that other semantic relations may also structure what counts as a hinge or cornerstone; however, I aimed to make the model as simple as possible, and tree networks of entailment do that job remarkably well.

"Bob opened the door" on its own does not entail much of anything except maybe that doors can be opened, and that Bob and doors exist. However, to make a Wittgensteinian point, if "Bob opened the door" were the only proposition you believed, it would also be mostly meaningless. For you to interpret this proposition in the way that an ordinary reader of English would, you need to believe a bunch of things about people opening doors, beliefs about Bobs who are people, beliefs about people being human organisms with all the human organs, and so on. All of *those propositions* together would end up entailing the hinge that everybody has parents. To illustrate, nobody who reads the sentence is likely to interpret it as "the automatic mechanism, that everybody at the office calls Bob, opened the door"; this automatic Bob would not have parents. Thus, to understand "Bob opened the door" in the way that we usually understand such a sentence, we have to believe a large set of propositions and as a result are indeed committed to accepting "everybody has parents".

In sum, one can also postulate the following schema to recognise cornerstones: How many propositions that I believe to be true would I have to reject if I rejected this proposition given that it is a necessary condition for the former? The larger the set of propositions to be rejected is, the more central a commitment is.[4] This entails a gradual notion of the hinge status. This leads us to another concern raised by Volpe: How do Wright's "regions of thought" relate to these large sets of propositions?

I opted for the abstract language of sets of propositions because I aimed to offer an account as general as possible. Wright's language of "regions of thought" is vague, and I wanted to avoid a discussion of how to divvy out such regions because I suspect there is also substantial overlap among them. I do acknowledge that our belief sets are divided into subdomains—or regions of thought—with their specific hinges given that there are domain-specific entailment relations (Ohlhorst, 2023, p. 8). However, mostly I had the total set of believed propositions in mind when talking about cornerstones. The differentiation of subdomains or regions of thought permits

[4] An analogous formulation can be found for rules and belief changes. For instance, if an epistemic agent uses ampliative inference—infers Type II from Type I propositions—then they are committed to a corresponding rule.

also the introduction of local hinges; a set of propositions arguably has to be large enough to fix determinate meanings and inferences among themselves to give rise to local hinges.

This also explains why I take logical necessities to be cornerstone propositions. Given that logical necessities are entailed by the empty set, and hence any set, if I were to reject one of them, then I would be committed to reject everything I believe because it is a (vacuously) necessary condition. Volpe suggests that relevant entailment would be a more natural alternative here; but I actually take the axiomatic hinge certainty status of logical necessities to be an interesting result of its own.[5]

I introduced transmission of warrant when closure would have done as well because it highlights the rational commitments that we undergo when believing things. The purpose was to connect the epistemic structure of entitlement to the doxastic structure of hinges. Hinge epistemology is often reluctant to be nailed down to specifics about what makes something into a hinge. The entitlement framework serves as the foundation to spell out the underspecified structure of hinge epistemology.

As Volpe notes, I am quite guarded regarding epistemic alchemy. The reason is that, like Wright, I cannot deny the force of McGlynn's (2014) arguments. By submitting warrant to logical consequence, it is inevitable to conclude that justification transmits from ordinary beliefs to hinge certainties. I emphasised that alchemy does not influence what you believe to any degree because you were already certain. But Volpe is right, that this is not enough.

What is missing, and what I was gesturing at with the metaphor of justificatory gilding versus gold-making, is the justificatory status of hinges. The justification I gain for hinges through evidential alchemy does not go beyond the entitlement I had already in its warrant: All the justification I gain is nothing beyond justification to trust with certainty. That is, it is no justification to believe it like an ordinary proposition.

The reason for this is the blatant circularity that would arise otherwise: I would be justified to believe a hinge that I have to presuppose to be true in order to gain this justification. This is why there is no real alchemy occurring. The lead of entitlement *to trust* remains leaden because it only warrants trust. The gilding is that it now includes evidential justification (to trust). I take this point to be significant, because I do not think that alchemy offers a solution to the sceptical problem. The alchemic justification does not have the force to dispel sceptical doubts because it remains a justification to trust. As a result, this does indeed entail a weakening of transmission of warrant across known entailment, namely regarding what it is warrant to do. Still many doxastic aspects like degrees of confidence are preserved: I am certain that an external world exists, and I have evidence for that, but nevertheless I am only justified to trust that the external world exists.

Let us now turn to the final concern about how sustainable *Trust Responsibly's* internalism is. Volpe raises the worry that relying on a naturalist account of epistemic virtue to address the problem of demarcation threatens to undermine my account's ability to address New Evil Demon-style problems. The need to exercise

[5] Differently from Coliva (2015), I am open as to which logical necessities follow from a total belief set.

(naturalised) reliabilist Type 1 virtues gives us the entitlement to take on trust the existence, reliability and normalcy conditions for these virtues.

First, Volpe asks how does this not raise alchemy-style concerns, given that I can investigate how reliable my Type 1 faculties are with my other virtues, which obviates the need for entitlement? This strategy would only work for very local scepticisms about particular faculties, taking the others for granted. Local scepticism about a particular faculty is not at the heart of the sceptical problem. Verifying each faculty's reliability and virtuousness would end up in an unsatisfying coherentism. Our trust in our Type 1 faculties can be limited in their domain by evidence of local unreliability; for instance, we do not trust our vision to be reliable in darkness.

Second, Volpe points out that entitlement for trusting Type 1 virtues depends on Burgean (2003) normalcy conditions. Suggesting that these normalcy conditions are the ones under which the content of the perceptual and epistemic states is fixed for a particular organism. Normalcy here is fixed in reference to the dyad of the organism and its environment. As Volpe notes, in such a case, New Evil Demon-style considerations disappear because a lifelong brain in a vat that happens to receive exactly the same stimuli as I do, has different beliefs than I do because of Putnamian (1973) considerations. For a brain that has been in a vat for its entire existence, these stimuli *are* normal, and hence it cannot be in the same mental state as I am. This is an interpretation of *normalcy* that is more individualistic than I took it to be.

I (mis)interpreted Burge to have a less individualistic notion of normalcy. Namely, I consider the normal environment for a cognitive faculty *qua* biological faculty to be the range of environments in which it evolved, and hence, I take Burge's individualism to be mistaken. As a result, what determines the content of an epistemic state is, to a significant degree, the evolutionary history of an organism. To wit, both my Type 1 and Type 2 cognitive faculties work and are reliable because they evolved in particular environments (cf. Burge, 2003, p. 532). As my work on core cognition argues, a certain range of our hinges is innate (Ohlhorst, 2024) and hence the product of evolution. Consequently, as evolved human organisms with evolved human cognitive faculties, we are entitled to trust that these faculties are reliable and that their normalcy conditions obtain. Compare this to a virtue's background conditions that Sosa (2021, p. 127) argues we have to trust to obtain.

As a result, in New Evil Demon-style scenarios, any human brain in a vat with the relevant human evolutionary history will be in the same or at least a sufficiently similar internal mental state as I am. Note that this extends to lifelong envatment or science-fiction scenarios like the copying of brains into vats or simulating them—as long as there is some causal link to the evolutionary history of brains determining normalcy.

Nevertheless, Volpe's critique shows a limitation of my naturalist proposal. Namely, as soon as we leave the naturalist realm of evolved brains and their copies and imitations, we leave the realm of naturalist Type 1 and Type 2 virtues, which we have no other means but to help to trust.

There is then a range of sceptical scenarios that this naturalist approach cannot directly address. Anything that we might be willing to call swamp brains, that have appeared through some purely random non-evolutionary process, but which happen to be in the exact same neurobiological internal state as I am in right now, would not be accounted for because their normalcy conditions are individualistic *in virtue of their brief history*. Similarly, victims of malicious demons, immaterial souls created with cognitive faculties

so defective as to be systematically mistaken (Descartes, 2013), or butterflies dreaming that they are me (Zhuangzi, 1968), are bound to have very different normalcy conditions. Consequently, my account of entitlement to trust in naturalist Type 1 and Type 2 virtues does not capture more remote non-naturalist sceptical scenarios.

Can we avoid the limitation to naturalist virtues? I will very briefly consider the prospects of a non-naturalist virtue hinge epistemology here. As we saw, the issue is that while the popular naturalist New Evil Demon-style sceptical scenarios can be covered, there are versions of the argument that block the unifying naturalist background story. We would then need a non-naturalist account of epistemic virtue. In what follows, I will offer some extremely cursory suggestions on how such an account might be constructed.

If we take a phenomenological, and thereby strictly internalist, perspective on our cognitive processes, then we have awareness of the operations of our reliabilist and responsibilist faculties without having full access to or control over their functioning. We are aware of cognitive processes occurring that process and transform our cognitive states (Ohlhorst, 2023, p. 130).

These processes present their outputs as facts with a certain structure, and we have no other choice but to accept their framing at first hand, lest we end up in total scepticism. A victim of a Cartesian demon—to take the most extreme example—cannot help but trust how their deeply flawed faculties present things to them lest they end up in cognitive paralysis. What I am getting at here is that we cannot shed the cognitive schema that our cognitive faculties present to us (Strawson, 1959). This unavoidability then gives us entitlement of conceptual schema that Wright originally proposed (Ohlhorst, 2023, p. 55; Wright, 2004, p. 203). Given that our conceptual schema is shared in both naturalist and non-naturalist sceptical scenarios, this approach could potentially subsume *Trust Responsibly's* naturalist account under it.

Funding Open Access funding enabled and organized by Projekt DEAL. NeuroSys as part of the initiative "Clusters4Future" is funded by the Federal Ministry of Research, Technology and Space BMFTR (03ZU2106EA).

Data Availability No data are associated with this work.

Declarations

Ethics approval Not applicable.

Conflicts of interest No conflict of interest to report.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Burge, T. (2003). Perceptual entitlement. *Philosophy and Phenomenological Research, 67*(3), 503–548. https://doi.org/10.1111/j.1933-1592.2003.tb00307.x

Coliva, A. (2015). *Extended Rationality*. Palgrave Macmillan.

Descartes, R. (2013). *Meditations on First Philosophy* (J. Cottingham, Ed.). Cambridge University Press.

Hazlett, A. (2025). Transcendental entitlement and reasons for belief. *Asian Journal of Philosophy, 4*(1), 10. https://doi.org/10.1007/s44204-024-00234-y

James, W. (2014). The will to believe. In *The Will to Believe: And Other Essays in Popular Philosophy* (pp. 1–31). Cambridge University Press.

McGlynn, A. (2014). On Epistemic Alchemy*. In D. Dodd & E. Zardini (Eds), *Scepticism and Perceptual Justification* (Vol. 6, pp. 173–189). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199658343.003.0009

Moyal-Sharrock, D. (2004). *Understanding Wittgenstein's On Certainty*. Palgrave Macmillan.

Ohlhorst, J. (2022). Is there a problem of demarcation for hinges? *International Journal for the Study of Skepticism, 12*(4), 317–330. https://doi.org/10.1163/22105700-bja10040

Ohlhorst, J. (2023). *Trust responsibly: non-evidential virtue epistemology* (1st edn). Routledge. https://doi.org/10.4324/9781003374466

Ohlhorst, J. (2024). Baby logic – A hinge epistemology. *Erkenntnis*. https://doi.org/10.1007/s10670-024-00858-9

Pascal, B. (2005). *Pensées* (R. Ariew, Trans.). Hackett Pub. Co.

Pedersen, N. J. L. L. (2009). Entitlement, value and rationality. *Synthese, 171*(3), 443–457. https://doi.org/10.1007/s11229-008-9330-x

Pritchard, D. (2016). *Epistemic Angst*. Princeton University Press.

Putnam, H. (1973). Meaning and reference. *Journal of Philosophy, 70*(19), 699–711.

Sosa, E. (2021). Epistemic explanations. *Oxford University Press*. https://doi.org/10.1093/oso/9780198856467.001.0001

Strawson, P. F. (1959). *Individuals*. Routledge.

Volpe, G. (2024). Non-evidential virtue epistemology: Some queries about cornerstones, epistemic alchemy, and scepticism. *Asian Journal of Philosophy, 4*(1), Article 6. https://doi.org/10.1007/s44204-024-00227-x

Wright, C. (2004). I—Crispin Wright: Warrant for nothing (and foundations for free)? *Aristotelian Society Supplementary Volume, 78*(1), 167–212. https://doi.org/10.1111/j.0309-7013.2004.00121.x

Wright, C. (2014). On Epistemic Entitlement (II). In D. Dodd & E. Zardini (Eds), *Scepticism and Perceptual Justification* (pp. 213–247). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199658343.003.0011

Zhuangzi. (1968). *The complete works of Chuang Tzu* (B. Watson, Trans.). Columbia University Press.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.