Justification as Ignorance and Logical Omniscience

Daniel Waxman
National University of Singapore

Abstract

I argue that there is a tension between two of the most distinctive theses of Sven Rosenkranz's *Justification as Ignorance*: (i) the central thesis concerning justification, according to which an agent has propositional justification to believe p iff they are in no position to know that they are in no position to know p; and (ii) the desire to avoid logical omniscience by imposing only "realistic" idealizations on epistemic agents.

Sven Rosenkranz's *Justification as Ignorance* is a fascinating contribution to epistemology, offering not only a powerful and distinctive account of justification but also advancing the literature along many fronts. To mention just a few topics that are illuminated: the notion of being in a position (to know, and in general); luminosity and its absence; logical omniscience and the limits of idealization in epistemology; the metaphysical grounds of epistemic states; and the internalism/externalism debate. In time-honoured fashion, I will try to raise some challenges for the views defended in the book. But let me say at the outset that I have learned much from thinking through the details of the many rich arguments the book contains.

$J = \neg K \neg K$ and idealization in epistemology.

I will try to bring out what I see as a tension between two of the most distinctive positions advanced in the book. If nothing else, I hope that my comments will spur Professor Rosenkranz to elaborate on how these parts of his view hang together.

For a salient agent, let K and J be propositional operators expressing 'the agent is in a position to know that' and 'the agent is propositionally justified in believing that'. The main thesis of the book is what we might call:

$J = \neg K \neg K$: For an agent to be propositionally justified in believing some proposition ϕ is for them to be in no position to know that they are in no position to know ϕ.

Hence, justification as (a certain kind of) ignorance. $J = \neg K \neg K$ is advanced in the spirit of an 'interpretative hypothesis'.[1] To deal with possible counterexamples involving

[1] Rosenkranz (2021, 107).

such characters as 'small infants, dogs, trees, madmen, drunkards and dead people', the claim is advanced as applying only to what we might call *realistically idealized agents*:

> suitably improved versions of ourselves whose epistemic powers finitely extend our own, who can grasp every thought expressible in the language, and who have other epistemic virtues such as freedom of irrationality, bias and compulsion, freedom of attention deficiencies, and freedom of other ills that affect the epistemic lives of ordinary subjects.[2]

The case for $J = \neg K \neg K$ in *Justification as Ignorance* is multifaceted. Rosenkranz argues that it returns plausible verdicts in many clear cases; that apparent counterexamples can be defused; and that it has substantial explanatory and theoretical virtues.

The second distinctive position of the book concerns the question of appropriate idealization in epistemic theorizing. The vast majority of work in epistemic logic (I think it is fair to say) involves the strong idealization that epistemic subjects are logically omniscient: they know or are in a position to know every logical truth. *Justification as Ignorance* firmly rejects this idealization as unrealistic, implausible, and for many purposes theoretically unhelpful.[3] To my mind, this stance is one of the great strengths of the book: it is an important corrective to approaches to epistemology which ignore the fact that, for creatures like us, logical knowledge does not come for free, but often involves significant cognitive achievement.[4] Rosenkranz rejects logical omniscience on the basis of a conviction that any acceptable idealization involves, at most, 'improved versions of ourselves whose epistemic powers finitely extend our own'.[5] We can isolate this conception of idealization as a standalone thesis:

Realistic Idealization: In order for an epistemic idealization to be acceptable, it must concern agents whose epistemic powers at most finitely extend our own.

For this reason, in his 'realistic' system of epistemic logic, Rosenkranz rejects the following rules.[6]

RN_K: $\phi / K\phi$

[2]Rosenkranz (2021, 108)

[3]Rosenkranz (2021, 29).

[4]Indeed, as argued in Field (1984), it may be that all mathematical knowledge is fundamentally logical knowledge.

[5]Rosenkranz (2021, 58). As Rosenkranz (2021, 28) puts it: 'we are chiefly concerned with knowledge available to us, or at least to creatures that we resemble to a sufficient degree; and subjects with infinite powers are simply too powerful to resemble us to such a degree'.

[6]While these rules are included in Rosenkranz's 'idealized' system, he is explicit (Rosenkranz (2021, 86)) that they involve 'unacceptable idealizations' and are 'far too strong', to be rejected in the final analysis.

RM_K: $\phi \to \psi / K\phi \to K\psi$

where, again, the 'K' operator expresses not knowledge but rather *being in a position to know*. Even for finitely improved versions of us,

> there will be some logical theorems, and some of the infinitely many logical consequences of what they are in a position to know, that are too complex or too difficult to discern, in order for those subjects to be in a position to recognize them as true in all situations. To suppose otherwise is to deprive the notion of an epistemic situation – and hence that of being in a position to know – of much of its use. Accordingly, RN_K and RM_K should be taken to fail.[7]

Say that a proof is *surveyable* for an agent if they are in a position to know its conclusion on the basis of working through it. Even for realistically idealized agents, Rosenkranz's thought seems to be, not all proofs are surveyable: there is an upper bound on the length or complexity or difficulty of proofs, beyond which the proof is too long or too complex or too hard to comprehend.

Naturally, we can imagine agents who can comprehend longer or more complex or harder proofs than can we. But if these agents are to have epistemic powers that merely *finitely* improve our own, there must nevertheless be an upper bound on the complexity of the proofs that they themselves are able to survey. There is thus a potentially infinite sequence of idealized agents, each of whose logical powers finitely exceeds the last (and thus our own). But no agent in this sequence has the ability to comprehend proofs of *unbounded* complexity; that would require an agent altogether different in kind, possessing abilities that *infinitely* improve ours.

A tension between $J = \neg K \neg K$ and Realistic Idealization.

Consider a realistic agent, say, me, realistically idealized so as to be free from irrationality, bias, compulsion, attention deficits, and so on. As above, there will be an upper bound on idealized-me's ability to comprehend logical proofs. To fix ideas, let's suppose that a rough bound is (perhaps generously!) something like a book-length proof consisting of advanced, research-level, logic or mathematics, involving complex definitions, appeals without proof to more elementary theorems, and so on. Now let us pick a particular mathematical claim – say the twin prime conjecture (t for short), for which there is no proof (even tersely presented) that would take less than a book to

[7] Rosenkranz (2021, 58)

write out.[8] If the case is as described, I think the following claims hold.

First, **I am in no position to know** *t*. This is straightforward from the setup of the case, since there is no surveyable proof of *t*.

Second, **I am not propositionally justified in believing** *t*. By hypothesis, I have not proven *t*, nor am I in a position to prove it. I have (let us suppose, as surely we may) no abductive grounds to prefer *t* over $\neg t$, and there are none that I am in a position to acquire; similarly (let us again suppose) there are no testimonial or other non-logical grounds to the same effect. Intuitively, it is overwhelmingly natural to say that in such a scenario I do not have justification to believe *t*. Theoretical accounts of propositional justification yield the same result. Glosses like the following are common: one has propositional justification to believe *p* iff one is in a position where it would be epistemically appropriate for one to believe *p*; or iff one has sufficient epistemic reason to believe *p*.[9] But for a mathematical or logical claim, it is the *availability of a proof* that canonically makes it epistemically appropriate, or provides epistemic reason, to believe. And in the case under consideration, my cognitive limitations mean that no such proof is available to me.

Third – and here is the kicker – **I am in no position to know that I am in no position to know** *t*. Why? Let us think of methods I might use to come to know that I am in no position to know *t*. Again, we might elaborate the case so that I get abductive or testimonial evidence, sufficient for knowledge, that $\neg Kt$. (Perhaps generations of the best mathematicians have tried in vain to come up with a proof; or perhaps Terence Tao tells me that he has found a proof of *t* or its negation, without telling me which, but credibly informs me that I would be in no position to comprehend it and that there is no shorter proof.) But equally well, we might elaborate the case so that no such evidence is available.

The only method I can think of that might be applicable in full generality for realistically idealized agents is something like this: (a) I first come to know some upper bound on the complexity of proofs I am able to survey; and (b) I then consider the totality of possible proofs below that bound, examine each of them, and discover that none is a proof of *t*. It may well be that step (a) is always achievable for sufficiently idealized agents, if, not unreasonably, they are credited with sufficient introspective and abductive powers. (I take it that it is not so hard for agents like you or I to come to appreciate our own logically limited natures.) The real problem, though, is with (b).

[8] If a truly logical (not mathematical) example is desired, just consider the conditional whose antecedent is the conjunction of the relevant mathematical axioms (or some appropriate finite surrogate, in the case of infinite theories) and whose consequent is the claim in question.

[9] See e.g. Pryor (2005, 181) and Ichikawa and Steup (2018). Pryor's gloss is offered for 'having justification to believe *p*', which he contrasts with 'actually appropriately holding a belief in *p*'; I take it that this is simply a notational variant of the usual propositional/doxastic distinction.

Remember that we are considering a realistically idealized version of me, with limited computational and logical abilities. Carrying out step (b) is wildly beyond anything that I am in a position to do. If it would push my cognitive capacities to their limit to work through a single advanced book-length mathematical proof, then there is surely no chance whatsoever that, given those same capacities, I can exhaustively enumerate and consider *all* proofs of a similar length or complexity.

We can put the point more pithily by extending the notion of surveyability as follows. For an agent A and property P, say that a set S is P-surveyable for A if A is in a position to sequentially run through the elements of S and check, of each, whether it has P. The relevant property for our discussion is 'being a proof of t', which I'll henceforth suppress.[10] So, if a set is surveyable for me, then I am in a position to know whether some/all/which members of S are proofs of t on the basis of exhaustively considering all the possibilities.

In these terms, the concern is fundamentally this. In order to be in a position to know that I am not in a position to know t, the set of surveyable proofs must itself be surveyable; but it is not; so I am not in a position to know that I am not in a position to know t. But if so, then it follows – I would say absurdly – from $J = \neg K \neg K$ that I am propositionally justified in believing t. Hence the advertised tension.

More on the unsurveyability of the set of surveyable proofs.

It would be desirable, if possible, to make the foregoing remarks more precise by providing formal models of (a) the cognitive limitations of realistically idealized agents and (b) the 'complexity' or 'difficulty' of proofs. A very natural, if perhaps crude, way to handle the notion of the complexity of a proof is in terms of its length. This is precisely how the problem is approached in the mathematical field of proof complexity theory, which discloses deep connections between logic and computational complexity theory.[11] If we additionally view agents as bounded in the length of proofs they are

[10] Checking whether a particular string is a proof of t is relatively computationally easy – it can be carried out in polynomial time relative to the length of the string.

[11] See for instance Cook and Nguyen (2010). One of the most robust findings is that, for almost all commonly studied proof systems, it is computationally hard to find whether there exists a short proof of some sentence. A problem is typically regarded as *feasibly computable* if it can be solved in polynomial time (i.e. the complexity class P); by contrast problems whose answers can be *verified* in polynomial time (i.e. the complexity class NP) often appear computationally intractable. While $P \subseteq NP$, it is famously open whether $P = NP$; one source of the common view that $P \neq NP$ is precisely the apparent difference in computational difficulty. See Dean (2021) for more.

But, as Buss (2012) puts it, "For almost all common proof systems (resolution, Frege, nullstellensatz, sequent calculus, cut-free sequent calculus, etc.), it is impossible to approximate shortest proof length to within a factor of $2^{log^{1-o(1)}n}$ in polynomial time, unless $P = NP$... the theorem states more than it is difficult to search for a short proof; instead, it is already hard to determine whether such a proof exists (assuming $P \neq NP$)." If anything, this is further evidence that the problem of surveying the set of surveyable proofs is far from computationally or logically trivial.

able to comprehend – with limited random access memory, so to speak – it is trivial to show that no agent will be able to cognize even two distinct proofs close to the upper bound of surveyability, let alone the totality of all surveyable proofs.[12]

On some ways of counting the length of a proof (e.g. the number of lines it contains), then there will be infinitely many proofs of any given length. But I don't think the tension I am exploring needs to exploit this infinitude. On other ways of counting (e.g. the number of symbols occurring in a proof), there will only be finitely many proofs of a given length, up to alphabetic variation. And indeed, perhaps knowledge that there is no surveyable proof of t does not require an *exhaustive* search through *all* surveyable proofs, but only some restricted, finite set of proofs, perhaps those which use premises or inferences that are 'relevant' to the conclusion. But even so, the relevant set of proofs will still be vast. Returning to our informal example above, think of how many possible book-length proofs there are that are relevant to number theory or to the theory of primes or even ('just') to the theory of prime progressions. Again, if comprehending a single book-length proof would push my cognitive capacities to their limit, it is surely an understatement to say that comprehending the class of all relevant shorter proofs is far beyond anything I am in a position to do.

Perhaps there is some realistically idealized agent A 'above me' in the sequence of idealized agents who really *can* run through all of the proofs that are surveyable-for-me. But that doesn't help: the very same cognitive resources that allow A to run through the totality of these proofs (as I cannot) will mean that A is able to cognize much more complex proofs than can I, and so surveyability-for-A exceeds surveyability-for-me. And then the same argument can be run: the set of surveyable-for-A proofs will be unsurveyable-for-A; so, if t is some claim with no proof that is surveyable-for-A, A will be in no position to know that they are in no position to know t.

Of course, it may be that this model of agents or this measure of proof complexity is too crude or otherwise deficient. Indeed, some of Rosenkranz's other commitments seem to require an alternative. For instance, he endorses (Rosenkranz (2021, 60)) the schema:

$$K\phi \leftrightarrow K\neg\neg\phi$$

and others which seem to imply that agents are in a position to carry out arbitrarily long proofs. Even though this arguably violates Realistic Idealization, Rosenkranz justifies it on the basis that

[12] Given plausible background assumptions, thinking of agents as bounded in this way is equivalent to conceiving of them as capable of carrying out only a finite number of inferential steps, as in Bjerring and Skipper (2019).

> the further idealization that is needed [...] solely concerns the subjects' powers to entertain propositions of arbitrary complexity built from a finite stock of constituents, and not their epistemic powers to evaluate them for truth or falsity.[13]

It is not entirely clear to me how this passage is intended; after all, in order for me to *know* the double negation of anything I am in a position to know, surely I must do more than merely *entertain* the double negated sentence. But perhaps the idea is something like this: for reflective agents who are in a position to know ϕ, $\neg\neg\phi$ is sufficiently 'close' to ϕ that no new proof idea is needed to demonstrate $\neg\neg\phi$; it is merely a mechanical or combinatorial exercise. But here, two points are worth making. The first is that this thought is not obviously defensible; after all, *any* proof can be broken down into transitions between sentences which seem as 'close' as do ϕ and $\neg\neg\phi$. So we had better not trivialize the cognitive work involved in stringing together a substantial number of steps of this kind, at least not if we wish to resist logical omniscience. The second point is that, even if the thought is defensible, it is not at all obvious how it helps to show that the set of surveyable proofs is surveyable. If the complexity of a proof is measured in terms of the number of distinct proof ideas that are involved in it (however exactly these are individuated), and there is an upper bound on the number of proof ideas I am in a position to comprehend, then surveying the entire set of relevant surveyable proofs will surely involve comprehending far more proof ideas than that upper bound.

Conclusion

So I think there is a robust case that, however exactly we flesh out the notion of a realistically idealized agent, there will be sentences, like t, which they are in no position to know that they are in no position to know. Given $J = \neg K \neg K$, the problematic conclusion follows that they are propositionally justified in believing such sentences.

I can think of only a few ways out. One is to try and fine-tune the notion of a realistically idealized agent in such a way that the set of surveyable proofs comes out as surveyable. As I've argued above, this is a tough needle to thread, especially if logical omniscience is to be avoided.

Another would be to find some other method by which realistically idealized agents are in general in a position to know that they are not in a position to know claims like t. I cannot see how this might be done: after all, surveyable proofs are ones that I am in a position to work through and thereby know the conclusion. How in general could I know that I am in no position to know t without ruling out the existence of a surveyable

[13] Rosenkranz (2021, 61)

proof? And how in general could I safely rule this out without running through the (relevant) possibilities? But perhaps there is some other method I am failing to consider.

Another option still would be to bite the bullet and hold that realistically idealized agents really do have justification to believe claims like t. As I argued above, this seems both intuitively implausible and difficult to reconcile with standard understandings of propositional justification. This option also threatens to compromise the logic of justification. For any realistically idealized agent, there will be some claim p such that neither p nor $\neg p$ has a surveyable proof.[14] Then the view is committed to saying that both p and $\neg p$ are propositionally justified, violating the D axiom for justification.[15] Finally, this response seems to me to sit particularly badly with a vision of epistemic agents who are in no position to *know* claims like t in virtue of their logical non-omniscience. Why should propositional justification come so far apart from what the agent is in a position to know in such cases?

My own tentative suspicion is that none of these approaches will succeed, and thus that there is a serious tension between $J = \neg K \neg K$ and Realistic Idealization. How exactly the tension is to be resolved (or whether it can be sidestepped) is a matter for further discussion.[16]

References

Bjerring, J. C. and M. Skipper (2019). A dynamic solution to the problem of logical omniscience. *Journal of Philosophical Logic 48*(3), 501–521.

Buss, S. R. (2012). Towards NP–P via proof complexity and search. *Annals of Pure and Applied Logic 163*(7), 906–917.

Cook, S. and P. Nguyen (2010). *Logical Foundations of Proof Complexity*. Cambridge University Press Cambridge.

Dean, W. (2021). Computational Complexity Theory. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2021 ed.). Metaphysics Research Lab, Stanford University.

Field, H. (1984). Is mathematical knowledge just logical knowledge? *The Philosophical Review 93*(4), 509–552.

Ichikawa, J. J. and M. Steup (2018). The Analysis of Knowledge. In E. N. Zalta (Ed.),

[14] A realistically idealized agent will only be in a position to carry out finitely many proofs, and thus there will be infinitely many claims of this form. For all we know, the twin prime conjecture is among them.

[15] The D axiom for justification is not beyond doubt, but rejecting it would be costly to Rosenkranz; it is entailed by both the systems proposed in Rosenkranz (2021, Ch 5). Thanks to a referee for pointing out that this option would lead to failures of D.

[16] Thanks to an anonymous referee, Lavinia Picollo, Mattias Skipper, and Jared Warren for helpful comments and discussion.

The Stanford Encyclopedia of Philosophy (Summer 2018 ed.). Metaphysics Research Lab, Stanford University.

Pryor, J. (2005). There is immediate justification. In M. Steup and E. Sosa (Eds.), *Contemporary Debates in Epistemology*, pp. 181–202. Blackwell.

Rosenkranz, S. (2021). *Justification as Ignorance*. Oxford University Press.