

Review of Yli-Vakkuri and Hawthorne's *Narrow Content**

Ethan Jerzak

October 17, 2020

It might seem, and has seemed to many, that what I think is up to me. I go about life representing the world with thoughts, and my intrinsic state fixes the content of those thoughts—fixes, that is, what they require of the world in order to be true. Call this idea *internalism about mental content* (henceforth, internalism). Juhani Yli-Vakkuri and John Hawthorne (henceforth, YVH) have written a book that attempts to refute internalism in all its reasonable manifestations.

There is much of value in *Narrow Content*. The central argument, with its protagonist Mirror Man, constitutes a formidable stumbling-block that all future internalists will have to contend with. And the authors have done substantial work to iron out a more precise characterization of the conceptual landscape than existed hitherto. That said, the book is not without frustrations, for reasons both stylistic and substantive. Stylistically, while the authors are admirably clear about defining the views at issue, they're not always great about explaining why the definitions are as they are, especially in cases where their framework is more complicated than what you might have expected. Substantively (and, I'll argue, relatedly), the authors largely neglect to address one of the most prominent roles certain philosophers (e.g. Dennett, Stalnaker, and Lewis) have taken content to play, that of explaining and predicting behavior. It's not clear, so I'll argue, that the central argument of the book refutes internalists of that kind.

The main thread

Narrow Content takes up a project invented by Putnam and Burge. Putnam (its closest spiritual predecessor) convinced many that the content of our thoughts about water depends on more than us; it depends also on what has caused our water experiences. Here on Earth, that's H_2O , so my thought that I behold water is true just in case I behold H_2O . My qualitatively identical twin, on a planet like Earth except that the watery stuff there comprises XYZ , has a thought internally just like mine. But if he beheld H_2O , his thought would be false, for his thoughts concern XYZ . Since our qualitatively identical thoughts in our qualitatively identical heads differ in truth, what they require of the world to be true (their content) is not fixed internally.

The opponents of Putnam/Burge whom YVH are concerned to refute are those who try

*Penultimate draft; forthcoming in *Philosophical Review*. Thanks to Zach Barnett, Simon Goldstein, John Hawthorne, Arc Kocurek, Sven Neth, Rachel Rudolph, and Dan Waxman for helpful discussions.

nonetheless to find a “theoretically interesting” notion of content shared between me and my twin. Two-dimensionalists, e.g., hold that while my twin and I entertain different contents in some sense, we share content in another. On the simplest version, *primary* contents (the narrow ones) are functions from world-agent pairs $\langle w, a \rangle$ to truth values. My twin and I both believe that a beholds the stuff that generally plays the water role for a in w . However his thought might still be false and mine true, for this content gets evaluated relative to different agents.

The central claim of *Narrow Content* is that enrichments like this will not do. YVH purport to describe a *single* agent, Mirror Man, who, at a single time, thinks two qualitatively indistinguishable thoughts that differ in truth value. Think of Mirror Man as Ethan and Twin Ethan crammed into one body. Mirror Man is a left-right symmetric agent. Every time he thinks a thought with his left side, he thinks a corresponding thought with his right. With his right side he perceives Kit Fine, and truly judges him to be human. With his left side he perceives a perfect wax copy of Kit Fine, and falsely judges it to be human. Two qualitatively identical thoughts, same world, same agent, different truth values. Thus either more indices are required, or content does not supervene narrowly.

This is the core of what YVH call the *parameter proliferation argument*. Further epicycles abound, all with a common structure: Whatever new indices the narrow content theorist adds to the supervenience base, YVH whip up some novel variation of the Mirror Man story to supply two thoughts with the same indices of occurrence but intuitively different truth-values.

To formalize this argument, YVH must give a definition of internalism that allows them to distinguish simultaneous, qualitatively identical thoughts within a single agent. As we’ll see, this constitutes some of the more interesting aspects of the book, and its more frustrating ones. They spend Chapter 1—all 45 pages of it!—defining internalism. Chapter 2 sketches their argument against narrow content. The next two chapters deal with specific job descriptions for content in light of this argument: serving as the referents of that-clauses (Chapter 3), and explaining *a priori* entailment between thoughts (Chapter 4). The final two chapters take up the prospects for weakened versions of internalism—weakened by dropping the *qualitative* supervenience requirement (Chapter 5), and by allowing relations between thoughts to figure into the supervenience base (Chapter 6).

Qualitative agential profiles

You might be surprised—I was surprised—that it takes 45 pages to define internalism. Why does it take so long to flesh out the basic idea that what I think depends only on me? Indeed, Williamson (e.g.) requires but one paragraph for this task:

A case [basically a centered world] α is internally like a case β if and only if the total internal physical state of the agent in α is exactly the same as the total internal physical state of the agent in β . A condition C is narrow if and only if for all cases α and β , if α is internally like β then C obtains in α if and only if C obtains in β . In other terminology, narrow conditions supervene on or are determined by the internal physical state: no difference in whether they obtain

without a difference in that state. (Williamson 2000, 52)

The kinds of conditions relevant for narrow content are ones like “believing that one beholds water”. If the holding of that condition depends only on the complete physical state of the relevant agent, that condition is (more or less) what YVH call “coarsely narrow”.

YVH’s definition requires more machinery. The internal part of Williamson’s cases are states of whole agents, and the relevant conditions apply (or not) to whole agents. This won’t do for YVH’s purposes; their definition applies to particular token thoughts within an agent. It relies on what they call the *qualitative agential profile* (QAP) of a thought, i.e. the conjunction of all intrinsic qualitative relations that hold between the thought and its bearer. A content assignment, assigning contents to thoughts, is *narrow* just in case it supervenes on QAPs. *Internalism* is the view that there is a theoretically interesting narrow content assignment satisfying plausible conditions.

With these definitions in hand, YVH can formulate the Mirror Man argument. Since Mirror Man is symmetric, his left thought l and right thought r have the same QAP. (Assuming, that is, that “being located in the left side of” does not constitute a qualitative relation—more on that later.) So any narrow content assignment must assign l and r the same content, in which case the thoughts, having as they do the same index of evaluation, must be true or false together, contradicting our intuitive judgments of the case.

Why so complicated?

Why, apart from being indispensable for their argument, do YVH think we need to switch from coarse narrowness to QAPs? The answer comes on page 48, in the form of a case purportedly pulling narrowness apart from coarse narrowness. They ask us to imagine two duplicate brains each with a red and a blue hemisphere. Each hemisphere contains exactly one thought. Take a content assignment according to which both brains believe p and q , but in Brain 1 the *blue* hemisphere’s thought gets assigned p whereas in Brain 2 the red hemisphere’s does.

This, YVH say, is consistent with that content assignment’s being coarsely narrow but not with its being narrow in their sense. The idea is that Brain 1 and Brain 2 as wholes believe all the same things, and so constitute no counterexample to coarse narrowness. But since which half realizes which belief changes without any corresponding change in the QAP, the content assignment is not narrow.

Ok, fine, the two notions come apart—but which is closer to the intuitive idea? YVH argue that it must be theirs, on the basis of the following modification: Any possible duplicate of Brain 1 gets p assigned to the blue hemisphere iff it is located within a mile of a baboon (otherwise the red). “In this example”, YVH contend, “it is clear that the ur-content of a thought depends on the environment of the agent of the thought, so is not intuitively narrow. It is also not narrow according to our definition. Yet content is coarsely narrow in the example” (48). QED; the finer-grained definition that accommodates Mirror Man is actually necessary to elucidate the underlying concept.

However, those who accept a minimal kind of holism about propositional attitudes are unlikely to be moved by this argument. YVH bake into the description of the case that Brain 1 and Brain 2 have thought tokens located in various parts of their brains, and trade on the intuition that if which part realizes which belief depends on whether a baboon is nearby, that content assignment is intuitively wide. But the very idea that distinct thoughts can be pinpointed to specific brain regions is one that many would happily deny. Stalnaker, for instance, denies that we have *any* way of isolating a particular thought token without reference to content:

It is not at all clear in the case of beliefs, intentions, and other states and attitudes that one can identify something that is the belief or intention in abstraction from its content, something about which we can ask, what would the content of that belief have been if it had been a belief I had on Twin Earth. (Stalnaker 1990)

Stalnaker's thought is that we have a grip on the notion of contentful mental states largely by reference to the explanation and prediction of the behavior of agents. To ascribe to Fido the thought that a bone-treat is forthcoming is to expect Fido to behave in certain bone-treat-directed ways—by, e.g., tail-wagging, greedy staring in the direction of the bone-treat closet, pitiful whines in the direction of his owner, etc.

YVH are pretty clearly adopting the language-of-thought model that Stalnaker lambastes in Stalnaker (1984). They are happy to assume—indeed as parts of the very descriptions of their critical cases—that thoughts come in tokens, have parts, have physical locations, are intelligible as thoughts independently of their contents (so that different content assignments can assign different contents to the same one), etc. *Officially* they are not committed to a specific language of thought thesis, and they acknowledge, e.g. on page 20, that “perhaps some thoughts are diffusely located, perhaps underwritten by the agent's pattern of behavioural dispositions”—but key moves made later on in the book belie this allowance.

This red/blue hemisphere business precisely one such move. We are to accept as part of the description of the case itself that there is a thought lurking in the red half, and another lurking in the blue half, about which it makes sense to ask about their differing contents. It is hard to imagine Stalnaker acquiescing without complaint. The relevant agent may well believe p and q , but that's as fine-grained as it goes. About no particular entity in his colorfully bifurcated brain does it make any sense to ask, “is *this* the thought whose content is p ?”

Some complaints in this vicinity troubled me, I confess, on my first encounter with Mirror Man himself. What makes Mirror Man a single agent, rather than a pair of Siamese twins? What *does* Mirror Man as a whole believe? Are his halves functionally integrated? Is it an accident that his halves are symmetric, or is he robustly set up that way? No one can object to YVH providing us with a physical description of an entity and then asking for our intuitions about its beliefs and desires; but to bake into the very setup that Mirror Man is a single agent, in which two token thoughts separately occur, each located determinately in a different hemisphere—this is to paper over by fiat what I would think are the most interesting questions to ask about their creature.

And anyway, I suspect that many who incline towards internalism do so partly on the basis of behavioristic inclinations. Every undergraduate has thought: “But Oscar and Twin-Oscar would *act* in all of the same ways in the presence of their redistributed liquids! Neither of

them would complain upon quenching his thirst with the wrong one! Doesn't that mean *something* importantly contentful is shared?" Correct or not, that this is a natural way to state the internalistic hunch ought to make one a bit suspicious of the machinery YVH require for their supposed refutation.

Left and right

One final thought about the strength of the thesis *Narrow Content* purports to establish. Its central argument relies on the idea that left-right orientation does not constitute a qualitative property. If the property of being located on the *left* side of its agent were a qualitative property of a thought token, then Mirror Man's left and right thoughts would differ in QAP. YVH consider on p. 90 a symmetry-breaking response in terms of the different *phenomenologies* of left and right, but reject it: they are committed, and must be committed, to the idea that "a pair of left and right gloves could be intrinsic duplicates".

There are, I am told, chemical molecules called *enantiomers* that are perfect molecular mirrors of each other. Oranges and lemons, for example, both contain a molecule called limonene. It comes in two orientations: one finds limonene+ in oranges, and its enantiomer, limonene-, in lemons. The gustatory difference between the two fruits owes, apparently, to this difference in orientation.¹

Suppose (now counterfactually) that oranges and lemons were perfectly identical, save for the difference in limonene+ vs limonene-. Then if YVH are correct, an orange and a lemon might be perfect qualitative duplicates. Limonene+ and limonene- do not differ in any respect save orientation; their different flavors arise from how they fit into our differently oriented tastebuds.

However, one looking to bake a lemon tart would consider himself duped, were his grocer to sell him an orange when he requested a lemon. He could quite reasonably complain that there was something wrong with the fruit he was handed. Now, perhaps for esoteric metaphysical reasons, that complaint is *strictly speaking* incorrect. Nothing is wrong with the fruit; it's a perfect qualitative duplicate of one that would yield a proper lemon tart. Something is wrong only with its relation to the baker's taste buds.

If that's right, then, contrary to what one might have expected, the condition of being a lemon (or a limonene+ molecule) is wide. It depends not only on the intrinsic qualitative properties of the fruit, but also on how its molecules happen to be oriented. But this hardly suggests that its lemonhood is constitutively dependent on its environment in any particularly deep sense. Suppose that left-right orientation is fixed by convention. There is, let us say, a left-handed glove in a Paris vault that fixes what "left" means. Then it's plausible that lemonhood, while not officially narrow, is wide only in a rather uninteresting sense: the fruit's intrinsic qualitative properties, together with one further non-qualitative fact relating it to the Paris Glove, plausibly determine its lemonhood.

¹No chemists were consulted in the writing of this review, and if the foregoing contains errors in this regard, take the above to describe a mere metaphysical possibility. Since a box prepends each of YVH's major claims, I trust they will not complain on this account.

When Oscar was reported to have different beliefs from his twin, this was taken to show that content constitutively depends on the surrounding environment. It's actual intercourse with actual water that makes my water-thoughts concern H_2O . This is what people have in mind when they say that content is wide.

Consider now a slight weakening of internalism: A content assignment is *quite narrow* iff it supervenes on a thought's QAP, together with one further fact: whether it occurs in the same side of its agent as the direction the Paris Glove's thumb points when observing it from the top. Granted this allows some amount of extra-qualitative stuff to factor into the supervenience base. But it's a far cry from the constitutive environmental dependence that fans of wide content have advocated. Thought depends on its environment to no greater degree than a lemon's lemonhood depends on its environment. Bare sameness of orientation with respect to the Paris Glove all that's required; other than that, QAP might well suffice. If that is all that the rejection of Narrow Content amounts to, it may not strike many as much.

Bibliography

Stalnaker, Robert. 1984. *Inquiry*. Cambridge, MA: MIT Press.

———. 1990. "Narrow Content." In *Propositional Attitudes: The Role of Content in Logic, Language, and Mind*, edited by C. Anthony Anderson and Joseph Owens. Stanford: Csl.

Williamson, Timothy. 2000. *Knowledge and Its Limits*. Oxford: Oxford University Press.

Yli-Vakkuri, Juhani, and John Hawthorne. 2018. *Narrow Content*. Oxford University Press.