



Comment on Yli-Vakkuri and Hawthorne, *Narrow Content*

Alex Byrne¹

Accepted: 2 July 2020 / Published online: 22 September 2020
© Springer Nature B.V. 2020

Abstract This comment mainly examines Yli-Vakkuri and Hawthorne’s preferred framework for examining whether narrow content is viable, arguing that their framework is not well-suited to the task; once a more traditional framework is adopted, Y&H’s case against internalism is strengthened.

Keywords Narrow content · Internalism · Externalism

Narrow Content is a tremendously good book; of necessity, this comment will leave large parts undiscussed. I will concentrate on Y&H’s preferred framework for discussing the issue of broad versus narrow content, arguing that it is not well-suited to the task; once a more traditional framework is adopted, Y&H’s case against internalism is strengthened. I shall end by briefly mentioning an appealing internalist picture that their otherwise comprehensive critique does not address.

1 Framing

Y&H begin by characterizing the mental phenomena of interest, “intentional states”, paradigmatically “thoughts, beliefs, and hopes” (Yli-Vakkuri and Hawthorne 2018; 1).¹ As they say, there are a number of grammatical constructions that we use to talk about intentional states. In the case of belief, these include

¹ All page references to this book unless otherwise noted.

✉ Alex Byrne
abyrne@mit.edu

¹ Department of Linguistics and Philosophy, MIT, Cambridge, MA 02139, USA

nominalizations like ‘Juhani’s belief that snow is white’, ‘the belief that snow is white’, gerund phrases like ‘believing that snow is white’, and sentences with ‘that’-clausal-complements like ‘Juhani believes that snow is white’. ‘That’-clauses are not essential, as in ‘Juhani has many true beliefs’ or ‘Juhani hopes for snow’, but as Y&H note, philosophers have favored “‘that’-clause-involving constructions for describing intentional states”, taking the clause to be “a canonical way of specifying the content” of the state, and they are happy to join the majority. Y&H adopt a technical understanding of ‘thought’, on which it applies to “all intentional states that are aptly described by ‘that’-clauses” (1).

The main issue of the book—which receives sophisticated refinement and elaboration—is “whether contents are fixed by inner goings-on” (15). That is: are an agent’s thoughts—intentional states with content specified by that-clauses—determined by how the agent is internally? For example, could there be two people, alike with respect to “inner goings-on”, one of whom believes that snow is white (or has a belief with the content that snow is white), and the other of whom doesn’t? It is not news that the standard post-Putnam-and-Burge answer is “no”. That is, an uncompromising *internalism about mental content* is false. Neither is it news that internalists have a variety of more-or-less concessive responses, yielding a range of theses that purport to preserve the core spirit of internalism. The aim of *Narrow Content* is to shrink a viable form of internalism down to nothing.

Here is a familiar way of setting up the narrow content debate that is *not* Y&H’s. A gerundial phrase like ‘believing that snow is white’ refers to a *state* or *condition* or *property*. ‘Oscar believes that snow is white’, and ‘Oscar is in the state of (is in the condition of/has the property of) believing that snow is white’ are equivalent. If Oscar and Twoscar both believe that snow is white, then they are in the very same state, namely the state of believing that snow is white. There is nothing episodic or eventful about this state, as indicated by the absence of progressive tenses: *Oscar is believing that snow is white. Believing that snow is white is *relational* state: it is the state of bearing the relation expressed by ‘believes’ to the content or proposition *that snow is white*. A perspicuous rendering of:

- (a) x Vs that p ,
is:
- (b) $V(x, \text{that } p)$.
(See, e.g., Stalnaker 1988: 150.) Or, alternatively:
- (c) $\lambda yV(y, \text{that } p). x$.

Granted this setup, the basic internalist thesis, in the case of the belief that snow is white, is that $\lambda yV(y, \text{that } p)$, the state of believing that snow is white, is *intrinsic*, in the sense of being necessarily shared between perfect duplicates or doppelgängers, agents with the very same “internal goings-on”.

On this conception, the relevant content properties that the internalist claims are intrinsic are properties of agents—*believing that snow is white*, and the like. Y&H call this conception *coarse narrowness*. This, they argue, “cannot be used to draw various distinctions that our notion of narrowness can and that we think are essential for theorizing in this area” (47).

Here is Y&H's preferred alternative. They assume that (a) "is truth-conditionally equivalent" (21) to:

- (d) $\exists e(\text{AGENT}(e, x) \wedge \text{CONTENT}(e, \text{that } p) \wedge \text{V}(e))$,²

where the values of 'e' are "representational events (states or episodes) that occur in an agent at a time" (18). *CONTENT* relates a representational event or thought to exactly one content or proposition, which Y&H call the *ur-content assignment*. A *content assignment* in general is simply anything formally like *CONTENT*—of course there are numerous content assignments of no theoretical interest at all.

Y&H's framing takes a leaf from the neo-Davidsonian program (Davidson 1967), on which (at least many) verbs have an implicit argument place for an event. Thus a neo-Davidsonian might render:

- (e) Juhani buttered the toast.

As:

- (f) $\exists e(\text{AGENT}(e, \text{Juhani}) \wedge \text{PATIENT}(e, \text{the toast}) \wedge \text{BUTTER}(e))$.

Here the transitive verb 'to butter' vanishes, to be replaced by a one-place predicate 'BUTTER' ("is a buttering") of events. A neo-Davidson will say that (something like) (d) and (f) are the "logical forms" of (a) and (e), and Y&H are sympathetic. However, as they note, they need no such asymmetrical claim about logical form. For their purposes, the equivalence of (a) and (d) (in particular) will do.

Instances of the schematic (a) and (d) are:

- (g) Juhani believes that snow is white.

And:

- (h) $\exists e(\text{AGENT}(e, \text{Juhani}) \wedge \text{CONTENT}(e, \text{that snow is white}) \wedge \text{BELIEF}(e))$.

'Butter' is a dynamic verb, used to describe an action unfolding in time—Juhani was buttering the toast for five minutes, say. 'Believe', on the other hand, is a stative verb, used to describe someone's enduring state or condition. So although there is a "buttering event" occurring from midnight to five minutes past, when Juhani was buttering the toast, there appears to be no "believing event" occurring during the period when Juhani believes that snow is white. In other words, if stative verbs like 'believe' have a Davidsonian implicit argument place, this does not bring events into the picture. Borrowing a term from Bach (1986), let us use 'eventuality' as an umbrella term for the items quantified over by (4) and (6). This is a minor terminological departure from Y&H, who are explicit that they use 'event' promiscuously, to include entities that "are more naturally described as 'states' than as 'events'" (18, fn. 2).

Y&H label the eventualities invoked by sentences ascribing intentional states 'thought tokens' or 'thoughts' (18). Thoughts are the "primary bearers of content" (18) and, as already noted, they are supposed to "occur in an agent at a time." (Here 'occur'

² Changing Y&H's numbering.

should be construed promiscuously, so states can be said to “occur”). In this framework, the internalist holds that the content of a thought is determined, not by its intrinsic properties, but by its intrinsic structural position inside the agent, which Y&H call its “Qualitative Agential Profile” (QAP). Analogously, if John and Juhani have duplicate houses, the brick in the top left corner of John’s front porch has the same intrinsic structural position as the corresponding brick in Juhani’s house (cf. 30). If John and Juhani are duplicate agents, then any thought α in John has a corresponding thought β in Juhani with the same QAP—the thoughts are “agential duplicates”.

Coarse narrowness—*believing that snow is white* is intrinsic—has accordingly been traded in for Y&H narrowness: any “thought” inside an agent that corresponds to *believing that snow is white* has its content determined by its intrinsic structural position.

This only makes sense if thoughts—eventualities of the sort quantified over in (f)—occur (wholly) *in* agents. If they do not, then duplicating John will not guarantee that his duplicate has corresponding thoughts, any more than duplicating John guarantees that his duplicate has corresponding graduate students. And there seems to be no obvious reason why the pertinent eventualities do occur in agents—that sounds like something an externalist enamored with neo-Davidsonism might resist. Certainly that is not true for eventualities in general. Consider the classic toast-buttering example (f). Here the buttering eventuality occurs (at least partly) outside the agent, where the toast and butter are, and so duplicating Juhani will not guarantee that his duplicate is the agent of a corresponding buttering.

What’s more, there is some evidence that thoughts in Y&H’s sense (assuming that there are such things) are not spatially located at all, a fortiori not spatially located in agents. They have a temporal location, because a neo-Davidsonian treatment of:

- (i) Last week, Juhani believed that snow is white,

requires that believing-eventualities can occur at times, say last week. But sentences like (g) do not happily take location modifiers:

- (j) *Over there/in his office Juhani believed that snow is white.³

And finally, at a superficial glance an eventuality corresponding to (g) is not Juhani’s belief that snow is white:

- (k) *Juhani’s belief that snow is white occurred/happened last week.⁴

2 Coarse narrowness versus Y&H narrowness

I have just argued, in effect, for framing the dispute in terms of the more traditional conception of “coarse narrowness”, where what is at issue is the intrinsic/extrinsic status of properties or conditions of agents, like *believing that snow is white*. Y&H

³ See Maienborn 2007, 2019.

⁴ For a deeper look, see Moltmann Forthcoming: 196-7.

offer some considerations against this traditional framing. First, they claim that it is too theoretically restrictive (49). In particular, it does not help us assess what Y&H call *ecumenical internalism*, the view that “the ur-content assignment is broad but...there is some other content assignment that is both narrow and occupies some interesting theoretical role” (97). This characterization of ecumenical internalism in terms of a content assignment to “thoughts” (eventualities of a certain kind) presupposes the questionable neo-Davidsonian apparatus, but the basic idea has no such presupposition. For instance, consider a view on which (i) believing that snow is white is extrinsic, and (ii) being in that state entails being in the intrinsic state of *believing that the snowish-stuff is white*, where the proposition that the snowish-stuff is white is the “internal” component of the broad content that snow is white. This is a kind of ecumenical internalism. We can assess this view without departing from the traditional conception, and indeed Y&H in effect provide an assessment themselves.

Y&H apparently see a problem here because they construe ‘coarse narrowness’ very narrowly (as it were): the theorist of coarse narrowness is “at best attempting to ask whether the ur-content assignment is narrow” (48). That is, once (i) is established, content is not coarsely narrow and the debate is over. But ‘coarse narrowness’ can be given a broader interpretation, so that if (ii) is true, there is a component or dimension of belief content that is coarsely narrow. Perhaps we sometimes need to invent special terminology (‘snowish-stuff’) to express this content, but we don’t need the apparatus of eventualities in order to state the view.

Y&H’s second objection is that the tools of coarse narrowness are “not even adequate to the task” of theorizing about ur-content (48), which they support with the following “toy example”:

...imagine two duplicate brains, Brain 1 and Brain 2, with exactly two beliefs each, one occurring in a red hemisphere, the other in a blue hemisphere. Suppose that the ur-content of Brain 1’s red hemisphere belief is p and that the ur-content of Brain 2’s red hemisphere belief is a distinct content q , whereas the ur-content of Brain 1’s blue hemisphere belief is q and the ur-content of Brain 2’s blue hemisphere belief is p . This is already a counterexample to the narrowness of ur-content in our sense, but it remains consistent with the claim that content is coarsely narrow. Suppose furthermore that the ur-contents of the red hemisphere and blue hemisphere beliefs of any possible duplicate of Brain 1 are determined by whether that duplicate is within a mile of a baboon: in a world in which a duplicate of Brain 1 is within a mile of a baboon, its blue hemisphere belief’s ur-content is p and its red hemisphere belief’s ur-content is q , and otherwise its blue hemisphere belief’s ur-content is q and its red hemisphere belief’s ur-content is p . In this example 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. Once the case is in view, it is clear that our conception of narrowness is preferable to coarse narrowness. (48–9)

If the example essentially relies on neo-Davidsonianism to underwrite the “occurrence” of a particular belief “in” the red hemisphere, then it isn’t very persuasive. But we can redescribe it without such a commitment. Suppose, then, that there is a language of thought (a hypothesis to which Y&H are somewhat partial; see, e.g., 20). Brain 1’s belief box is divided, with one half in the red hemisphere and the other half in the blue hemisphere. In the red belief box there is a token of Mentalese sentence-type *S*, and in the blue box there is a token of sentence *S**. For any duplicate of Brain 1, what the tokens of *S* and *S** mean depend on the proximity of a baboon: a close baboon makes (the tokens) *S* and *S** mean *q* and *p*, respectively; a distant baboon makes *S* mean *p* and *S** mean *q*.

Brain 1 is very far from a baboon: *S* means *p* and *S** means *q*. A duplicate, Brain 2, is right next to one: *S* means *q* and *S** means *p*. Brain 1 (or the animal whose brain it is) believes *p* and believes *q*; Brain 2 believes *q* and believes *p*. Brain 1 and Brain 2 have the same beliefs, and so the tools of coarse narrowness fail to certify the case as a kind of externalism. And yet, as Y&H, say, it clearly is. Does this show that coarse narrowness is not up to snuff?

I don’t think so. One reply is that although Brain 1 and Brain 2 don’t illustrate that the baboon theory of content determination is coarsely broad, on any version of the baboon theory worth discussing there will be another case which *does* illustrate that. Take, for example, the result of performing a hemispherectomy on Brain 1 and Brain 2: both brains lose the blue hemisphere. Y&H don’t say how the baboon influences content, but a natural prediction is that the operation results in Brain 1 believing *p* and Brain 2 believing *q*—duplicate brains with different beliefs, and so a case of coarse broadness.

In case this first reply does not cover all cases of theoretical interest, there is another. If content is narrow, then an explanation of Agent 1’s actions in terms of its propositional attitudes (e.g. beliefs) should also apply to the corresponding actions of a duplicate, Agent 2. Why did Agent 1 go the pub? Because, *inter alia*, it believed *p*. If the explanandum adverts to narrow content then the same explanation should go through for Agent 2, similarly situated with respect to a pub. But now imagine Agent 1 has (intact, two-hemisphere) Brain 1, and Agent 2 has Brain 2. At the level of neural mechanisms, Agent 1 goes the pub because the sentence *S* in its red hemisphere engages with its motor control system, propelling Agent 1 pubwards; *S** in its blue hemisphere is idle. But then Agent 2’s *S** is also idle: it did not go to the pub because it believed *p*. So without departing from the spirit of coarse narrowness, we can certify this case as broad.

One might fairly conjecture that similar strategies can cope with Y&H’s final programmatic worry, that coarse narrowness cannot cope with “the theoretical roles of interest in this book hav[ing] to do with explaining various properties of thoughts, including their epistemological roles, phenomenology, and so on” (49), although of course everything turns on the details.

3 Mirror man and thought relativism

Jettisoning the dubiously appropriate neo-Davidsonian apparatus in favor of old-fashioned coarse narrowness helps Y&H’s case. Let me build up to that with some more notation and terminology from *Narrow Content*.

A content or proposition p determines a function f_p from possible worlds to truth values, such that $f_p(w) = T$ if p is true at w , and $f_p(w) = F$ otherwise. f_p is the *intension* of p . If there are “alethic parameters” other than worlds—times, for example—the idea can be generalized: the intension of p is a function from *indices* $\langle w, t, \dots \rangle$ to truth values. The set of all indices \mathbf{I} is the Cartesian product of the sets containing the alethic parameters—one set for worlds, another set for times (if needed), and so on:

$$\mathbf{I} = \mathbf{X}_1 x \dots x \mathbf{X}_n$$

Y&H “assume, with semantic orthodoxy, that the set of all metaphysically possible worlds, which we will call ‘ \mathbf{W} ’, is the first alethic parameter” (26); second, that there is “a natural way of associating each thought α with a unique index i_α , the index of α ” (27). And:

Let us say that a thought α is *Doppelgänger*-related to a thought β iff QAP (α) = QAP (β) and [the truth value of $\alpha \neq$ the truth value of β]. That is, two thoughts are *Doppelgänger*-related to each other iff they are agential duplicates that differ in truth value. (68)

Suppose that truth is relative only to worlds: \mathbf{W} is the only alethic parameter. A Putnam-style Twin-Earth scenario shows that a pair of ‘water’-thoughts can be *Doppelgänger*-related, and that externalism is true for at least some thoughts. (Other sorts of examples, notably Burge’s, show that the conclusion can be extended to every thought (70–1).)

Increasing the alethic parameters can help the internalist. If the truth of a thought is also relative to the agent of the thought, then the fact that Earthling’s ‘water’-thought and duplicate-Earthling’s ‘water’-thought differ in truth value don’t show that they have the same content. Y&H’s character Loop Lady is brought into show that supplementing \mathbf{W} with an agent parameter \mathbf{A} isn’t sufficient (73–4). Shuttling from Earth to Twin Earth, Loop Lady’s ‘water’-thought on Earth and her ‘water’-thought on Twin Earth are agential duplicates but the world and agent of the two thoughts are identical. Hence, this is a case of agential duplicate thoughts which have different contents, because the thoughts differ in truth value relative to the same index.

Now a regress is supposed to begin, with the internalist adding parameters to fend off yet more externalist examples. Y&H’s Mirror Man, “an agent who is perfectly qualitatively symmetrical along some plane”, shows that adding a third parameter for time won’t help (76–7). After a few more iterations of this dialectic, Y&H consider the “radical view” that the “set of all possible thoughts is an alethic parameter”, which they call *thought relativism*:

According to thought relativism, the thought itself is included in its index: necessarily, for all thoughts α , $i_\alpha = \langle \mathbf{W}(\alpha), \mathbf{A}(\alpha), \mathbf{T}(\alpha), \dots, \alpha \rangle \dots$
 ...Thought relativism is a highly unorthodox form of relativism, but it has at least one notable virtue (from the internalist's point of view): it renders the internalist's position immune against all counterexamples of the form we have been discussing. (80)

As Chalmers points out, "There are worse problems for a view to have!" (Chalmers 2018). But the availability of thought relativism is an artefact of Y&H's dialectical framing, where the existence of thoughts occurring "in" agents is supposedly secured by neo-Davidsonianism. Without that, there is no evident reason to believe in internal "thoughts" in the first place. Mirror Man remains as a counterexample to internalism supplemented with alethic parameters for world, agent and time, but thought relativism isn't an available response.

4 I = W?

Y&H are sympathetic to the view that truth is only relative to worlds: **I = W**. As they note, "There is a case to be made that the world parameter is special" (66). In my view, Y&H's sympathies are well-placed.⁵ Reasons to think that "the world parameter is special" arise, I think, as soon as the 'world' terminology is introduced.

Recall Stalnaker's well-known exchange with Lewis:

The argument Lewis gives for thesis (1) [that possible worlds exist], identifying possible worlds with ways things might have been, seems even to be incompatible with his explanation of possible worlds as more things of the same kind as I and all my surroundings. If possible worlds are ways things might have been, then the actual world ought to be *the way things are* rather than *I and all my surroundings*. *The way things are* is a property or a state of the world, not the world itself... This is important, since if properties can exist uninstantiated, then *the way the world is* could exist even if a world that is that way did not. One could accept thesis (1)—that there really are many ways that things could have been—while denying that there exists anything else that is like the actual world. (Stalnaker 2003: 27–8)

Stalnaker's point is well-taken: the actual world, the way things are, is not "I and all my surroundings". But is Stalnaker's further claim correct, that a possible world is a (maximally specific) property that the world could have had? As someone once said, the world is all that is the case. *The way the world is*, or *the way things are*, seems best explicated in terms of facts or true propositions, not properties or states. *This is the way the world is*: Stalnaker is a philosopher, kangaroos have tails,...; *this is a way things might have been*: Stalnaker is a plumber, kangaroos have no tails,...

If that is right, on the unforced way of taking the introductory explanation of "possible worlds", it appeals both to (monadic) truth and propositions. A possible

⁵ See also Cappelen and Hawthorne 2009: ch. 1.

world is a maximal consistent proposition, and the actual world is the true one.⁶ Truth is not relative to worlds in any interesting sense: p is true in w iff $w \vdash p$ iff if w had obtained (been true), p would have been true (cf. 67).⁷

Y&H commendably try to give their opponents enough rope. But one should not lose sight of the strong case for **W** being the only alethic parameter. And if **I** = **W** a number of the internalist positions discussed in the book do not get off the ground.

5 Phenomenal intentionality

Finally, and all too briefly, phenomenal intentionality. As Y&H say:

Many of the philosophers who find the thesis of privileged access to ur-content compelling are motivated to think that the ur-content of a thought at a time must somehow be determined by the phenomenal life of the agent—or ‘how things seem’ to the agent—at that time’ (113)

That does seem, as Y&H say, “completely incredible” (114). But there is a weaker and much better motivated “phenomenal” internalist thesis that Y&H do not discuss. Consider this argument:

- P1. How things perceptually seem determines a content: *perceptual content*, the way the perceiver’s environment is presented as being.
- P2. How things perceptually seem is internally determined.⁸
- C. Perceptual content is internally determined.

Both premises are quite popular. Those who think that perceptual content can concern particular individuals (say, the individual rabbit Peter whom one is looking at) will deny P2, but they might also accept a layer of perceptual content that is existentially quantified (cf. Davies 1991). For a theorist who thinks that (one layer of) perceptual content is not object-involving, a revised version of the argument could well seem appealing. And even if perceptual content is broad because it is object-dependent, that might be the sole source, with the representation of perceptual qualities like colors, textures and shapes being internally determined

⁶ Cutting propositions coarsely by modal equivalence.

⁷ Perhaps Stalnaker’s invocation of properties should not be taken too seriously. At least, switching from properties to propositions preserves his insight that possible (and non-actual) worlds exist even though there are (quantifying over absolutely everything) no concrete universes that correspond to them. (On the property conception: world-states exist uninstantiated; on the propositional conception: world-propositions exist without being true.)

⁸ Imagine the following Burge-style case. Arthur and twin-Arthur are both looking at the same tomato. The reddish-yellow tomato is close to a borderline case of red. In Arthur’s linguistic community (ours), ‘red’ applies to the tomato. In twin-Arthur’s community, where the intension of ‘red’ is slightly different, it does not. The tomato looks red to Arthur; does the tomato also look red to twin-Arthur? (Of course, twin-Arthur is not in a position to put it that way.) If the tomato does not look red to twin-Arthur, then there is a sense in which things do *not* perceptually seem the same way to Arthur and his twin, and P2 (under that interpretation) is false. We need not settle the issue here because there is a natural understanding of P2 on which it is not refuted by Burge-style cases—and that is the one at issue.

(see, e.g., Pautz Forthcoming). Whether internalism should be junked entirely depends on whether perceptual phenomenology is determined by inner neural goings-on. *Narrow Content* leaves that last question hanging.⁹

References

- Bach, E. (1986). The algebra of events. *Linguistics and Philosophy*, 9, 5–16.
- Cappelen, H., & Hawthorne, J. (2009). *Relativism and monadic truth*. Oxford: Oxford University Press.
- Chalmers, D. (2018). Review of Yli-Vakkuri and Hawthorne, *Narrow Content*. *Notre Dame Philosophical Reviews*. <https://ndpr.nd.edu/news/narrow-content/>.
- Davidson, D. (1967). The logical form of action sentences. In N. Rescher (Ed.), *The logic of decision and action*. Pittsburgh: University of Pittsburgh Press.
- Davies, M. (1991). Individualism and perceptual content. *Mind*, 100, 461–484.
- Maienborn, C. (2007). On Davidsonian and Kimian states. In I. Comorovski & K. V. Heusinger (Eds.), *Existence: Semantics and syntax*. Dordrecht: Springer.
- Maienborn, C. (2019). Events and states. In R. Truswell (Ed.), *The Oxford handbook of event structure*. Oxford: Oxford University Press.
- Moltmann, F. (Forthcoming). Attitudinal objects. In Tillman, C. (Ed.) *The Routledge Handbook of Propositions*. London: Routledge.
- Pautz, A. (Forthcoming). Naïve realism and the science of consciousness. *Analytic Philosophy*.
- Stalnaker, R. (1988). Belief attribution and context. In R. H. Grimm & D. D. Merrill (Eds.), *Contents of thought*. Tucson, AZ: University of Arizona Press. (Page reference to the reprint in Stalnaker 1999).
- Stalnaker, R. (1999). *Context and content: Essays on intentionality in speech and thought*. Oxford: Oxford University Press.
- Stalnaker, R. (2003). *Possible worlds. Ways a world might be*. Oxford: Oxford University Press.
- Yli-Vakkuri, J., & Hawthorne, J. (2018). *Narrow content*. Oxford: Oxford University Press.

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

⁹ Many thanks to Justin Khoo and Bob Stalnaker.