

# Hyperintensionality and Overfitting

Francesco Berto\*

## Abstract

A hyperintensional epistemic logic would take the contents which can be known or believed as more fine-grained than sets of possible worlds. I consider one objection to the idea: Williamson's Objection from Overfitting. I propose a hyperintensional account of propositions as sets of worlds enriched with topics: what those propositions, and so the attitudes having them as contents, are about. I show that the account captures the conditions under which sentences express the same content; that it can be pervasively applied in formal and mainstream epistemology; and that it is left unscathed by the objection.

## 1 The Hyperintensional Revolution

Believers in propositions think that different sentences (or, sentence types) can sometimes say the same thing, or express the same content. We want a semantics to capture propositional content as what is said; and so, to capture same-saying. Call *Standard Possible Worlds Semantics* (SPWS) the view that such contents are sets of possible worlds giving truth conditions; or, equivalently, intensions: functions from possible worlds to extensions – in the relevant case, truth values. Call an operator  $H$  *hyperintensional* when  $H\varphi$  and  $H\psi$  can differ in truth value although  $\varphi$  and  $\psi$  express necessarily equivalent contents; a propositional content  $P$ , when  $P$  can differ from  $Q$  although they are necessarily equivalent. Necessary equivalence is usually understood as co-intensionality: truth at the same worlds. A hyperintensional semantics would account for some hyperintensional operators by postulating hyperintensional contents, i.e., contents that are individuated more finely than sets of possible worlds.

---

\*Arché, University of St Andrews, UK, [fb96@st-andrews.ac.uk](mailto:fb96@st-andrews.ac.uk) and Institute for Logic, Language and Computation (ILLC), University of Amsterdam, The Netherlands [f.berto@uva.nl](mailto:f.berto@uva.nl).

As propositions are often taken also as the objects of various attitudes, modal-epistemic logic in the Hintikka (1962) tradition represented knowledge and belief as modals – restricted quantifiers over possible worlds – and the things which can be known or believed as SPWS propositions. This was part of the 20th Century’s ‘intensional revolution’: a collective effort to account for a number of concepts (essence, causation, supervenience, conditionality, information) in terms of intensions. Troubles have emerged piecemeal, but have a common source: those notions appear to be hyperintensional. Nowadays there are lots of hyperintensional approaches to content, e.g., structured propositions of various sorts (Soames, 1985, 2010; King, 1996, 2007; King et al., 2014; Duží et al., 2010, 2023), non-normal or impossible worlds semantics (Jago, 2014; Priest, 2016; Berto and Jago, 2019), situation semantics (Barwise and Perry, 1983), truthmaker semantics (Fine, 2017; Fine and Jago, 2018), non-classical logics (Anderson and Belnap, 1975; Anderson et al., 1992; Dunn and Restall, 2002; Standefer, 2023). They differ vastly, but have this much in common: they distinguish contents identified in SPWS. Some talk of a ‘hyperintensional revolution’ (Nolan, 2014).

One early issue in epistemic logic was that the Hintikkan agents are ‘logically omniscient’: they know or believe all logical consequences of what they know or believe. One  $X$ s that  $\varphi$  ( $X$  being an ascription of knowledge, belief, perhaps some other attitude) at world  $w$  when the SPWS proposition that  $\varphi$  is true at all worlds epistemically accessible from  $w$  (compatible with one’s evidence at  $w$ , one’s belief system there, or whatnot). If entailment is truth preservation at all worlds of all models, one who  $X$ s that  $\varphi$  automatically  $X$ s all the entailed  $\psi$ s. Surely we aren’t like that. Also, all necessary truths turn out to be uninformative. That there are no solutions in positive integers for  $x^n + y^n = z^n$  when  $n > 2$  would be already known or believed. But it took a proof of over 130 pages to find out that it’s true.

That  $2 + 2 = 4$ , and that there are no solutions in positive integers for  $x^n + y^n = z^n$  when  $n > 2$ , are true at the same worlds and so, by SPWS, are the same proposition. Perhaps that’s where the problem lies? One may be tempted to adopt some hyperintensional view of content and do epistemic logic systematically on that basis. (E.g., for works recruiting Kit Fine’s truthmaker semantics to do just that, see Krämer (2022); Hawke and Özgün (2023).) Under the label of ‘overfitting’, section 2, introduces Timothy Williamson’s recent arguments against the very idea of being so tempted. Section 3 talks of what I take to be the core issue with SPWS: it offends what Steve Yablo has called ‘our sense of when sentences say the same thing’ (Yablo, 2014, 2). The overfitting strategy, it is argued, won’t easily explain this away. Section 4 briefly considers to what extent intensionalists may resort to guises to account for such a sense. Section 5 accounts for it

by borrowing from Berto (2022) a little hyperintensional theory, developed by Peter Hawke and myself, where propositions are taken as made of two things: their truth set (the set of worlds where they are true) and their *topic* or *subject matter* (what they are about) – a version of what is coming to be called ‘two-component semantics’ (Yablo, 2014, 2017, 2018; Hawke, 2016; Plebani and Spolaore, 2023; Ferguson, 2023b,c; Berto and Hornischer, 2023; Hawke et al., 2024). Section 6 sketches how the theory can be put to work in formal and mainstream epistemology. Section 7 concludes.

## 2 Overfitting

The intensional revolution was resisted by one of the world’s greatest philosophers: Quine. The hyperintensional revolution is resisted by one of the world’s greatest philosophers: Timothy Williamson. According to him, hyperintensionalists are guilty of *overfitting*, ‘the willingness to add extra parameters to an equation until its curve goes almost exactly through all the data points’ (Williamson, 2020, 264). In science, overfitting is a bad feature of data models:

The flexibility of a model can be roughly measured by the number of its degrees of freedom, of adjustable parameters in the model. [...] By adding more and more degrees of freedom, one can fit just about any data, but in a cheap way which typically brings no insight. The problem with having too many degrees of freedom is not just un informativeness. It is also insensitivity to errors in the data, since the model can accommodate any data, however anomalous. [...] A better methodology is to be very reluctant to add new degrees of freedom, doing it only after potential sources of error in the data have been investigated and shown not to explain the evidence just as well. (Williamson, 2021, 79-80)

Now hyperintensionalists, Williamson has it, base their accounts on intuitive data from language use, speakers’ patterns of acceptance and rejection of sentences, which seem to tell against a merely intensional individuation of content: ‘a key feature of the hyperintensional revolution is that it is driven by *examples*, especially by apparent counterexamples to intensional principles.’ (Williamson, 2021, 87) Distinctions are postulated at the level of semantics in an effort to accommodate such data. But the data, Williamson claims, can be shown to be spurious precisely in the cases which were supposed to motivate the need for hyperintensional fine-graining.

Let us see how the objection is unpacked. (Williamson, 2020), a book on indicative and counterfactual conditionals, defends the view that the semantics of the indicative ‘if’ is given by the material conditional. The counterfactual is then accounted for by combining the material conditional with a ‘would’ operator, taken as a normal or merely intensional modal. There’s a wealth of data providing apparent evidence against the material conditional analysis: competent speakers just don’t seem to use the indicative ‘if’ as we would expect if its content was captured by the material conditional (Edgington, 1995). In the past, materialists tried to deal with the recalcitrant data at the level of pragmatics (Grice, 1989). But Williamson has a new strategy: borrowing terminology from cognitive psychology (e.g. Tversky and Kahneman, 1974), he claims the data are generated by fallible cognitive or epistemic heuristics in place in our assessment of conditionals: “‘fast and frugal’ or ‘quick and dirty’ ways [of] answering questions which are reliable enough to be useful, but still not perfectly reliable’ (Williamson, 2021, 83).

Should one ‘wonder where to fit such heuristics for assessing sentences into a standard picture of linguistic architecture, with semantics built on syntax and pragmatics built on semantics’ (Williamson, 2020, 24), he answers that ‘the heuristics must come above the semantics, for normally one is in no position to decide between accepting and rejecting a sentence until one knows what it means.’ (Ibid). However, it comes below the pragmatics:

*Pragmatics* is the usual first resort for filling the gaps between semantics and language use. However, it is not what is wanted here. In analysing conversational phenomena, pragmatics (legitimately) takes for granted speakers’ capacities to make the very cognitive assessments we are now seeking to understand. At the basic cognitive level, what we seek is a matter of psychology rather than linguistics. (Williamson, 2020, 5)

Williamson labels the heuristic governing our primary way of assessing conditionals the Suppositional Rule. It prescribes to take an attitude unconditionally to ‘If  $\varphi$ , then  $\psi$ ’ iff one takes that same attitude to the consequent  $\psi$ , conditional on the supposition of the antecedent  $\varphi$  (Williamson, 2020, 19). E.g., *Accept ‘If  $\varphi$ , then  $\psi$ ’ iff you accept  $\psi$  on the supposition of  $\varphi$ .* Next, in chapter 3 of his book Williamson goes on to show that the Suppositional Rule is *inconsistent*: when applied to attitudes to logical consequences of hypotheses, it leads to contradictions. When applied to attitudes that admit of degrees, such as credences, it delivers probabilistic paradoxes.<sup>1</sup> The moral is not that we don’t really use the Rule to assess conditionals. Rather,

---

<sup>1</sup>One can check in particular sections 3.2-3.4 for the details of Williamson’s arguments.

‘we may have been using an inconsistent rule for “if” all along’ (Williamson, 2020, 41).

Next, the account of the counterfactual presented in Part II of the book agrees with the standard similarity-based SPWS of Stalnaker (1968) and Lewis (1973) in making all *counterpossibles* – counterfactuals with impossible antecedents – trivially true. Various hyperintensional accounts of counterfactuals (Kocurek, 2021, has a beautiful overview) take issue with this, starting from the intuitive data that we are not disposed to accept both of each pair of counterpossibles with the same antecedent but opposite consequents. For a famous example due to Daniel Nolan (1997), take ‘If Hobbes had (secretly) squared the circle, sick children in the mountains of South America at the time would have cared’ and ‘If Hobbes had (secretly) squared the circle, sick children in the mountains of South America at the time would not have cared’: we’ll reject the former and accept the latter. When we counterfactually suppose, *per absurdum*, that Hobbes managed to come up with a square equal in area to a given circle using only ruler and compass, we conclude that sick children of South America could not have cared less in the hypothetical scenario. But in chapter 11 of his book, Williamson argues that the suppositional heuristic for the counterfactual is inconsistent just as well (section 11.4). In chapter 12, he argues that putative counterexamples to the intensionality of counterfactuals proposed by Fine (2012) don’t work, and concludes that no evidence has been provided that counterfactuals are hyperintensional after all.

In (Williamson, 2021), a similar story is told concerning attitudes. One way in which accounts more fine-grained than SPWS have been motivated, at least since Carnap (1947)’s idea of an intensional isomorphism, has to do with contexts created by attitude ascriptions: we seem to sometimes truthfully say that one has an attitude towards the proposition that  $\varphi$  without having it towards an intensionally equivalent  $\psi$ : John Doe may believe that  $2 + 2 = 4$  without believing Fermat’s Last Theorem, etc. But, Williamson claims, Kripke (1979)’s Pierre puzzle should have alerted us to the possibility that our attitude ascriptions are also guided by heuristics which turn out to be occasionally inconsistent:

[Kripke] plausibly suggests that English speakers rely on some-

---

The results are not so surprising when one considers that the (acceptance instance of the) Suppositional Rule is in fact the so-called Ramsey Test for the assessment of indicatives (Ramsey, 1990). Peter Gärdenfors had already shown in the 80s that the Test is inconsistent with natural assumptions on belief revision (Gärdenfors, 1986). The probabilistic paradoxes are akin to the triviality results, due to Lewis (1976), Hajek (1989), and others, which hit the hypothesis that the probability of an indicative amounts to the probability of the consequent conditional on the antecedent.

thing like the schema “*A normal English speaker who is not reticent will be disposed to sincere reflective assent to ‘p’ if and only if he believes that p*”. Plausibly, users of other natural languages rely on analogous schemata. Call this family of schemata *the assent principle*. Combined with the convincing principle that correctly translating a belief-ascription preserves its truth-value, the assent principle generates inconsistency in describing the beliefs of a bilingual speaker in realistically possible circumstances (the famous case of puzzling Pierre). Indeed, as Kripke also explains, the problem arises even in the monolingual case.

Readers of ‘A Puzzle about Belief’ may interpret Kripke as suggesting that our ordinary concept of belief is incoherent, or something like that. In response, some may attempt to qualify the assent principle in more or less elaborate ways to avoid the contradiction. But there is a simpler possibility. The assent schema may be a normal heuristic for ascribing beliefs. [...] All the complications may be fall-backs we invoke when the basic heuristic fails. If this approach is correct, many of the apparent counterexamples to various theoretical claims presented in the voluminous and inconclusive literature on propositional attitude ascriptions may be errors generated by our implicit reliance on fallible heuristics. The many complicated accounts proposed for the semantics of propositional attitude ascriptions may just be artefacts of overfitting. (Williamson, 2021, 84)

The take-home message: we should be wary of ‘the danger of giving semantic solutions to epistemic problems’ (Williamson, 2020, vi), or we’ll be dismissing good theory (in particular, SPWS) on the basis of bad data: ‘the self-proclaimed hyperintensional revolution involves multiplying degrees of freedom in order to explain data which may well be unreliable. That looks like a classic case of overfitting.’ (Williamson, 2021, 93)

### 3 Saying the Same Thing

The Williamsonian stance is (consciously) at odds with contemporary research in linguistic semantics. A standard textbook (e.g., Chierchia and McConnell-Ginet, 1990) will tell you at the outset that a key task of semantics is to capture competent speakers’ intuitions of synonymy, antonymy, entailment, equivalence, presupposition, etc. The theory is to start from ordinary language use, patterns of assent and dissent, shared judgments on

what means what. (Where else could one start from? Self-evident axioms? A priori epiphanies?) It will end up generating analyses and predictions about (further) use, assent and dissent, etc., which may be corroborated or disconfirmed, in a feedback loop, perhaps on the way to some reflective equilibrium. Everyone agrees that no theory can take on board all our intuitions. For these are bound to be inconsistent; but in the unfortunate case that diletheism is wrong, no acceptable semantics can be inconsistent. However, as remarked by Rothschild (2021) in his *Mind* review of Williamson’s book, Williamson takes the intuitive judgments delivered by our cognitive procedures, not as a guide to the semantics, but as things to be accounted for on their own, possibly in clear disagreement with the endorsed view of content. Conversely, the latter is not to account for patterns of sentence use, if not in the roundabout way mediated by the cognitive procedures.

Two can play the game! Call *underfitting* the over-simplification of models neglecting good data. One should be wary of retaining bad theory in spite of good data. What are the good data? I think they have to do, in the memorable Yablovian words I quoted above, with ‘our sense of when sentences say the same thing’ (Yablo, 2014, 2). Sure, it’s our sense: the distinctions supposedly missed by SPWS will be motivated by our fallible judgments – again, where else is one to start from? But some data on (non-)same-saying may not be easily explained away as by-products of fallible heuristics. They may be stable across sentences of different kinds and modal profiles, systematic, and such that accounting for them at the level of content gives explanatory virtues, unmatched by alternative accounts relegating them outside of the semantics.

Different sentences seem to sometimes say different things in spite of being necessary (of the same kind of necessity), or co-necessary, thus co-intensional:

1.  $2 + 2 = 4$ .
2.  $3 + 3 = 6$ .
3. Equilateral triangles are equiangular.

These are mathematical necessities. Only one is about equilateral triangles, and made true by what they are like.<sup>2</sup>

4. Fido is a dog.

---

<sup>2</sup>An anonymous referee of this paper wonders whether abstract types like numbers or geometrical figures may even be in the domain of any possible world, and takes this to be a further issue for the SPWS capacity to tell apart contents of this kind.

5. Kitty is a cat.

6. Water is H<sub>2</sub>O.

For a number of essentialists, these are metaphysical necessities. Only one is about water.

7. If it snows, then it snows.

8. Either South Bend is in Indiana, or not.

9. The Liar sentence is not both true and false.

These are (classical) logical necessities (well, with some leeway with the truth predicate). Only one is about the Liar.

10. Clopen sets are sets which are both open and closed.

11. Non-normal modal logics are modal logics weaker than K.

These are unrestrictedly necessarily true, if definitions of this kind are. Only one is about clopen sets.

12. Grass is green.

13. That grass is green is true,

These are arguably co-intensional, true at the same worlds. Only one is about a proposition. It is *prima facie* an issue for SPWS that it conflates such intuitively distinct contents. It has been raised several times in the literature. It is, for instance, part of Scott Soames (1985, 1987)'s influential criticisms of the idea that propositional contents are adequately captured as sets of possible worlds.<sup>3</sup>

Can't the desired distinctions be made within a merely intensional setting? That equilateral triangles are equiangular and that  $2 + 2 = 4$  can be told apart, an intensionalist may say, by using extensions and/or mere intensions: those of the subsentential constituents of the relevant sentences. Say,

---

<sup>3</sup>Soames targets accounts of contents as (mere) sets of truth-supporting circumstances, even more fine-grained than possible worlds. If the criticisms work in general (which is controversial: see Ripley (2012)), they will take down some hyperintensional accounts as well. This is of less importance for us; first, because the focus here is on defending hyperintensionality in general; second, because the toy hyperintensional account proposed below doesn't take propositions merely as sets of truth-supporting circumstances anyway.



‘equiangular’ gets a function from possible worlds to extensions, somehow supposedly embedded in the former proposition, but not in the latter.

But this is a hyperintensional, not merely intensional view of content, probably in the vicinity of structured propositions. That equilateral triangles are equiangular and that  $2 + 2 = 4$  keep being true at the same worlds, but now they are distinct contents for they feature distinct constituents. The constituents may be individuated merely intensionally, but the view is incompatible with the claim that contents just are sets of worlds.

How would the heuristics strategy come to the rescue? Williamson (2021) discusses examples used to motivate hyperintensionality in metaphysics, e.g., by proponents of grounding theories. The examples involve operators like ‘because’, ‘it is essential to  $x$  that’, or ‘ $x$  brings it about that’, such that substitution of necessary equivalents in their scope appears not to be truth-preserving:

14. It is essential to Socrates that he is Socrates.
15. It is essential to Socrates that he is a member of {Socrates}.
  
16. The proposition that grass is green is true because grass is green.
17. Grass is green because the proposition that grass is green is true.
  
18. Mary brought it about that John was a contributor.
19. Mary brought it about that John was a self-identical contributor.

We’re inclined to accept the first item in each pair and reject the second, although the embedded sentences express co-intensional contents. We are guided by considerations of grounding and explanatory asymmetry, e.g., Socrates’ existence and identity grounds the existence and identity of its singleton, but not vice versa; truth is grounded in facts and not vice versa.

But we have the same pattern of inclinations towards these:

20. Vera is a vixen because Vera is a female fox.
21. Vera is a female fox because Vera is a vixen.
  
22. Richard brought it about that Edward was a king.
23. Richard brought it about that Edward was a male monarch.

However, these differ pairwise only by substitution of synonymous subsentential constituents:<sup>4</sup> they cannot express different propositions. Thus, there is something wrong with our inclinations. This might be taken as casting a shadow on cases (14)-(19) too: it may well be that ‘superficial linguistic features can easily deceive us into accepting unsound arguments for hyperintensionality.’ (Williamson, 2021, 90) The diagnosis: in explanatory reasoning, the psychologically salient direction of explanation is from the simpler to the more complex, or from the more to the less familiar: ‘The helpfulness of a (putative) explanation is sensitive to its superficial linguistic form. For explanations are meant to provide *understanding*; how far they do so depends partly on their superficial linguistic features.’ (91) Thus, e.g., ‘The proposition that grass is green’ is a longer sentence than ‘Grass is green’, and grasping its meaning involves having some idea of what a proposition is, hence (16) looks better than (17). Sure, ‘female fox’ is two words while ‘vixen’ is only one, but the former two words are more familiar than the latter to the average English speaker, hence (20) looks better than (21).

But the hyperintensional stance is an existential claim: some distinct propositional contents,  $P$  and  $Q$ , are true at the same possible worlds. So for the strategy to work in general against hyperintensionality, it has to be universal. How would it extend to all of our initial list of cases?

The relevant intuitions connect the contents of sentences to what they are about. Aboutness is ‘the relation that meaningful items bear to whatever it is that they are *on* or *of* or that they *address* or *concern*’ (Yablo, 2014, 1). What meaningful items are about is their *subject matter* or, as I will also say, their *topic*. Research on topics has been burgeoning in recent decades (Lewis, 1988a,b; Gemes, 1994, 1997; Humberstone, 2008; Fine, 2016, 2017; Hawke, 2018; Moltmann, 2018; Schipper, 2018, 2020; Plebani, 2020; Plebani and Spolaore, 2021). What are topics? Some link them to questions sentences can be taken, in context, as answering to. Lewis (1988b)’s seminal example was *the number of stars*. It maps to the question, ‘What’s the number of stars?’. That splits modal space: two worlds end up in the same cell when they agree on the answer. The splitting gives what ‘There are ten stars’ can be about. Topics are partitions or divisions of modal space. On the other hand, any old sort of thing can serve as a topic: ‘Our topic in this lecture is neural networks’; ‘Today we talk about Margaret Thatcher.’ So other approaches to topic are more object-oriented (Goodman, 1961; Perry, 1989), or state-of-affairs-oriented. A prominent one by Fine (2017, 2020)

---

<sup>4</sup>An anonymous referee of this paper mentions that they think a male monarch needn’t be a king. So they’d object to the Williamsonian view that in (22) and (23) one is substituting synonyms.

takes topics as given by states or (exact) truthmakers/falsemakers: ‘There are ten stars’ can be about some situation which makes it true and which, unlike an entire possible world, is wholly relevant and responsible for its truth. Whatever the favored view of topics, there is nothing internalistic about aboutness.

Appeals to our sensitivity to superficial linguistic structure, syntactic simplicity or complexity, or familiarity with some terms rather than others, can’t plausibly work for all of the above cases. ‘ $2 + 2 = 4$ ’ and ‘ $3 + 3 = 6$ ’ have the same simple syntactic structure; so do ‘Fido is a dog’ and ‘Kitty is a cat’. Neither includes words less familiar than the other: ‘vixen’ may sound less familiar than ‘female fox’ to the average English speaker, but it would be bizarre to claim that people are generally more familiar with the numeral ‘2’ than with the numeral ‘3’, or with ‘dog’ than with ‘cat’.

Nor do the cases above obviously involve considerations of explanatory perspicuity. They have been presented unembedded, not in contexts involving operators tied to explanatory reasoning. It’s not even directly about patterns of assent and dissent, or rational requirements constraining them. Qua competent speakers of English (in possession of the relevant, truthful information) we may accept all of the contents expressed by sentences (1)-(13). We may also be rationally committed to accepting them all – because they are true, necessarily true, logically entailed by other truths, necessarily equivalent to them and to each other, or whatnot. Still we will take them as expressing distinct contents, for they are about different things.

One may bring in an operator ‘says the same as’ and come up with widespread, but mutually inconsistent folk judgments on sentences of the form ‘That  $\varphi$  says the same as that  $\psi$ ’. One can come up with inconsistent intuitions involving more or less any operator. But how are we to overrule our judgments of non-same-saying in all of the above cases? If it were to be applied to them, Williamsonian talk of ‘fast and frugal heuristics generating illusions’, taken from cognitive psychology, may start to look misleading. The Kahneman-Tversky illusions, such as the conjunction fallacy (people are prone to judge it more likely that Linda is a bank teller and active in the feminist movement than that Linda is a bank teller: Tversky and Kahneman (1974)) are normally overcome by the experimental subjects, after they have understood the relevant explanation. It’s not that easy to concoct an explanation that will make competent speakers accept that, when it is said that  $2 + 2 = 4$  and then that equilateral triangles are equiangular, the same thing has been said twice, setting aside superficial linguistic structure or explanatory salience.

## 4 Guises and Disguises

So one may raise an Objection from Underfitting against intensionalists who stick to SPWS propositions in the face of certain resilient intuitions. Intensionalists will probably agree that an explanation is called for, but they'll be happy to locate the explanatory material outside of semantic content.

One way may be to resort to guises or modes of presentation, taken as constraints on mental representations at work in pragmatics, or perhaps at the cognitive level intermediate between semantics and pragmatics Williamson has called our attention to. Guises may be legitimate devices for a number of purposes. They may be useful to hyperintensionalists as well: (neo-)Russellians on propositions like Salmon (1986) and Soames (1987) use them although Russellian structuralism makes for a hyperintensional individuation of content, way more fine-grained than SPWS. But can guises accommodate all good intuitions of aboutness and same-saying? The historical debate on guises quickly became subtle, but here's a reconstruction in broad strokes.

Guises were not supposed to be constituents of content. While using them in the business of addressing Frege's puzzle, Salmon (1986) distinguished the information 'semantically encoded' in a sentence from the 'pragmatically imparted' one. People were (and, are) impressed by arguments for the direct reference theory of names and by Kripke's criticisms of descriptivism, thus reluctant to add components to the semantics of names besides their denotation – especially components that would look like Fregean descriptive senses. Guises were introduced to be activated in opaque intentional contexts. The issue was that, even on a hyperintensional Russellian account of propositions, that Hesperus is Hesperus is the same proposition as that Hesperus is Phosphorus, whereas *prima facie* it appears that we can truthfully say that the ancients believed the former, not the latter.

But intuitions of aboutness and same-saying are shaky here, in a way they are not for at least some of our examples (1)-(13) above. Perhaps 'Hesperus is Hesperus' and 'Hesperus is Phosphorus' *are* about the same thing, namely planet Venus, or whatever single topic is suitably associated to it. It's difficult to argue that this is true of ' $2 + 2 = 4$ ', ' $3 + 3 = 6$ ' and 'Equilateral triangles are equiangular'. Does the difference in what these say reduce to our ways of mentally representing the same things? If so, which things?

Even when guises are used only to explain why we accept some attitude ascriptions and reject others, it may be taken as a mandatory feature of guise theory that it account for some compositional phenomena, in particular involving embeddings. But this easily makes guises look like Fregean senses in disguise. A critique along these lines can be found in a famous review of Salmon's book by Forbes (1987). An objection to Forbes can be found in

Branquinho (1990): Forbes' Fregean theory and Salmon's Russellian theory disagree in their assignment of truth values to attitude ascriptions. However, Graham Oppy (1992) argued, I think successfully, that the structural problem remains:

[I]t is as obvious that there must be a compositional theory involving Salmonian guises which issues in an assignment of assertability-values to sentence-context pairs as it is that there must be a compositional meaning theory for languages which issues in an assignment of truth-values to sentence-context pairs. (How else could we account for the fact that speakers can recognise the assertability-values of novel sentences? How else could speakers have the ability to produce and understand a potentially infinite range of sentences with attached assertability-values?) Moreover, it is equally clear that this theory will have exactly the same structure as the neo-Fregean theory of Fregean propositions. That is, it is clear that what Salmon's theory does is to shift some of the structure which is found in the Fregean theory from semantics to pragmatics. (Oppy, 1992, 4)

Now compositionality may be insufficient for falling on the side of content: there may well be compositionality at the epistemic and/or pragmatic levels. Still, I suspect something similar to what happened in the debate on guises may happen when someone attempts to protect SPWS by systematically explaining away all the putative counterexamples as mere cognitive differences alien to content. To be credible, the explanation will have to start looking like a disguised theory of content: as semantics, under another name.

Williamson (2020)'s account of the indicative as material conditional plus heuristics is not quite there. Endorsing the simplest semantics for the indicative 'if', he offloads the explanatory work semanticists expect from a semantics to the epistemic heuristics. The latter can only look simple to the extent that the account of the relevant epistemic procedures is underdeveloped.<sup>5</sup> I

---

<sup>5</sup>'The distance between usage and semantics in Williamson's account is a sign, not of developed science, but rather of an underdeveloped theory. [...] What Williamson presents as a good prediction of his theory, as opposed to a bad prediction of the material conditional view, is mostly a demonstration of the wide range of data that can be accommodated by exploiting the underspecification of heuristics. This problem is compounded by the fact that Williamson also takes instances in which conditionals seem to act like material conditionals to vindicate his overall account. Given the freedom Williamson's account allows in explaining linguistic judgements, it is no surprise that the survey of empirical problems for the material conditional view in chapters 7 and 8 is triumphalist.' (Rothschild, 2013, 11-12)

conjecture that, once fully developed, it may look at least as complicated as the rival, more complex semantics for the conditional he criticizes, and structurally like something many semanticists would call a semantics, but for the epistemic relabeling.

In the next section, I'll introduce an overfitting-free, hyperintensional account of content (a distillate of ideas from Berto (2022)) for a simple propositional language, capturing our robust judgments of (non-)same-saying. In the section after that I'll sketch how it can be pervasively put to use in epistemic logic.

## 5 A Little Hyperintensionality...

One who says 'Midori is an accountant' addresses a certain topic: one talks about Midori's job, what Midori does, or just Midori. And one says that things are such-and-so with respect to that topic. What one says is true just in case Midori's job is or includes being an accountant. We may then understand a proposition  $P$  as a pair,  $P = \langle W_P, T_P \rangle$ , and so as made of two components: (1)  $W_P$  is the truth set of  $P$ : the set of Worlds where it's true. (2)  $T_P$  is the Topic of  $P$ : what it's about, or directed to.  $W_P$  is just our old SPWS proposition ('thin proposition', as Yablo (2014) has it). The whole  $P$  also features a topic (a 'thick' proposition, or 'directed' if one likes: a proposition that points at a subject matter).

The literature on subject matters generally agrees on the space of topics having a natural mereological structure (Yablo, 2014; Humberstone, 2008; Yablo, 2014; Fine, 2016). Topics can have proper parts; distinct topics may have common parts. *Mathematics* includes *arithmetic*. *Mathematics* and *philosophy* share subject matter, having (certain parts of) *logic* in common. Correspondingly, what a proposition is about can overlap with, or be properly included in, what another one is about.

This also gives natural ideas of same-saying, saying more, etc. We claim that  $\varphi$  says at least as much as  $\psi$  (what  $\psi$  says is part of what  $\varphi$  says), when (1)  $\varphi$  entails  $\psi$  and (2) what  $\varphi$  is about includes what  $\psi$  is about. (As per Yablo's motto: 'Content-inclusion is implication plus subject matter inclusion', (Yablo, 2014, 15).) We claim that  $\varphi$  and  $\psi$  say the same when they are (1) co-intensional and (2) about the same things.

Next, the truth-functional logical vocabulary should be topic-transparent: it must add no subject matter of its own. There's some, though not universal, agreement on this: Yablo (2014) makes a forceful case for negation; Hawke (2018) and Fine (2020) make forceful cases for all the truth-functional connectives. The topic of  $\neg\varphi$  should be the same as that of  $\varphi$ : 'Grass is not

green’ is exactly about what ‘Grass is green’ is about. It certainly doesn’t address the topic of *negation*. Conjunction and disjunction merge topics: ‘Carlos is short and handsome’, ‘Carlos is short or handsome’ are about the height and looks of Carlos. The topics of  $\varphi \wedge \psi$  is that of  $\varphi \vee \psi$ : the fusion of the topic of  $\varphi$  and that of  $\psi$ .<sup>6</sup>

Here’s a simple hyperintensional semantics for a plain sentential language, using such ideas to capture same-saying. It is taken from ch. 2 of Berto

---

<sup>6</sup>On request by an anonymous referee, in this long footnote I’ll rehearse from the literature some arguments for the claim that the truth-functional connectives are topic-transparent (see, e.g. the aforementioned Yablo (2014); Hawke (2016); Fine (2020), but also Plebani and Spolaore (2021, 2023), Hawke et al. (2024)).

Well, why would one think otherwise, to begin with? The idea is inspired by the *Tractatus*’ ‘fundamental insight’, namely that ‘the “logical constants” do not represent’ (Wittgenstein, 1921/22, 4.0312). (Remember 5.44: ‘if there was an object called “ $\sim$ ”, then “ $\sim \sim p$ ” would have to say something other than “ $p$ ”. For the one proposition would then treat of  $\sim$ , the other would not.’). And one may worry that, by following Wittgenstein, one is pushing talk purporting to be about the meaning of the logical vocabulary in the realm of pseudo-talk aiming at the ineffable.

But we needn’t follow Wittgenstein that far. Surely we can talk about negation – it’s only that we don’t do it by negating sentences. Just in the same sense in which ‘Grass is green’ and ‘Grass is not green’ can be about *grass, what grass is like, or features of grass, etc.*, so can ‘Negation is a connective’ and ‘Negation is not a connective’ be about *negation, what negation is like, or features of negation, etc.* But ‘Grass is not green’ can’t be about *that* just because ‘not’ shows up in it.

One who thinks otherwise is probably taking what a sentence can be about – what topics it can be concerned with, what issues it can address, what questions it can be taken as an answer to in a conversational context – as plainly fixed by the contents (perhaps the reference or denotation) of its subsentential constituents; which, in general, it isn’t (‘Midori loves Akiko’ and ‘Akiko loves Midori’ have constituents with the same meanings – Midori, Akiko, loving – but can be about different topics). It may well be that negation is an operator from truth values to truth values, and the ‘not’ in ‘Grass is not green’ means or denotes that, but of course that’s not what ‘Grass is not green’ is about. ‘Negation is an operator from truth values to truth values’: this can be about negation. And in ‘Negation is not an operator from truth values to truth values’, it’s ‘negation’, not ‘not’, that allows the sentence to be about negation. ‘Negation is an operator from truth values to truth values’ is about that same topic: it says the opposite, about *it*.

Besides, there’s no easy candidate for a topic that’s systematically introduced in discourse by a negation, a conjunction, or a disjunction. For there to be such a thing, one would need discourse contexts where, e.g., ‘Grass is green’ is on-topic, but ‘Grass is not green’ isn’t, or vice versa, because the topicality component added by ‘not’ makes a difference – and it’s hard to come up with such contexts. Surely at most one will be true in any context (unless dialetheism); but that is handled by truth conditions, not by topics. Similarly, there’s no easily found context where ‘Midori is an accountant’ and ‘Akiko is an accountant’ are both on-topic but their conjunction isn’t. But if ‘and’ added topic, there should be such contexts (Berto, 2022, ch. 2). Surely there will be contexts where some  $\varphi$  and  $\psi$  are both acceptable for they’re found probable enough, while their conjunction isn’t; but probability of truth, too, is the business of truth conditions, not topic.

(2022). The language  $\mathcal{L}$  has a countable set  $\mathcal{L}_{AT}$  of atoms,  $p, q, r$  ( $p_1, p_2, \dots$ ), negation  $\neg$ , conjunction,  $\wedge$ , disjunction  $\vee$ , the box of necessity  $\square$ , two-place operators  $\approx$  and  $\triangleright$ , round parentheses as auxiliary symbols  $(, )$ . We use  $\varphi, \psi, \chi, \dots$ , as metavariables for formulas of  $\mathcal{L}$ . The well-formed formulas are the atoms and, if  $\varphi$  and  $\psi$  are well-formed, so are the following:

$$\neg\varphi \mid \square\varphi \mid (\varphi \wedge \psi) \mid (\varphi \vee \psi) \mid (\varphi \approx \psi) \mid (\varphi \triangleright \psi)$$

Outermost brackets are usually omitted. We identify  $\mathcal{L}$  with the set of its well-formed formulas. Read ‘ $\varphi \approx \psi$ ’ as saying that  $\varphi$  and  $\psi$  have the same topic, ‘ $\varphi \triangleright \psi$ ’ as saying that the content of  $\psi$  is part of that of  $\varphi$ ; then  $\varphi \trianglelefteq \psi := (\varphi \triangleright \psi) \wedge (\psi \triangleright \varphi)$  expresses same-saying. We use  $x, y, z$  ( $x_1, x_2, \dots$ ) for topics;  $w, w_1, w_2, \dots$  for possible worlds;  $P, Q, R, \dots$  for propositions. The semantics will recursively assign a thick proposition  $[\varphi]$  to well-formed sentences  $\varphi$  of  $\mathcal{L}$ . We use  $[\varphi] = \mathbb{W}_{[\varphi]}$  for the truth conditions of  $\varphi$ ,  $[\varphi] = \mathbb{T}_{[\varphi]}$  for its topic. (This slick notation is due to Peter Hawke, who co-authored that chapter.)

A *frame* for  $\mathcal{L}$  is a triple  $\mathfrak{F} = \langle W, \mathcal{T}, \oplus \rangle$ .  $W$  is a non-empty set of possible worlds;  $\mathcal{T}$  is a non-empty set of topics; and  $\oplus : \mathcal{T} \times \mathcal{T} \rightarrow \mathcal{T}$  is topic fusion: an idempotent ( $x \oplus x = x$ ), commutative ( $x \oplus y = y \oplus x$ ), associative ( $(x \oplus y) \oplus z = x \oplus (y \oplus z)$ ) operation making topics part of larger topics. So  $\langle \mathcal{T}, \oplus \rangle$  is a join semilattice. For simplicity, fusion is unrestricted:  $\forall xy \in \mathcal{T} \exists z \in \mathcal{T} (z = x \oplus y)$ . One can then define topic parthood as  $x \leq y := x \oplus y = y$ , a partial ordering.

A *model*  $\mathfrak{M} = \langle W, \mathcal{T}, \oplus, \mathbf{c}, \mathbf{t} \rangle$  adds to a frame two interpretation functions. To each atom  $p$ , the first assigns a truth set,  $\mathbf{c}(p) \subseteq W = [p]$ , giving the Conditions under which  $p$  is true; the second, a Topic  $\mathbf{t}(p) = [p] \in \mathcal{T}$ .<sup>7</sup> The two are extended to the truth-functional composites as follows:

- $[\neg\varphi] = W \setminus [\varphi]$
- $[\varphi \wedge \psi] = [\varphi] \cap [\psi]$
- $[\varphi \vee \psi] = [\varphi] \cup [\psi]$
- $[\neg\varphi] = [\varphi]$
- $[\varphi \wedge \psi] = [\varphi \vee \psi] = [\varphi] \oplus [\psi]$

The left-hand recursion gives the usual Boolean algebra of thin propositions corresponding to the truth-functional vocabulary. The right-hand recursion secures the topic-transparency of that vocabulary.  $\triangleright$ ,  $\square$ ,  $\approx$  are global operators:

<sup>7</sup>Notice that distinct atoms can be about the same topic. How to account for topicality for the subsentential constituents of sentences, of course, is an issue that doesn’t pop up for  $\mathcal{L}$ . Interesting ideas can be found in Hawke (2016); Badura (2021); Plebani and Spolaore (2023).



- $[\varphi \supseteq \psi] = W$ , if  $[\varphi] \subseteq [\psi]$  and  $[\psi] \leq [\varphi]$ . Else:  $[\varphi \supseteq \psi] = \emptyset$ .
- $[\Box\varphi] = W$ , if  $[\varphi] = W$ . Else:  $[\Box\varphi] = \emptyset$ .
- $[\varphi \approx \psi] = W$ , if  $[\psi] = [\varphi]$ . Else:  $[\varphi \approx \psi] = \emptyset$ .

These only get truth sets, as we just care about the conditions under which the relevant sentences are true.<sup>8</sup> So, relative to an interpretation in  $\mathfrak{M}$ , the (thick) content of a sentence  $\varphi$ , what it says, is  $[\varphi] = \langle [\varphi], [\varphi] \rangle$ . Entailment is, completely standardly, truth preservation at all worlds of all models:  $\varphi_1, \dots, \varphi_n \models \varphi$  if for every model  $\mathfrak{M}$  and  $w \in W$ , if  $w \in [\varphi_i]$  for every  $i$ , then  $w \in [\varphi]$ . Validity is truth at all worlds of all models:  $\models \varphi$  if for every model  $\mathfrak{M}$  and  $w \in W$ ,  $w \in [\varphi]$ .

Here's some (in)validities delivered by the semantics (the proofs are easy):

1.  $\models (\varphi \wedge \psi) \supseteq \varphi$
2.  $\not\models \varphi \supseteq (\varphi \vee \psi)$
3.  $\models \varphi \trianglelefteq \varphi \wedge \varphi$
4.  $\models (\varphi \wedge \psi) \trianglelefteq (\psi \wedge \varphi)$
5.  $\not\models \varphi \supseteq (\varphi \wedge (\varphi \vee \psi))$
6.  $\not\models \varphi \supseteq \neg(\neg\varphi \wedge \psi)$
7.  $\models \varphi \approx \neg\neg\varphi$
8.  $\models (\varphi \wedge \psi) \approx (\varphi \vee \psi)$
9.  $\Box\varphi \wedge \Box\psi \not\models \varphi \approx \psi$
10.  $\Box(\varphi \equiv \psi) \not\models \varphi \approx \psi$

Here's how these capture intuitions of same-saying. To begin with, *saying that* should transmit down to the parts of what is said, though not to the mere entailments (Yablo, 2014, ch. 1). Aisha says: 'Midori is tall and thin'. Bethany says: 'Midori is tall'. What Bethany said has already been said by Aisha, who also said more: that Midori is thin as well. (Why Bethany said that after Aisha is an interesting issue better left to pragmatics.)

---

<sup>8</sup>And, currently I take no stance on how topicality may work for them. However, for some very promising ideas on the topics of non-extensional operators, in a two-component account of content like the one just sketched, see Ferguson (2023a,b,c).

But Bethany hasn't thereby said that Midori is tall or a footballer. Sure, that easily follows from what she has said. Bethany may be rationally committed to that, supposing one is committed to all the logical consequences of what one says. She hasn't said *that*, however. (1) and (2) capture this by marking a difference between conjunction and disjunction: the content of a conjunction includes, and not just entails, that of its conjuncts; but the content of a disjunct does not perforce include that of the disjunction, in spite of entailing it. They also tell us why: the other disjunct can bring in extra topic. Following Yablo (2014) again,  $\varphi \vee \psi$  can *say less about more* than  $\varphi$ : the disjunction can address a larger topic than that of one of its disjuncts, even while being less informative in that it rules out fewer worlds.  $\varphi \wedge \psi$  can *say more about more* with respect to  $\varphi$ : it can both address a larger topic and be more informative, if it rules out more worlds. (1) and (2) are a widely recognized mark of a topic-sensitive semantics:

A paradigm of inclusion, I take it, is the relation that simple conjunctions bear to their conjuncts – the relation *Snow is white and expensive* bears, for example, to *Snow is white*. A paradigm of noninclusion is the relation disjuncts bear to disjunctions; *Snow is white* does not have *Snow is white or expensive* as a part. (Yablo, 2014, 11)

A guiding principle behind the understanding of partial content is that the content of  $A$  and  $B$  should each be part of the content of  $A \wedge B$  but that the content of  $A \vee B$  should not in general be part of the content of either  $A$  or  $B$ . (Fine, 2016, 200)

Aisha says: 'Scottish grass is green'. Bethany says: 'Scottish grass is green and it is green'. This sounds marked as redundant. That's because what Bethany said is but what Aisha just said. (3) captures this:  $\varphi$  and  $\varphi \wedge \varphi$  say the same thing. As Bethany is sticking to what has been said redundantly, we may want to step outside semantics and into pragmatics to make sense of her move: perhaps she wanted to stress the pervasiveness of green as the colour of Scottish grass.

Aisha says: 'Midori is an accountant and a football player'. Bethany says: 'Midori is a football player and an accountant'. Perhaps Bethany wanted to stress that playing football is what really matters for Midori. We resort to pragmatics again, because what Bethany has said just is what Aisha has said. (4) captures this:  $\varphi \wedge \psi$  and  $\psi \wedge \varphi$  say the same thing. (That happens, to be sure, when 'and' encodes order-insensitive, truth-functional conjunction. Sometimes 'and' can encode some kind of – e.g., temporal – ordering: 'John

went to the hospital and got ill’ can then say something different from ‘John got ill and went to the hospital’.)

Aisha says: ‘Midori is happy’. She hasn’t thereby said that Midori is happy and either Midori is happy or extremally disconnectedness is no hereditary property of topological spaces. (5) captures this. Aisha says: ‘The car is out of fuel’. She hasn’t thereby said it’s not the case that the car has fuel but the gauge is stuck. (6) captures this. (You may already have guessed where these take us once knowledge or belief ascriptions step in. We’ll get there in the next section.)

(7)-(10), taken together, give the core of the hyperintensional semantics. (7) and (8) express the transparency of the truth-functional logical vocabulary: that grass is green and that grass isn’t green are about the exact same topic. Of course, they say opposite things about that same topic, hence they are different contents. That Carlos is short and handsome and that Carlos is handsome and short are about the exact same topic, too, say, Carlos’ height and looks. (9) and (10) guarantee the possibility of topic-diverging necessities and co-necessities: that  $2 + 2 = 4$  and that  $3 + 3 = 6$  are both necessary, but about different things. That vixens are female foxes and that clopen sets are both open and closed – ditto. That grass is green and that it’s true that grass is green are about different things too, albeit co-intensional.

This is the barest sketch of a hyperintensional semantics. The view is developed in more detail (and confronted with some problems) in Berto (2022).

## 6 ... Goes a Long Way

Topic-sensitivity can do a lot of work in epistemic logic. In Berto (2022), operators are added to a language essentially like  $\mathcal{L}$  above, expressing conditional belief ( $B^\varphi\psi$ : one believes that  $\psi$  conditional on  $\varphi$ ), belief revision ( $[\varphi]\psi$ : after revising one’s beliefs by  $\varphi$ ,  $\psi$  is the case), knowability relative to information ( $K^\varphi\psi$ :  $\psi$  would be knowable for one given information  $\varphi$ ), suppositional thinking ( $I^\varphi\psi$ : supposing  $\varphi$ , one imagines that  $\psi$ ). They are all modals (variously restricted quantifiers over possible worlds, or constructions thereof) whose truth conditions are given in terms of topic-sensitive contents; hence they are labeled as Topic-Sensitive Intentional Modals (TSIMs).

Here’s a first application: we can ground a distinction, often made in the literature, between two kinds of logical closure principles for attitudes. E.g., Holliday (2012) takes  $\wedge$ -Elimination within knowledge operators (in the form:  $K(\varphi \wedge \psi) \supset K\psi$ ) as a *pure* (contrast *deductive*) epistemic closure principle (see also Yablo (2017)). A deductive closure principle from  $\varphi_1, \dots, \varphi_n$  to  $\psi$  has it that if an agent comes to believe  $\psi$  starting from  $\varphi_1, \dots, \varphi_n$ , by competent

deduction, and all the while knowing each of  $\varphi_1, \dots, \varphi_n$ , then the agent knows  $\psi$ . This can always go wrong for realistic agents: the deduction may be too complex for our Joe Bloggs. But  $\wedge$ -Elim, qua pure closure principle, is such that ‘an agent cannot know  $\varphi \wedge \psi$  without knowing  $\psi$  – regardless of whether the agent came to believe  $\psi$  by “competent deduction” from  $\varphi \wedge \psi$ ’ (Holliday, 2012, 15). Pure closure is (as one referee appropriately asked to mention) irrespective of the inferential skills or behaviour of the relevant agent.

This is so, TSIM theory explains, because when  $P$  is exactly about  $x$ , and one thinks (believes, knows, supposes, etc.) that  $P$ , one must be – taking a couple of metaphors from (Yablo, 2014, 39) – ‘attentive to everything within  $x$ ’. But one can be ‘oblivious to matters lying outside of  $x$ ’ although there are propositions  $Q$  entailed by  $P$  which are about those (subject) matters. That is, the topic-sensitivity of attitudes, inherited by that of the propositions making for their contents, delivers some natural logical closure and non-closure properties. *Thinking that* (believing that, knowing that, supposing that, etc.) should transmit down to the parts of what one thinks (Yablo calls this ‘immanent closure’) though not to the mere entailments.

On the immanent closure side, e.g., when Aisha believes (plainly, or conditional on something else) that Midori is tall and thin, she thereby automatically believes that Midori is tall ( $B^\varphi(\psi \wedge \chi) \vDash B^\varphi\psi$  is one validity in the TSIM semantics). That works also for supposedly anarchic mental activities like imagining, which is subject to voluntary control in ways belief is not: try and *imagine* that Midori is tall and thin without imagining that she is tall. That would be a bit like imagining that Midori is tall without imagining that she is tall, wouldn’t it? And so we find Williamson endorsing  $\wedge$ -Elim as a pure or immanent closure principle, first for knowledge:

... Knowledge of a conjunction is *already* knowledge of its conjuncts. ... There is no obstacle here to the idea that knowing a conjunction *constitutes* knowing its conjuncts, just as, in mathematics, we may *count* a proof of a conjunction as a proof of its conjuncts, so that if  $p \wedge q$  is proved then  $p$  is proved, not just provable. (Williamson, 2000, 282-3)

He then generalizes and conjectures that  $\wedge$ -Elim may hold for all positive attitudes (Ibid.): in believing a conjunction, one believes the conjuncts; in conceiving a conjunction, one conceives the conjuncts, etc.

This is so, TSIM theory explains, because what  $\varphi$  is about is (a proper) part of what  $\varphi \wedge \psi$  is about. And by thinking about the whole, one has already thought about the parts: there’s nothing more for one to do, such that if one failed to do it one would be thinking that  $\varphi \wedge \psi$  without thinking

that  $\varphi$ . Thinking that  $\varphi$  is understood here as having a contentful mental state, endowed with intentionality and directed towards what it's about. This draws a wedge with a merely syntactic conception of thought. If thinking that  $\varphi \wedge \psi$  was having a sentence (say, of mentalese) tokened in the head, one may need to do something to move from thinking that  $\varphi \wedge \psi$  to thinking that  $\varphi$ . And if one failed to apply  $\wedge$ -Elim to one's mentalese sentence, one would have ' $\varphi \wedge \psi$ ' in the head without having ' $\varphi$ ' there.

But immanent closure is weaker than full closure under entailment, and the TSIMs generally are not fully closed. Immanent closure is well-suited to capture pure closure as characterized above. One may think that  $\varphi$  without thinking that  $\psi$  although the former entails the latter, because one is not thinking about what  $\psi$  is about. That may happen for different reasons. One may think that  $\varphi$  (plainly, or given or supposing or conditionally on something else) without thinking that  $\varphi \vee \psi$  although the former easily entails the latter ( $X^x\varphi \not\equiv X^x(\varphi \vee \psi)$  is one invalidity in the semantics for a number of TSIMs  $X$ ), because one lacks some concept needed to grasp what  $\psi$  is about; and one cannot have attitudes such as knowing, believing, or even supposing, towards contents one cannot grasp.<sup>9</sup> Williamson again:

$\wedge$ -elimination has a special status. It may be brought out by a comparison with the equally canonical  $\vee$ -introduction inference to the disjunction  $p \vee q$  from the disjunct  $p$  or from the disjunct  $q$ . Although the validity of  $\vee$ -introduction is closely tied to the meaning of  $\vee$ , a perfect logician who knows  $p$  may lack the empirical concepts to grasp (understand) the other disjunct  $q$ . Since knowing a proposition involves grasping it, and grasping a complex proposition involves grasping its constituents, such a logician is in no position to grasp  $p \vee q$ , and therefore does not know  $p \vee q$ . In contrast, those who know a conjunction grasp its conjunct, for they grasp the conjunction. (Williamson, 2000, 282-3)

Aisha may believe that Midori is tall without believing that either Midori is tall or extremally disconnectedness is no hereditary property of topological spaces, because Aisha has no idea what topological spaces are. Topology is an alien topic to Aisha. Thus, the TSIMs are good for modeling agents with certain conceptual limitations. As Williamson remarks, these can affect a 'perfect logician'. So they must be of a different kind from limitations due

---

<sup>9</sup>For a recent work on how to tackle the problem of conjunction elimination and disjunction introduction in attitude logic, that also addresses how to accommodate (hyper-intensional counterparts of) these, see Jespersen (2023).

to the boundaries of one's deductive capacities. Aisha's conceptual incompetence has little to do with her bounded inferential resources: the disjunction is just one basic inferential step away.

You may think that Aisha is rationally committed to believing the disjunction. You may also think that, by believing that Midori is happy, Aisha is committed to believing that Midori is happy or Midori is happy and extremely disconnectedness is no hereditary property of topological spaces: one is rationally committed to being on top of *all* concepts. 'Conceptually omniscient' agents – agents who can think about anything, that is, who can entertain any propositional content – represent a normative ideal.

Or, you may have a Harmanian view of rationality: sometimes one should not think that  $\psi$  although one thinks that  $\varphi$  and the former entails the latter, *and* one is even perfectly on top of the concepts needed to think about what  $\psi$  is about. Chased by a predator, you run towards a small stream of water. Will you make it if you try and jump to the other side? Before blindly trying, you quickly simulate jumping in your mind; you bring in your knowledge or beliefs on the width of the stream, your physical abilities, etc., and you come to believe that you will make it if you jump. Are you committed to coming to believe that either you will make it or grass is green, or there's life on Kepler-442b, or... Well, sure, you can entertain all such contents. But Harmanian normativity has it that you'd better focus on jumping now, before the predator is on you. Thus the TSIMs are good for capturing agents whose mental states are sensitive to relevance, in that they keep their suppositional and belief management procedures *on-topic*.

There's a wave of pushbacks against Harman in formal epistemology (Christensen, 2004; Smithies, 2015; Titelbaum, 2015). It is sometimes stressed that the theoretical costs of coming up with formal epistemologies modeling non-logically-omniscient agents outweigh the benefits of sticking to standard Bayesianism or normal modal-epistemic logic. I don't think this debate should prevent one from attempting to formally capture ideas concerning Harmanian agents. Epistemic logicians won't wait for normative theorists to come to an agreement on the principles of rationality before they start building their models.

Aisha knows that the car is out of fuel on the basis of having checked the gauge, which reads 'empty'. Is she thereby positioned to know that it's false that the car has fuel but the gauge is stuck? Bethany may ask Aisha: the car is old; could it be the gauge is stuck? Aisha may then retract her knowledge claim. This connects to the debate on epistemic closure, 'one of the most significant disputes in epistemology over the last forty years' (Kvanvig, 2006, 256). One may list Williamson (2000), Hawthorne (2004), Roush (2010), Kripke (2011) among the yea-sayers; Dretske (1970), Nozick (1981), Lawlor

(2005), Holliday (2015), Hawke (2016), Alsdpector-Kelly (2019) among the nay-sayers. I think the jury is out on this.

But look at the putative counterexamples to closure: you know it's a zebra on the basis of your sensory perception; you're not thereby positioned to know it's no cleverly disguised mule (Dretske, 1970). You know the table is red on the same kind of basis; you're not thereby positioned to know it's no white table under a deceiving red light (Cohen, 2002). You know where your car is parked on the basis of your memory of having left it there a few minutes ago. You're not thereby positioned to know it's false that it has been stolen and it's not there (Vogel, 1990). And so on. These are all of the form: one can know on a certain basis that  $\varphi$  without being positioned to know, on that same basis, that  $\neg(\neg\varphi \wedge \psi)$ . So they are all rooted in the addition of subject matter the information positioning you to know that  $\varphi$  is supposedly insensitive to, or incapable to provide you with evidence for, or, plainly, not *about*.

Closure nay-sayers have it that whatever justifies your belief in the former fails to transmit to the latter. They will also think that there's no other route for you to get to know the latter. They may be wrong on this. But TSIM theory can capture, via our invalidity (6) from the previous section, how knowing something on a certain basis doesn't mean being positioned to know something else which is logically entailed, on *that same* basis. In particular, in the TSIM knowability-relative-to-information setting,  $K^x\varphi \not\equiv K^x\neg(\neg\varphi \wedge \psi)$  (see Berto and Hawke (2021) and ch. 4 of Berto (2022) for some details).

## 7 Conclusion

Hyperintensional distinctions, motivated by purely semantic considerations of aboutness, can be fruitfully put to use in formal and mainstream epistemology. A hyperintensional account of content may be safe from the Objection from Overfitting, insofar as it's based on robust judgments of same-saying. These may be difficult to account for at a cognitive level alien to content, without introducing devices which, once developed in some detail, will look as semantics under another name and/or just as complex as the supposedly more complex hyperintensional rivals.

## Acknowledgments

Ideas on the Objection from Overfitting that made it into this paper were presented to the Logic & Metaphysics Group in Arché at St Andrews and during my José Gaos Chair lectures at the UNAM in Mexico City. I am grateful to both audiences, as well as to three anonymous referees of this Journal, for helpful remarks and objections. This research is published within the project ‘The Logic of Conceivability’, funded by the European Research Council (ERC CoG), Grant Number 681404.

*Statement: The author declares that there are no financial or non-financial conflicts of interest directly or indirectly related to the work submitted. The work includes no research involving human participants and/or animals.*

## References

- M. Alspector-Kelly. *Against Knowledge Closure*. Cambridge University Press, Cambridge, 2019.
- A.R. Anderson and N. Belnap. *Entailment: The Logic of Relevance and Necessity I*. Princeton University Press, Princeton, 1975.
- A.R. Anderson, N. Belnap, and J.M. Dunn. *Entailment: The Logic of Relevance and Necessity II*. Princeton University Press, Princeton, 1992.
- C. Badura. More aboutness in imagination. *Journal of Philosophical Logic*, 50:523–47, 2021.
- J. Barwise and J. Perry. *Situations and Attitudes*. CSLI Publications, Stanford, 1983.
- F. Berto. *Topics of Thought. The Logic of Knowledge, Belief, Imagination*. Oxford University Press, Oxford, 2022.
- F. Berto and P. Hawke. Knowability relative to information. *Mind*, 130:1–33, 2021.
- F. Berto and L. Hornischer. Cognitive synonymy: A dead parrot? *Philosophical Studies*, 180:2727–52, 2023.
- F. Berto and M. Jago. *Impossible Worlds*. Oxford University Press, Oxford, 2019.



- J. Branquinho. Are Salmon's 'guises' disguised Fregean senses? *Analysis*, 50:19–24, 1990.
- R. Carnap. *Meaning and Necessity*. Chicago University Press, Chicago, 1947.
- G. Chierchia and S. McConnell-Ginet. *Meaning and Grammar. An Introduction to Semantics*. MIT Press, Cambridge, MA, 1990.
- D. Christensen. *Putting Logic in its Place: Formal Constraints on Rational Belief*. Oxford University Press, Oxford, 2004.
- S. Cohen. Basic knowledge and the problem of easy knowledge. *Philosophy and Phenomenological Research*, 65:309–29, 2002.
- F. Dretske. Epistemic operators. *Journal of Philosophy*, 67:1007–23, 1970.
- J.M. Dunn and G. Restall. Relevance logic. In D. Gabbay and F. Guenther, editors, *Handbook of Philosophical Logic*, volume 6, pages 1–136. Kluwer Academic, Dordrecht, second edition, 2002.
- M. Duží, B. Jespersen, and P. Materna. *Procedural Semantics for Hyperintensional Logic*. Springer, Dordrecht, 2010.
- M. Duží, B. Jespersen, and M. Kosterec. *Transparent Intensional Logic: Selected Recent Essays*. College Publications, London, 2023.
- D. Edgington. On conditionals. *Mind*, 104:235–329, 1995.
- T. Ferguson. Subject-matter and intensional operators i: conditional-agnostic analytic implication. *Philosophical Studies*, 180:1849–1879, 2023a.
- T. Ferguson. Subject-matter and intensional operators ii: Applications to the theory of topic-sensitive intensional modals. *Journal of Philosophical Logic*, 52:1673–1701, 2023b.
- T. Ferguson. Subject-matter and intensional operators iii: State-sensitive subject matter and topic-sufficiency. *Review of Symbolic Logic*, Online First:doi:10.1017/S1755020323000230, 2023c.
- K. Fine. Counterfactuals without possible worlds. *Journal of Philosophy*, 109:221–46, 2012.
- K. Fine. Angelic content. *Journal of Philosophical Logic*, 45:199–226, 2016.

- K. Fine. A theory of truthmaker content I: Conjunction, disjunction and negation. *Journal of Philosophical Logic*, 46:625–74, 2017.
- K. Fine. Yablo on subject matter. *Philosophical Studies*, 177:129–71, 2020.
- K. Fine and M. Jago. Logic for exact entailment. *Review of Symbolic Logic*, 12:536–56, 2018.
- G. Forbes. Review of *Frege’s Puzzle*, by Nathan Salmon. *Philosophical Review*, 96:455–8, 1987.
- P. Gärdenfors. Belief revisions and the ramsey test for conditionals. *Philosophical Review*, 95:81–93, 1986.
- K. Gemes. A new theory of content I: Basic content. *Journal of Philosophical Logic*, 23:595–620, 1994.
- K. Gemes. A new theory of content II: Model theory and some alternatives. *Journal of Philosophical Logic*, 26:449–76, 1997.
- N. Goodman. About. *Mind*, 70:1–24, 1961.
- H.P. Grice. *Studies in the Way of Words*. Harvard University Press, Cambridge, MA, 1989.
- D. Hajek. Probabilities of conditionals. *Journal of Philosophical Logic*, 18:423–8, 1989.
- P. Hawke. Questions, topics and restricted closure. *Philosophical Studies*, 73:2759–84, 2016.
- P. Hawke. Theories of aboutness. *Australasian Journal of Philosophy*, 96:697–723, 2018.
- P. Hawke and A. Özgün. Truthmaker semantics for epistemic logic. In F. Faroldi and F. Van de Putte, editors, *Outstanding Contributions to Logic: Kit Fine on Truthmakers, Relevance, and Non-Classical Logic*, pages 295–335. Springer, Dordrecht, 2023.
- P. Hawke, L. Hornischer, and F. Berto. Truth, topicality, and transparency: one-component versus two component semantics. *Linguistics and Philosophy*, Online first, 2024.
- J. Hawthorne. *Knowledge and Lotteries*. Oxford University Press, Oxford, 2004.

- J. Hintikka. *Knowledge and Belief. An Introduction to the Logic of the Two Notions*. Cornell University Press, Ithaca, NY, 1962.
- W. Holliday. *Knowing What Follows: Epistemic Closure and Epistemic Logic*. PhD Dissertation, Stanford, 2012.
- W. Holliday. Fallibilism and multiple paths to knowledge. In T. Gendler and J. Hawthorne, editors, *Oxford Studies in Epistemology*, volume 5, pages 97–144. Oxford University Press, Oxford, 2015.
- L. Humberstone. Parts and partitions. *Theoria*, 66:41–82, 2008.
- M. Jago. *The Impossible. An Essay on Hyperintensionality*. Oxford University Press, Oxford, 2014.
- B. Jespersen. Is act theory a propositional logic without logic? *The Journal of Philosophy*, 120:484–510, 2023.
- J. King. Structured propositions and sentence structure. *Journal of Philosophical Logic*, 25:495–521, 1996.
- J. King. *The Nature and Structure of Content*. Oxford University Press, Oxford, 2007.
- J. King, J. Speaks, and S. Soames. *New Thinking About Propositions*. Oxford University Press, New York, 2014.
- A. Kocurek. Counterpossibles. *Philosophy compass*, page Online First, 2021.
- S. Krämer. Mighty belief revision. *Journal of Philosophical Logic*, 51:1175–1213, 2022.
- S. Kripke. A puzzle about belief. In A. Margalit, editor, *Meaning and Use*, pages 239–83. Reidel, Dordrecht, 1979.
- S. Kripke. Nozick on knowledge. In *Philosophical Troubles: Collected Papers, Volume 1*. Oxford University Press, Oxford, 2011.
- J. Kvanvig. *The Knowability Paradox*. Oxford University Press, Oxford, 2006.
- K. Lawlor. Living without closure. *Grazer Philosophische Studien*, 697:25–49, 2005.
- D. Lewis. *Counterfactuals*. Blackwell, Oxford, 1973.

- D. Lewis. Probabilities of conditionals and conditional probabilities. *Philosophical Review*, 95:581–9, 1976.
- D. Lewis. Relevant implication. *Theoria*, 54:161–74, 1988a.
- D. Lewis. Statements partly about observation. *Philosophical Papers*, 17: 1–31, 1988b.
- F. Moltmann. An object-based truthmaker semantics for modals. *Philosophical Issues*, 28:255–88, 2018.
- D. Nolan. Impossible worlds: a modest approach. *Notre Dame Journal of Formal Logic*, 38:535–72, 1997.
- D. Nolan. Hyperintensional metaphysics. *Philosophical Studies*, 171:149–70, 2014.
- R. Nozick. *Philosophical Explanations*. Harvard University Press, Cambridge, MA, 1981.
- G. Oppy. Semantics for propositional attitude ascriptions. *Philosophical Studies*, 67:1–18, 1992.
- J. Perry. Possible worlds and subject matter. In *The Problem of the Essential Indexical and Other Essays*, pages 145–60. CSLI publications, Stanford, 1989.
- M. Plebani. Why aboutness matters: Meta-fictionalism as a case study. *Philosophia*, Online First: doi 10.1007/s11406-020-00272-9, 2020.
- M. Plebani and G. Spolaore. Subject matter: A modest proposal. *The Philosophical Quarterly*, 71:605–22, 2021.
- M. Plebani and G. Spolaore. Predicative subject matter. *Philosophical Studies*, 181:247–65, 2023.
- G. Priest. *Towards Non-Being: the Logic and Metaphysics of Intentionality*, 2nd ed. Oxford University Press, Oxford, 2016.
- F.P. Ramsey. General propositions and causality. In D.H. Mellor, editor, *Philosophical Papers*, pages 145–63. Cambridge University Press, Cambridge, 1990.
- D. Ripley. Structures and circumstances: Two ways to fine-grain propositions. *Synthese*, 189:97–118, 2012.

- D. Rothschild. Do indicative conditionals express propositions? *Nous*, 47: 49–68, 2013.
- D. Rothschild. Living in a material world: A critical notice of suppose and tell: The semantics and heuristics of conditionals by timothy williamson. *Mind*, page Online First, 2021.
- S. Roush. Closure on skepticism. *Journal of Philosophy*, 107:243–56, 2010.
- N. Salmon. *Frege's Puzzle*. MIT Press, Cambridge, MA, 1986.
- A. Schipper. Aboutness and negative truths: a modest strategy for truth-maker theorists. *Synthese*, 195:3685–722, 2018.
- A. Schipper. Aboutness and ontology: a modest approach to truthmakers. *Philosophical Studies*, 177:505–33, 2020.
- D. Smithies. Ideal rationality and logical omniscience. *Synthese*, 192:2769–93, 2015.
- S. Soames. Lost innocence. *Linguistics and Philosophy*, 8:59–71, 1985.
- S. Soames. Direct reference, propositional attitudes, and semantic content. *Philosophical Topics*, 15:47–87, 1987.
- S. Soames. *What Is Meaning?* Princeton University Press, Princeton, NJ, 2010.
- R. Stalnaker. A theory of conditionals. In Nicholas Rescher, editor, *Studies in Logical Theory (American Philosophical Quarterly Monographs 2)*, pages 98–112. Blackwell, Oxford, 1968.
- S. Standefer. Hyperintensionality in relevant logics. In N. Alechina, A. Herzig, and F. Liang, editors, *Logic, Rationality and Interaction - LORI 2023*, pages 238–50. Springer, Dordrecht, 2023.
- M. Titelbaum. Rationality's fixed point. In T. Gendler and J. Hawthorne, editors, *Oxford Studies in Epistemology*, volume 5, pages 253–94. Oxford University Press, Oxford, 2015.
- A. Tversky and D. Kahneman. Judgment under uncertainty: Heuristics and biases. *Science*, 185:1124–31, 1974.
- J. Vogel. Are there counterexamples to the closure principle? In M. Roth and G. Ross, editors, *Doubting: Contemporary Perspectives on Skepticism*, pages 13–29. Kluwer, Dordrecht, 1990.

- T. Williamson. *Knowledge and Its Limits*. Oxford University Press, Oxford, 2000.
- T. Williamson. *Suppose and Tell. The Semantics and Heuristics of Conditionals*. Oxford University Press, Oxford, 2020.
- T. Williamson. Degrees of freedom: Is good philosophy bad science? *Disputatio*, 13:73–91, 2021.
- L. Wittgenstein. *Tractatus logico-philosophicus*. Routledge & Kegan Paul, London, 1921/22.
- S. Yablo. *Aboutness*. Princeton University Press, Princeton, 2014.
- S. Yablo. Open knowledge and changing the subject. *Philosophical Studies*, 174:1047–71, 2017.
- S. Yablo. Reply to Fine on *Aboutness*. *Philosophical Studies*, 175:1495–512, 2018.