

## The Nature of Mental Content

© 2007 by Damon Woolsey

“Mental content” is the representational content of mental states. When I am in pain, for example, I am in a mental state that represents something: the *pain*. Or when I think about the moon, I am in a mental state that represents, or is about, the moon. Strange as it might seem, for all that has been said about mental content, no one has a complete story of *what* it might be. Externalists hold that some or all mental content is external, or wide, since it supervenes on relations between the mind and the external world. They have a difficult time, however, explaining how a relation can hold between a mind and something that doesn’t exist, as when I am hallucinating, or thinking about Aphrodite. Internalists, on the other hand, maintain that all mental content is intrinsic, or narrow, meaning that it supervenes on the intrinsic properties of the mind. They have a harder time explaining how the mind can “refer” to external objects; in other words, how is it that a particular mental representation is *about* something external to the mind, if the representation itself is narrow. How could it be, for example, that my thought about the moon is about *the moon*, unless that content depends on an external relation?

Intuitions play a huge role in the debate over the nature of mental content. Consider, for example, Robert Stalnaker’s take on this issue:

Isn’t it obvious that semantic properties, and intentional properties generally, are relational properties...And isn’t it obvious that relations depend...on more than the intrinsic properties of one of the things related...This...should follow from any account of representation that holds that we can talk and think...about things and properties outside of ourselves. (1989, p. 288)

In a similar vein, an internalist could hold it to be a near analytic truth that mental content is intrinsic: it is non-controversial that the mind is strongly correlated with (or supervenes on, or is

reducible to) the brain, thus mental states are strongly correlated with (or supervene on, or are reducible to) brain states. Furthermore, the notion of “content” implies being “contained in” (or “part of”) something, and so mental content is contained in (or part of) a mental state. In any event, it just doesn’t make any sense to say that mental content can exist “outside” of a mind. So isn’t it obvious that mental content is intrinsic?

Of course, I don’t hold either of the above to be knock-down arguments for their respective positions; rather, they seem to result from competing intuitions. The difference between them is that Stalnaker has focused on the relation between mental states and external objects, leaving open the question of what minds are, whereas the internalist has focused on the mind itself, leaving open the question of what comprises mental content. From this point I suppose that the externalist owes us an appropriate view of what a mind is, whereas the internalist owes us an explanation of what comprises mental content; but I do not intend to provide such explanations here. There is, I believe, something more fundamental underlying these differing intuitions that needs to be brought to light. My goal here is to do this by examining and responding to two popular arguments for externalism, whereupon I will develop a conceptual thesis of meaning and reference that is immune to these arguments, because it treats meaning and reference not as two separate entities, but as two sides of the same coin. The problem with the prevailing views is that both sides treat the “aboutness” of mental states (a.k.a. “intentionality”) as a *metaphysical relation* between the mind and external objects. On my view, there simply is no relation.

### **The First Argument for Externalism (FAE)**

I begin with what I take to be the general form of the argument for externalism (hereafter FAE) given by Putnam (1975, 218 ff.):

1. Assumptions:
  - (i). Knowing the meaning of a term is to be in a certain intrinsic mental state.
  - (ii). The meaning of a term (intension) determines the extension of that term.
2. According to (i) and (ii), a term as used by two different agents in the same intrinsic mental state must have the same extension in both cases.
3. In certain counterfactual situations, however, two agents can be in the same intrinsic mental state, but their terms can have different extensions.
4. Mental state is therefore not sufficient for determining extension, and so we must reject either (i) or (ii).
5. We should keep (ii).
6. Therefore, intension is not determined by mental state.

This appears to be a valid argument, but there are several responses that can be made to it. First, one could simply accept Putnam's argument and remain an internalist. One such view that allows for this is *Epistemicism*, where any apparent vagueness in the extension of a term is due to an incomplete grasp of that term's intension. Thus there is a gap between (i) and (ii) large enough to make premise 2 false. However, Epistemicism carries with it a weighty ontological commitment to distinct boundaries. On such a view, for example, the difference between a heap of sand and a non-heap of sand consists in a *single* grain of sand (cf. Sorenson, p. 1). Another possible response is that (ii) might be too strong, and Putnam doesn't provide any independent grounds for keeping it rather than (i). Denying (ii) would have the result that intension does not sufficiently determine extension – but so what? We could simply accept that intension *underdetermines* extension, and that there is a “loose fit” between our words and the world. Do we really require that the meanings of our words have absolute surgical precision? As

Wittgenstein pointed out: “We do not know the boundaries because none have been drawn. To repeat, we can draw a boundary – for a special purpose. Does it take that to make the concept usable? Not at all!” (p. 28, § 69) This, if correct, suggests that the third premise is false. Since intension underdetermines extension, there is no good reason why an intension couldn’t apply to a variety of natural kinds. For example, we could say that ‘water’ picks out H<sub>2</sub>O, XYZ, LMNOP, and any other drinkable “watery-ish” substance that we can imagine.<sup>1</sup>

To assume that intensions are individuated according to natural kinds is simply false; some words refer to a disjunction of natural kinds (e.g. jade), and for *most* of our words there simply isn’t any clear boundary between what they do and do not apply to (e.g. ‘games’). It just so happens, however, that different natural kinds most often have profound differences in readily observable characteristics, which led to distinctions early on in the development of language; but this isn’t always the case. One needn’t be a scientist in order to see that there is a difference between iron and carbon, for example, but distinguishing between jadeite and nephrite is another matter. So suppose that XYZ is a colorless, odorless, tasteless, refreshing liquid, such that it is perceptually indistinguishable from H<sub>2</sub>O. Furthermore, suppose that, in addition to H<sub>2</sub>O, XYZ filled some (or part) of the lakes, rivers, etc., in the actual world. Wouldn’t we treat H<sub>2</sub>O and XYZ in the same way as jadeite and nephrite? Wouldn’t we call the first two by the name ‘water’, just as we call the second two by the name ‘jade’? I think that what we *would* do in such a case tells us a lot more about the nature of mental content than what we *actually* do.

Another possible response would be to accept (i) and (ii) just as they are, and simply point out that on *any* account of reference, we can talk about substances in other worlds just as well as

---

<sup>1</sup> I should mention that our knowledge of the chemical structure of H<sub>2</sub>O, combined with the fact that H<sub>2</sub>O seems to be the only watery-ish substance in our universe, has led to the chemical structure of H<sub>2</sub>O becoming part of the meaning of ‘water’. It will be helpful if we assume that this is not the case, because if it were, then Putnam and his twin-earth counterpart really would not be in the same mental state.

our own. Thus, the determination of extension by mental state should work across all possible worlds by the same mechanism as it does in our world. This would again make the third premise false, because that premise is based on restricting the determination of extension by mental state to the substances in our world, and then determining transworld reference according to type-identical relations with the natural kind substances in our world. This is what Putnam (1975, pp. 225 and 232) calls the “same<sub>L</sub>” relation, meaning that “x is the same liquid as y”. According to Putnam, this is a *theoretical* relation, such that one needn’t know what the nature of x is (e.g. that x is H<sub>2</sub>O), in order to give an ostensive definition of it (e.g. ‘that is water’). So even *prior* to the discovery of the chemical structure of H<sub>2</sub>O, when giving an ostensive definition of water, one is somehow “picking out” H<sub>2</sub>O. This means that a necessary and sufficient condition for being water is bearing this sameness relation to what the members of our linguistic community call ‘water’, and *that* stuff (whether we know it or not) has the chemical structure H<sub>2</sub>O. But is this picture correct?

According to Putnam, “‘water’ is stuff that bears a certain similarity relation to the water *around here*.” (1975, p. 234; *emphasis in original*) What I take issue with is the restrictive clause “around here”; no principled reason has been given in *this* argument that would explain why we should accept such a restriction. It is true that arguments for this restriction *have* been given elsewhere (and we shall consider one shortly), but Putnam stated at the outset that we should “assume that nothing is clear in advance.” (1975, p. 216) But without the restrictive clause “around here”, the third premise is simply false. In the first premise we are to assume that (i) intrinsic mental states determine intensions, and (ii) intensions sufficiently determine extensions. If we make these assumptions (as Putnam asks us to do), and *stick with those assumptions* (thereby rejecting the “around here” restriction), then one will reach the conclusion that ‘water’

picks out any drinkable “watery-ish” substance that we can imagine. To see this more clearly, let us revisit Putnam’s twin-earth thought experiment. Since Putnam and his twin-earth counterpart are *ex hypothesi* both in the same intrinsic mental state, then by assumptions (i) and (ii) the extensions of the term they are using *must not differ*. Since the *chemical structure* of the substances that Putnam and his twin identify *does* differ, by *modus ponens* the extension of the term ‘water’ must be a *disjunction* of H<sub>2</sub>O, XYZ, LMNOP, etc. We can conclude as follows: If intrinsic mental state determines intension, then presumably it does so according to a particular set of phenomenal characteristics. Therefore, if intension determines extension, it must identify any and all substances that exemplify that particular set of phenomenal characteristics – unless one accepts the “around here” restriction.

Of course, it may be that we can only properly refer to substances with which we have an appropriate *causal connection*. This, if true, would enable one to formalize the “around here” restriction that Putnam employs, but would require further argumentation. And since the FAE relies on this restriction, it must be an argument independent of the FAE. The reason is that the FAE purports to be a *modus tollens* on assumptions (i) and (ii), thereby purporting to refute internalism on its own grounds, but this will only work if one employs a causal theory of reference (hereafter CTR). And a CTR contradicts the internalist assumption that mental content is *intrinsic*, since a causal relation is just that: a *relation*. If the mechanism by which mental states refer is an external relation, then it is not intrinsic, end of argument. Thus, providing an independent argument for a CTR makes the FAE superfluous. Perhaps the FAE could be a useful pedagogical tool for *showing* that a CTR refutes internalism to those who don’t *understand* this point. Nevertheless, a conclusive argument for a CTR *is* sufficient for refuting internalism. The important point is that a CTR, by itself (i.e. without argumentation), is nothing more than a

plausible *alternative* to internalism. One cannot pretend to assume that mental states sufficiently determine extension, and then sneak in a causal theory of reference – that begs the question.

One cannot use a CTR to argue *against* internalism unless one can show that it is capable of accounting for important relevant factors that an internalist theory cannot. And since a CTR is incompatible with internalism, an argument for a CTR is *ipso facto* an argument against internalism. Barring such an argument, all one has done is provide a plausible *alternative* to internalism. But here the burden is on the internalist to provide a non-relational account of reference. I stated above that, if narrow mental states determine intensions, then presumably they do so according to particular sets of phenomenal characteristics. This would seem to indicate that there is a comparison made between the phenomenal characteristics of objects as they are represented in the mind, and the phenomenal characteristics of objects as they are perceived. But there are obvious flaws with such a view of reference, and so philosophers have used the term ‘intentionality’ as a placeholder for an as-yet-unknown mechanism whereby mental states refer (cf. Putnam 1981, pp. 1-2). This leads us to Putnam’s (1981) argument for a CTR.

### **The Prior Argument for Externalism (PAE)**

Putnam’s (1981) argument for a CTR, the one that should be given prior to (or in conjunction with) the FAE, is what I will call the Prior Argument for Externalism (hereafter PAE). The argument, as I understand it, is quite short:

1. Similarity is neither necessary nor sufficient for reference.
2. Barring similarity, there is no other plausible intrinsic mechanism whereby mental states can refer to external objects (no “magical theories of reference”).

3. Thus, there must be an external connection (e.g. a causal connection) between mental states and what those states refer to.

The first premise is beyond question, since there are many different ways of representing things (e.g. images, names, etc.) that bear little or no resemblance to the things themselves, while many things that bear a strong resemblance to other things (e.g. identical twins) are not representations of the things they resemble. And the second premise is merely a recognition of the fact that if a mental state is to refer to anything other than narrow mental contents, then the mechanism whereby that state refers must serve as a link *between* that state and something in the external world. But does this mechanism have to be a causal connection? It would seem so. In regard to language, for example, Putnam (1981, p. 11) states that there are “language entry rules” that serve as a link between experiences of external objects and linguistic expressions about those objects. Likewise, there are “language exit rules” that serve as a link between linguistic expressions about objects to actions directed on those objects. This may be extended to other representational modalities as well (e.g. images); perhaps there are rules for generating imagistic representations, along with other rules for responding to them. These rules should govern which causal connections are *sufficient* for reference and which are not, so the veracity of a CTR depends on there being such rules. And if it turns out that there are no such rules, then a CTR won’t work. However, not much progress has been made in producing such a list of rules. So while a CTR might be a good working hypothesis, one shouldn’t place too much emphasis on it until these rules are specified. Even Kripke (pp. 91-97) admits the difficulties here. The moral is that we shouldn’t be too eager to trade one vague notion (i.e. intentionality) for another (i.e. a CTR).

At this point I would like to take a closer look at these causal connections. First I will simply point out the obvious: that not just any causal connection is sufficient for reference. Consider, for example, a woman (let's call her Jane). Certainly there are a great multitude of causal connections between Jane and things that do not represent her, like her twin sister June; we would not say that June is a representation of Jane, and neither is the chair that Jane sat in this morning. So what would be a sufficient condition for something (e.g. a picture) to be a representation of Jane? It cannot be that the causal connection results in the picture looking like her, since similarity is not sufficient for reference. And if one says it is the fact that the picture is *a picture of Jane*, then one is just arguing in a circle. Perhaps it is just the simple fact that we *treat* the picture *as* a picture of Jane. It is nothing about the picture, nor about the way in which it was produced; it is just the way we *respond* to it. So while the fact that something is causally connected to Jane *may* be necessary for it to count as a representation of her, it is *we* who decide, of the things that are causally connected to Jane, which ones count as representations of her and which do not. This is just to say that we take some of these things to be significant in this respect (i.e. pictures) and not others (e.g. chairs). But even here there will be innumerable cases where we are not sure what to say. Is an impression that Jane's body made in the sand a representation of her? What about her shadow? Or the dress she wore yesterday that still retains a semblance of the shape of her body? If you want to say that these things represent her because they resemble her shape, then you have not understood that similarity simply isn't sufficient for reference. And what if she writes her name on a slip of paper – does that represent her? What if she writes the name 'Jane' while thinking of her *cousin* Jane? Does that also represent her?

The point is that causal connections, like mental representations, do not intrinsically determine reference. The causal connections that do seem to be sufficient for reference are

merely the ones that we take to be significant. And there doesn't seem to be any way to make a clear distinction between what is and what is not a representation. Recall what Wittgenstein said about drawing boundaries – this is one of those cases. The truth is that casual connections are important, to *us*. But a picture of Jane isn't a picture of Jane unless we conceive of it as a picture of Jane; otherwise it's just a pattern on a piece of paper. In fact, it isn't even a picture unless we conceive of it as a picture, it's just a *thing*. We give meaning to words and objects through our interactions with them, not the other way around. Discovering the causal history of something, say some “picture” of Jane that resulted from the random discoloration of “defective” photo-paper, can lead us to retract the meaning we first gave to it (i.e. that it was a picture of Jane). The only thing that would change in such a case, upon discovering the causal history of that piece of photo-paper, is the way we think about it. I *could* still think of it as a picture of Jane, albeit one that was produced by random chance. I don't think that this would be a wholly inappropriate use of words, just so long as we were clear about what was meant. Of course, what was meant was that there is no significant causal connection between the picture and Jane. So it seems that, while a causal connection is not sufficient for representation, it *is* necessary; and it is necessary because (and only because) causal connections are profoundly important to *us*. After all, ‘reference’ and ‘representation’ are just words in our language – they needn't correspond to some metaphysical mechanism.

The obvious objection at this point would be that I previously claimed that ‘water’ picks out both H<sub>2</sub>O and XYZ, but have now acknowledged that a causal connection is necessary for “picking out”; thus I have contradicted myself. But this is not quite right. In the first case, I said that intension underdetermines extension, such that a variety of substances *can* fall under the same term. In the second case, I said that if XYZ had occurred in our world, we would use the

term ‘water’ to refer to a disjunction of H<sub>2</sub>O and XYZ. Thus, underdetermination means that it is possible for our terms to be applied to various substances that we encounter, and this is a quite different claim than the one made in the objection. This objection rests on what I will call the Intensional Error: that intensions determine extensions as if they ran out ahead of us and divided all objects into well-defined sets (cf. Wittgenstein pp 64-65; §§ 188-190). In this form it is obviously false (and of course nobody believes it), but it commonly takes the following guise: that our intensions carve out well-defined sets of objects. This leads to two mistakes: first, a confusion of the intension with the set of objects that it would pick out *in use*; and second, that intensions correspond to well-defined sets of objects. In the first case, on the assumption that intension determines extension, there should be a clear distinction between the intension and the set of objects that it would pick out in use. In the second case, as noted above, most of our words simply don’t match up with the world in such a way as to form well-defined sets of objects (e.g. ‘games’). The problem is that, given the way we *think*, it really does seem that all our terms refer to well-defined sets. This is because our intensions have a certain structure, and we innocently mistake *that* structure for the mind-independent structure of the world.

So intensions determine extensions “on the fly”, so to speak, and the term ‘water’ refers to any “watery-ish” substance that we may come across. When we imagine XYZ, we are to imagine some watery-ish substance that has a chemical structure other than that of H<sub>2</sub>O. I claim that the term ‘water’ would refer to XYZ if we ever *did* come into contact with it. Reference is about the way we use words, not some metaphysical connection with well-defined sets of objects. And it would be no objection whatsoever to claim that we will never come across XYZ since it does not exist in our world, and therefore the term ‘water’ never has and never will apply to it. If this

objection did carry any weight, then there is no causal connection whereby the externalist can refer to XYZ; thus his argument is at best meaningless.

Our inquiry here is more about the way we *think* about the world, not the way the world *is*. Unfortunately, I don't think a decisive argument can be given for the point I am trying to make. But it seems to me that what philosophers have done is to project their own thought processes onto the world, and they are now searching for the metaphysical correlates of those processes. They are looking "out there" for something that is "in here". I am not claiming that there is no such thing as reference and representation; the only thing I deny is that they have metaphysical correlates. So in the remainder of this paper I would like to present my view and show why Putnam's PAE fails to uproot it.

Putnam makes an analogy between pictures and mental images to extend his argument against the intrinsic representation of one object by another (e.g. a picture of x) to mental content (e.g. a mental image of x), but his analogy makes mental images sound epiphenomenal, as if they were mere accompaniments to cognition. Rather, they seem to be an active and integral part of thinking. All things considered, I think it best to avoid such pictures-in-the-head talk in favor of *ways of conceiving*, or *modes of thought*. One way of thinking about something is imagistically, as when I think of Jane's face. Another is acoustically, as when I think of her voice. But it doesn't make any sense to say that, when I think of Jane's face, I generate a mental image that I then take to be "about" Jane's face. I don't believe that Putnam's analogy holds, because a picture only represents something else because we think of it that way. When I look at a picture of Jane, for example, I think of it *as* a picture *of* Jane. But when I think about Jane, I do not think of Jane and then think of my thought as a thought of Jane – I just think of Jane. Now, as to what my "Jane" thoughts refer to, as in *which external object am I thinking about*, I believe that the

question is ill-conceived. I will return to this point shortly, but for now, I will simply provide an *alternative* to Putnam's view.

Putnam (1981) observes that "If there are mental representations that necessarily refer (to external things) they must be of the nature of concepts and not of the nature of images." (p. 17) But he goes on to say that "Concepts...are not mental presentations at all. Concepts are signs used in a certain way." (p. 18) But here he seems to be confusing the terms 'concept' and 'word'. A *word* is a sign – a concept *determines* the way in which a word is used. So when he notes that "signs do not themselves intrinsically refer", he is right about words, and wrong about concepts. (p. 18) Putnam's objection to this is that his concepts of elm trees and beech trees are identical. Therefore, the fact that they refer to different things cannot be explained by a difference in his concepts. But the truth is that he has no ability to "refer" to elm trees any more than he has the ability to "refer" to beech trees. Recall that a causal connection is *not sufficient* for reference; so the bare fact that there is a causal connection between Putnam's 'elm' and, say, some botanist's 'elm', is *not sufficient* for Putnam's 'elm' to refer to anything at all – there must be something more involved, and *Putnam just ain't got it!*

I stated above that, if narrow mental states determine intensions, then presumably they do so according to particular sets of phenomenal characteristics; *that* is the business of concepts, and that is how concepts determine the way a word is used. If Putnam cannot distinguish between two different trees, and apply 'elm' to one and 'beech' to the other, then his concepts "elm" and "beech" consist entirely of the terms 'elm' and 'beech' (thus, contra Putnam, these concepts are not completely identical), along with his concept of "tree" (which is something they have in common). He could, however, use these terms at his local neighborhood nursery (where they sell trees), or the library (where there are books about trees), or the Biology Department at Harvard

University (where there are experts on trees). At these places he will find people who have different ways of using these terms, and he can learn new uses for them. So what, then, is the nature of concepts?

As a first attempt to understand what concepts are, I will provide a comparison between them and neural structures. Concepts do not correspond to discrete neural structures; rather, they correspond to the interconnectivity and weight of those neural connections that dispose neurons to activate in response to certain inputs, such that they produce certain outputs. Input can come from the sense organs or from elsewhere in the neural net. Likewise, output can be to the various motor organs or elsewhere in the neural net. Think of it like a simple parallel distributed network that can discriminate between two different objects (e.g. mines and rocks) based on “sensory” input. The connectivity and weight assigned to the “neural” architecture provides it with two different concepts (i.e. “mine” and “rock”).<sup>2</sup> This architecture includes shared pathways for sensory input (sonar frequencies), and discrete pathways for outputs (one corresponding to ‘mine’ and another to ‘rock’). In-between these is a network of pathways, many of which are common to both output states, while others are not. So there is no discrete physical structure corresponding to any one individual concept, but they can be modeled mathematically according to their relationships to inputs and outputs (which is *not* to say that they are mathematical in nature). Corresponding to this neural structure we have inputs (in the form of perception), outputs (in the form of usage), and concepts (what comes in-between). This provides a simple model for what concepts are. For example, the neural correlate of my concept of “Jane” is distributed throughout my brain, and when my brain is *in* that state, it corresponds to a mental state *about* Jane. Earlier I said that the question, what do my “Jane” thoughts refer to, is ill-

---

<sup>2</sup> I should point out, however, that the concepts “mine” and “rock” of this simple network are not the same as our concepts “mine” and “rock”, because we have different perceptual mechanisms, and these concepts have relations to other concepts (e.g. “ship”, “explosion”, “crew”, “death”, etc.) that the simple network doesn’t have.

conceived. The truth is I don't think that there is any deep fact of the matter here; there are just our *perceptions* of the external world, the way we *think* about those perceptions, and the way we *act* within that world. Perhaps every theory has a point where it begins to sound implausible, and this would take a lot more explaining than I can do in such a short paper. Nevertheless, at least I am kind enough to make this explicit, and even go so far as to direct the reader's attention toward it, rather than trying to bury it. So I must simply beg the reader's pardon while I sidestep this issue.

It is a further question of how phenomenal properties are correlated with neural architecture – a question that is again beyond the scope of this paper. And how this all works is going to depend on one's metaphysical views. A physicalist might say that such concepts are epiphenomenal, but I don't think so. I think that reason points us to some sort of dual-aspect view, or neutral monism. On my view, concepts stand on the boundary between the physical and the mental, and an adequate theory of concepts would be at least as hard to provide as an adequate solution to the mind/body problem. It may even be logically impossible, I am uncertain on this point. I don't think that we can apply language to concepts, because that would be to form a general concept of concepts, which would be on a par with trying to state the general form of a proposition in the form of a proposition. As Wittgenstein found, "One thinks that one is tracing the outline of the thing's nature over and over again, and one is merely tracing round the frame through which we look at it." (2001, p. 41, § 114) Nevertheless, we can at least begin to understand the source of our difficulty.

Note that concepts, as I have defined them, are *intrinsic*. Further, there is no right or wrong way to apply them – they simply apply to whatever it is that they apply to. So a concept is in-itself just a mechanism that "does what it does". You might think of it as something like a private

rule, although this would be misleading because obeying a rule requires, as Wittgenstein observed, objective criteria (cf. p. 202, § 69) This objective criteria is provided by the *external* behavior of others with whom I communicate. Thus it is the “common behavior” of mankind that makes language possible. But behavior is just an outward manifestation of a concept, since it is part of the business of a concept to generate behavioral output. So while a concept is intrinsic, agreement in behavior provides objective criteria for their application, and we came to have the concepts we have by growing up in our particular linguistic communities. But since there are no discrete identifiable neural structures that correspond to them, we can only attribute concepts to others based upon our interpretations of their behavior. And the possibility of interpreting the behavior of others rests in having concepts by which to make those interpretations. This explanation has no foundation; in regard to intrinsic concepts and external behavior, one does not take priority over the other. There is but one thesis I am advancing here, and it is this: that a concept, as what comes in-between perception and behavior, is both the thought and the mechanism whereby that thought refers.

### References

- Hilary, Putnam. 1975. The Meaning of Meaning. In *Philosophical papers: Volume 2. Mind, Language and Reality*, auth. Hilary Putnam, pp. 215-271. NY: Cambridge University Press.
- 1981. Brains in a Vat. In *Reason, Truth, and History*, auth. Hilary Putnam, pp. 1-21. NY: Cambridge University Press.
- Kripke, Saul. 1980. Naming and Necessity. MA: Harvard University Press.
- Stalnaker, Robert. 1989. On What’s in the Head. *Philosophical Perspectives*, Vol. 3, Philosophy of Mind and Action Theory, pp. 287-316.
- Sorensen, Roy. 2001. *Vagueness and Contradiction*. NY: Oxford University Press.
- Wittgenstein, Ludwig. 2001. *Philosophical Investigations*. MA: Blackwell Publishing Ltd.