

## On Deriving Essentialism from the Theory of Reference

Jussi Haukioja

It is widely agreed that the causal theory of reference for natural kind terms can be used as a central ingredient in an argument that, if successful, shows that theoretical identity statements such as (1) are necessary, if true:

(1) Water is H<sub>2</sub>O.

However, Nathan Salmon (1981) has famously argued that one cannot derive interesting essentialist conclusions *merely* from a causal theory of reference à la Kripke and Putnam. When it comes to natural kinds such as water in particular, Salmon claimed that, to argue for natural kind essentialism, a non-trivial essentialist premiss concerning substances (or species, natural phenomena, etc.) is also needed. My aim in this paper is to challenge this widely accepted claim and sketch, in some detail, a possible way of arguing for natural kind essentialism, which does not require any additional essentialist premisses to be brought in. Rather, this explanation sees essentialism as being already built in to the semantics of natural kind terms.

But first, we should get clear on exactly what kind of an essentialist view is at issue here. Claims such as “Water is necessarily H<sub>2</sub>O” can be read in two importantly different ways. On one reading, they express *de re* necessities between kinds (thought of as *abstracta*), and have the logical form of an identity statement: the kind *water* is identical with the kind *H<sub>2</sub>O*. On the

other, they express the necessity of a quantified biconditional: on this reading (1) expresses the proposition:  $\forall x[x \text{ is (composed of) water} \leftrightarrow x \text{ is (composed of) H}_2\text{O}]$ .

In discussions of natural kind essentialism, the two readings are not always carefully distinguished. Indeed, Kripke (1980) seems to move back and forth between the two kinds of reading, in discussing various examples. Putnam's (1975) natural kind essentialism, on the other hand, clearly holds that statements of the latter, quantified biconditional kind, are metaphysically necessary, when true, and this is the reading Salmon adopts in responding to Putnam. I will too, in this paper, understand natural kind essentialism (NKE) in this sense. As stressed by Mackie (2006, 170), the latter reading is weaker than the former, in some important ways. One may well hold that statements of the latter kind are true, without wanting to include kinds in one's ontology, and thereby without accepting the truth of statements of the former kind. Moreover, NKE in the sense relevant here does not explicitly ascribe essential properties to kinds, let alone members of kinds. What it does, rather, is specify properties that are essential for kind membership: having a certain chemical structure is necessary for belonging in the extension of the term "water".

Here is Salmon's (1981, 166-167) reconstruction of a Putnamian argument for the relevant natural kind essentialist claim (numbering mine):

(2)  $\Box$ [Something is a sample of water iff it is a sample of *dthat* (the same substance that *this* is a sample of)]

(3) *This* (liquid sample) has the chemical structure H<sub>2</sub>O.

(4) Being a sample of the same substance as something consists in having the same chemical structure.

Therefore,

(5) □ (Every sample of water has the chemical structure H<sub>2</sub>O)

While something like (2) is (the *qua* problem aside) plausibly derivable from a causal theory of reference for natural kind terms, and (3) reports the relevant empirical finding, (4) is neither a product of the theory of reference, nor a part of the empirical finding about water, but an independent, non-trivial essentialist premiss. Salmon's argument has gained widespread acceptance: while most causal theorists are happy to accept something like premiss (4), they seem to agree, by and large, that some such non-trivial essentialist premiss is needed in order to argue for essentialism on the basis of semantics.

Occasionally some doubts are expressed about this, however. For example, after having expressed sympathy with Salmon's argument, Alexander Bird and Emma Tobin add the following:

However, one might alternatively take the view that semantics and metaphysics are more intimately related. For example, one might hold that the semantics itself makes certain metaphysical assumptions [...] According to the view that the correct semantics implies an essentialist metaphysics, the semantic claim that 'water' is a rigid designator is itself metaphysically loaded. But this is not to say that Putnam, Kripke, and other philosophers are smuggling in metaphysical assumptions. Rather, it is to say that competent users of the term tacitly make those assumptions. Putnam's Twin Earth thought experiments make those assumptions explicit. (Bird & Tobin 2008, 58)

Bird and Tobin do not, however develop the idea further, or attempt to defend it – they merely mention it as a possible position. My aim in this paper is to develop, in some detail, one way of arguing for such a position. In my view, there is an independently attractive explanation of why statements like (1) are necessary if true, one that does not rely on antecedent non-trivial essentialist premises on the part of the theorist. Rather, on this explanation a certain kind of essentialist preference is already built in to the semantics of natural kind terms, in the way ordinary competent speakers use these terms.

My argument will be based on three assumptions: two theoretical ones, and one empirical one. I have argued for both of the theoretical assumptions in recent work; I will not present the full arguments for these assumptions here, but rather state them in a way that makes them directly relevant to the question at hand, as well as give a condensed version of the arguments supporting them. The empirical assumption, I claim, is quite strongly supported both by everyday experience (including thought experimentation) as well as empirical data from psychology. In effect, I am here arguing for a conditional claim: *if* one accepts my three assumptions – assumptions which I take to be plausible and by no means highly revisionary – *then* a certain kind of derivation of NKE becomes available, and this derivation proceeds without reliance on antecedent essentialist premisses.

### **1. Theoretical Assumption 1: the explanation of necessary a posteriori identity statements**

The derivation (2) – (5) relies implicitly on the notion of rigid designation. The necessity in (2) arises from the assumption that “water” is a rigid designator, designating the same substance in all possible worlds. The criteria for *being* the same substance, in turn, are thought to be independent of the semantics: once the word “water” latches on to an actual sample, the sameness relation described in (4) will, as it were, take care of the rest. The causal theory does not provide the sameness-relation; this needs to be assumed as an additional premiss.

However, there is a problem. Although Kripke and Putnam clearly did take natural kind terms to be rigid designators, and took this feature to play a central role in explaining the necessity of theoretical identifications, they never provided a definition of rigidity for general terms such as “water” (or predicates such as “is [composed of] water”). Kripke only gave one for singular terms. And as is by now widely acknowledged as a result of a fairly voluminous

discussion<sup>1</sup>, there's no obvious and unproblematic way of extending the notion of rigidity from singular terms to general terms.

A seemingly natural way of defining rigid designation for general terms – and one that is often assumed in discussion – would be as follows: a general term is rigid iff it denotes the same property in all possible worlds. However, it is widely recognized that this runs the risk of overgeneralizing, even trivializing the notion of rigidity: not only will natural kind terms such as “water” and “tiger” be rigid according to this account, so will terms such as “bachelor” (denoting bachelorhood in all possible worlds) and “thermometer” (denoting the property of being a thermometer), and so on. Some argue that this is a feature, rather than a bug, of this account. However, I've argued elsewhere (Haukioja 2012) that rigidity, so defined, cannot do the theoretical work that is relevant for my purposes here, namely that of explaining the necessity of a posteriori identity statements.

In a nutshell, the problem is the following. The usual assumption, found already in Kripke, is that identity statements between rigid expressions are necessary, if true. But if we accept the above characterisation of rigidity for general terms, this is obviously *not* the case: we should then expect “all and only renates are cordates” to be necessary (if true) as well, since on the view in question “renate” and “cordate” come out rigid. But “renates are cordates” is not necessary, if true. So, whether or not there are other reasons to adopt the notion of rigidity for general terms under discussion, it should be clear that it is not *this* semantic feature that is doing the work in deriving the essentialist claims that concern us here.<sup>2</sup>

1 Some of the central works are: Soames 2002; Martí 2004; Devitt 2005; LaPorte 2013.

2 Two clarifications are in order. First, (2) – (5) does not explicitly appeal to the rigidity of “H<sub>2</sub>O” (which I think is telling, cf. below), and the derivation is therefore not directly of the sort mentioned here. However, Kripke and Putnam clearly do assume that true identifications between rigid terms should be necessary, and some examples discussed in the literature (such as “Brontosaurus is Apatosaurus”) make explicit use of the assumption. Secondly, kind identities are sometimes (often by proponents of the notion of rigidity under discussion here, cf. LaPorte 2013) construed as identity statements between *abstracta* rather than as quantified biconditionals (cf. above). However, this involves taking kind terms to be *singular* terms denoting kinds, rather than general terms; moreover, on this understanding it is no longer obvious how a kind identity could be a straightforward empirical discovery, since we do not observe kinds, but only their instances (for

What *is* doing the work, then? My answer is as follows. Natural kind terms differ semantically from most other general terms in being *actuality-dependent*. Roughly, a term is actuality-dependent iff some feature of the object or objects in its *actual* extension enters into determining its extension in *non-actual* contexts of interpretation. For example, according to causal theories of reference, “water” is actuality-dependent, since its extension in non-actual worlds is partly determined by the structure of the stuff we *actually* refer to, with our tokens of “water”. “Bachelor”, by contrast, is not actuality-dependent, since its extension in non-actual worlds is not determined on any features of the actual bachelors, but is rather determined by the very same criteria as its *actual* extension. It is the actuality-dependence of “water” that explains the necessity of (1): “H<sub>2</sub>O” merely spells out the nature of the stuff we refer to as “water”, and since “water” is actuality-dependent, it picks out H<sub>2</sub>O in all possible worlds, making (1) necessary.<sup>3</sup> In fact, “H<sub>2</sub>O” is not actuality-dependent.

To add some details to this rough characterisation that will be relevant later on: any referring expression will have an *applicability role* associated with it, be it the satisfaction of a descriptive characterisation, a causal role, or something wholly different. (Such applicability roles need not be *semantically* associated with the relevant expressions.) The applicability role associated with expression E specifies the criteria that an object, substance, etc. has to satisfy in order for E to apply to it. Actuality-dependent expressions can now be defined as those expressions which apply, in non-actual worlds, to the things which possess the properties which *realise* the relevant applicability role in the *actual* world. In other words, an actuality-dependent expression has, as its extension, the things which instantiate the actual realiser property of the applicability role, or, in short its actual *ARR* (Applicability Role Realiser), where the realiser property is not uniquely determined by the applicability role. For example,

---

more on this see Haukioja 2012).

3 See Haukioja 2012 for more details.

when Oscar and Toscar (on Earth and Twin Earth, respectively), produce the string “water”, the applicability roles associated with their terms will be identical, but because of a difference in their environments, different properties will realise that role, and the extensions of their tokens of the string will thereby be different across possible worlds.<sup>4</sup> An actuality-independent expression, in contrast, applies in non-actual worlds on the basis of the very same applicability role as in the actual world.

The derivation of NKE on the basis of actuality-dependence is, to a first approximation, the following:

(6)  $\square$  (Something is a sample of water iff it instantiates the actual ARR of "water")

(7) The actual ARR of "water" is the chemical structure  $H_2O$

Therefore,

(5)  $\square$  (Every sample of water has the chemical structure  $H_2O$ )

(6) follows from the actuality-dependence of “water”. If (7) is properly described as a straightforward empirical discovery, this derivation of NKE will not require any antecedent essentialist premises. But for all that has been said so far, the following worry remains. If the ARR of “water” has precise enough content such that (7) is properly described as a straightforward empirical discovery, can we plausibly claim that (6) follows from the semantics? For example: if the ARR of “water” is something like *the chemical structure shared by the stuff in the actual extension of “water”*, (7) is surely an empirical discovery free of hidden essentialism, but (6) simply does not follow from the semantics. To address this worry, we will need to look at what it is that *makes* an expression actuality-dependent, and

---

<sup>4</sup> I remain neutral here with respect to questions of how the applicability role is represented by the speaker (if at all), and if it is represented, whether it deserves to be thought of as narrow content.

what determines *how* the extension of an actuality-dependent expression is determined in non-actual worlds.

## 2. Theoretical Assumption 2: determination of reference

What determines what our terms refer to in general, and what they refer to in non-actual worlds in particular? Here, it is important to note that such questions are ambiguous between two quite different readings. The role of internal vs. external factors can be raised on both readings, and accordingly two different internalism/externalism distinctions can be made, and we should take care to distinguish between them.<sup>5</sup>

First, note that actuality-dependent expressions have, in a very familiar sense, an externalist semantics. If “water” is actuality-dependent, this entails that its reference is, at least in non-actual worlds, determined by factors external to the speaker, namely by the actual ARR of “water”. By contrast, actuality-independent expressions can have an internalist semantics. (But they need not – this will depend on whether the relevant applicability role is internally determined. For example, on a strong social externalist view the applicability role of “bachelor” may be partly externally determined, although the term is actuality-independent). This internalism/externalism distinction is concerned with the question of whether or not the criteria for the applicability of an expression are determined by the speaker's individual psychology; let us call these views *first-order* internalism and externalism (cf. Cohnitz & Haukioja 2013).

But another question concerns what makes it the case *that* a given expression has the kind of semantics that it has. What makes it the case, for example, that “water” is actuality-dependent (supposing that it is), and thereby has first-order externalist semantics, while

---

5 The two distinctions are elaborated in much more detail in Cohnitz & Haukioja (2013), although the connection to actuality-dependence is not explicitly made there.



“bachelor” is actuality-independent? Again, an internalist and an externalist answer is possible. According to *meta-externalism*, the fact that expressions have the semantics that they do is, at least partly, determined by factors external to the speaker's individual psychology, whereas according to *meta-internalism*, such facts are determined by facts internal to the speaker in question.

In Cohnitz & Haukioja (2013) we argue at length against meta-externalism, on the basis that a meta-externalist position fails to make sense of the role of reference, and theories of reference, in explaining successful linguistic communication. We also raise a suspicion that, because the two internalism/externalism distinctions have not been distinguished carefully enough, many have illicitly taken our evidence for first-order externalism to also count as evidence for meta-externalism.

The alternative meta-internalist view allows that the extensions of at least some expressions are partly externally determined – it is consistent with first-order externalism – but holds that the fact *that* this is the case is determined by something to do with the speaker's individual psychology. Meta-internalism, as such, leaves room for a variety of different views on *which* internal features determine the semantics of a speaker's terms. Our preferred version is a dispositionalist one. What makes it the case, in general, that a given expression is to be given a first-order internalist or externalist semantics has to do with our<sup>6</sup> *patterns of application and interpretation* (including our dispositions to apply and interpret our expressions in conditions that are non-actual, as well as our dispositions to revise our application and expression in response to empirical information, cf. below). For some expressions, such as “bachelor”, our patterns and dispositions of application and interpretation

---

6 I will write, in plural, of *our* dispositions, since I believe such dispositions are, as a matter of fact, widely shared by actual speakers. However, strictly speaking a meta-internalist view will claim that the semantics of a speaker's expressions, including whether and how the reference of the speaker's terms is dependent on his/her physical *and* social environment, is wholly determined by the individual psychology of the speaker.

are unaffected by contingent features of the actual world around us—the properties associated with the expression are taken to be decisive for the question of whether the expression applies to a given individual or not. With other expressions we have dispositions to 'shift the burden' of determining their applicability partly to external factors. In the case of “water”, for example, we have dispositions to evaluate the correctness of actual and counterfactual applications of the term according to whether or not the term is applied to samples which share the underlying structure of the substance that is causally connected in the appropriate way to our actual usage of “water” (or something similar: the details will depend on one's preferred theory of reference). We also have dispositions to *re-evaluate* our application of such terms in the face of new empirical information about what the world is like. This is what makes “water” actuality-dependent.

To illustrate, suppose that we took a representative sample of bachelors, studied them empirically, and found out that every single one of them has a certain neural structure – call it N – while no individuals outside the sample have N. Would we then start to categorise people as bachelors in other possible worlds according to whether or not they have neural structure N or not? Surely not. That is just not how “bachelor” works: we would go on categorising people as bachelors, in the actual world as well as in counterfactual ones, according to their age, gender, and marital status.<sup>7</sup> “Bachelor”, unlike “water”, has no aspirations to denote a natural kind.

---

<sup>7</sup> Or, consider another version: suppose that we found that, say, 99.7% of unmarried adult males have N, while 0.3% do not; at the same time, we find N present in a tiny proportion of married females. Would we revise our categorisation of this small minority of men as bachelors and instead include the married females? Surely not. (Thanks to NN for this twist on the example). Were we to make a similar discovery about the golden stuff, say that we find some of the metal we categorised as gold to have a different atomic number than 99.7% of the rest, while a small sample of a greenish looking metal turns out to have the same atomic number as the other 99.7%, we would, I think, revise our categorisation of the 0.3% and consider the greenish stuff as gold.

If what I've just said is roughly correct, then the meanings of actuality-dependent terms include, as it were, a placeholder to be filled in by the actual world, and this placeholder exists because of our dispositions to use these terms. The question of which expressions, if any, *are* actuality-dependent is then ultimately an *empirical* question concerning our linguistic dispositions. Likewise, the question of what kinds of things are eligible to fill this placeholder, *i.e.* what the applicability roles, and thereby the ARR of our actuality-dependent terms are,<sup>8</sup> will be determined by our dispositions.

### **3. Empirical Assumption: essentialist dispositions**

Do we, as a matter of fact, have the kinds of burden-shifting dispositions mentioned above, with respect to natural kind terms? Everyday experience, at least everyday experience about philosophers, indicates that we do. The kinds of intuitions we, for the most part, report having in response to Twin Earth-style thought experiments are precisely of this kind. But of course such intuitions can be, and have been, denied: one may worry about how representative a sample philosophers are here. Furthermore, one might doubt whether it really is reasonable to assume that ordinary speakers have the quite complex kinds of dispositions that are required for my view to work.

However, a considerable amount of relevant empirical evidence exists in psychology, under the heading of “psychological essentialism” (for summaries, see Gelman 2004; Gelman & Ware 2012). This is the hypothesis that humans generally believe and assume that natural categories such as substances and biological species have an underlying reality that cannot be directly observed, and that is causally responsible for the outward characteristics and shared features of the members or instances of that category. This hypothesis is well supported by a range of empirical data, across cultures and educational backgrounds – indeed, even subjects

---

<sup>8</sup> More precisely: what the ARRs of our actuality-dependent terms are, *given* a way the actual world is like.

with no formal education and no familiarity with modern science share such beliefs. Moreover, the dominant version of psychological essentialism is *placeholder* essentialism, according to which humans believe *that* categories possess essences without knowing *what* those essences are (Medin & Ortony 1989).

It should be obvious that psychological essentialism (especially of the placeholder variety) fits extremely well with the view I've been developing here. It can, of course, be questioned whether psychological essentialism, as a theory that posits essentialist beliefs in children and adults, is the best, or sufficient, *psychological* theory for explaining the relevant data (cf. Malt 1994; Strevens 2000). What matters most for my purposes here is that the studies supporting psychological essentialism are *categorisation* studies, where subjects are typically invited to make *verbal* judgements about various scenarios. As such, these studies provide indirect evidence for the view that natural kind terms are actuality-dependent: we are, in our usage of natural kind terms, guided by a tacit assumption about underlying essences. Moreover, the studies give us an indication on what the ARR's for our natural kind terms are: they are, at least roughly, those properties, whichever they are, which are actually causally responsible for the shared observable characteristics of members of the relevant natural kind.

These studies do not provide evidence for the kinds of *corrective* dispositions that were mentioned above: dispositions to re-evaluate and possibly reverse one's applications of natural kind terms in the face of new empirical evidence about what the actual world is like. However, a study I was involved in (Jylkkä, Railo & Haukioja 2009) did look at precisely these kinds of dispositions, in non-philosophers, and found that lay speakers do, by and large, have the kinds of dispositions which would make natural kind terms actuality-dependent.

#### 4. Putting the assumptions together

If all three assumptions are accepted, we get an explanation of why identifications such as (1) are necessary if true. We saw earlier that, given my first theoretical assumption about actuality-dependence, we get a derivation of NKE as follows:

(6)  $\square$  (Something is a sample of water iff it instantiates the actual ARR of "water")

(7) The actual ARR of "water" is the chemical structure  $H_2O$

Therefore,

(5)  $\square$  (Every sample of water has the chemical structure  $H_2O$ )

Now, given the empirical assumption and the second theoretical assumption, we can safely assume that natural kind terms such as "water" are actuality-dependent. Moreover, we get a more precise specification of the ARR of "water": ordinary speakers using the term have dispositions that make it the case that that "water" applies, in non-actual worlds, to substances on the basis of whether or not a substance in that world possesses those underlying properties which are, in the actual world, causally responsible for the observable characteristics of water.

In other words, we get the following derivation of NKE:

(8)  $\square$  (Something is a sample of water iff it possesses that underlying property which is actually causally responsible for the observable characteristics of the stuff in the actual extension of "water")

(9) The underlying property which is actually causally responsible for the observable characteristics of the stuff in the actual extension of "water" is the chemical structure  $H_2O$

Therefore,

(5)  $\square$  (Every sample of water has the chemical structure  $H_2O$ )

(8) follows from the theory of reference, if my three assumptions are granted. No additional premisses concerning essences of kinds are needed for the derivation to go through; (8) only relies on an independent assumption about the criteria of sameness for *properties*. (9), on the other hand, is a straightforward empirical discovery about the actual world. Here we have, then, a derivation of natural kind essentialism that requires no antecedent essentialist premisses. The explanation does, to be sure, rely on essentialist dispositions or assumptions, on the part of *ordinary speakers*: this is part of what makes “water” actuality-dependent, and thereby makes (8) true.<sup>9</sup> But, crucially, the explanation does not rely on any essentialist assumptions that we, *as theorists*, have to accept as true in order for the derivation to go through. Rather, (8) is due to the essentialist assumptions ordinary speakers have, in their use of “water” – assumptions which become apparent in our judgements about Twin Earth-style scenarios.<sup>10</sup> Given that the actual world co-operates with such assumptions sufficiently to make some theoretical identification statement such as (1) true, the identification statement will then be true across possible worlds, delivering NKE.

---

9 In an interesting recent paper, Sarah-Jane Leslie (2013) argues that the kinds of intuitions elicited by Twin Earth-style thought experiments are, indeed, reflections of psychological essentialism, but that we should take them to be systematically false, given that findings in natural sciences such as biology and chemistry seem to conflict with such essentialism. We should note, however, that the kinds of essentialist intuitions we have (and which are manifested in thought experimentation) are remarkably flexible. “ $H_2O$ ” is, in Putnam's discussion, just as much a placeholder as “c-fibres” is in Kripke's – indeed, Putnam considers the epistemic possibility of “water” turning out to be a jade-like disjunctive kind. We have, in our ordinary usage of natural kind terms, a strong preference for unified underlying essences, but are willing to settle for something far less unified and even disjunctive, if that is the most that the actual world gives us. As long as we, as speakers, are disposed to let an underlying property determine the modal profile of a natural kind term, that property qualifies as a possible essence.

10 Note that such essentialist assumptions need not be (and in all likelihood are not) very specific at all, and definitely nothing as specific as Salmon's (4). Rather, the essentialist assumption is merely a preference for some explanatory unity underlying the observable characteristics of members of the kind (cf. previous footnote). The precise nature of the explanatory unity – whether it has to do with chemical structure or something completely different – is an empirical matter.

## 5. Conclusion

One may feel that the kind of essentialism delivered by my derivation falls short of “real” essentialism. I sympathise with this: indeed, I noted early on that the kind of NKE I have here discussed is a relatively weak claim. This is, I think, the most that one can get, by Twin Earth-style argumentation – a stronger natural kind essentialist view about kinds as *abstracta* would need to be argued for on different grounds.<sup>11</sup>

If an argument for a stronger essentialist view can be given, this is not in conflict with my view. If such an argument is *not* available, then we should adopt a *conferralist* view (cf. Sveinsdottir 2008, Sveinsdottir 2013), according to which our practices and assumptions confer essentiality on certain properties. There is an element of anti-realism in such a view, but it is important to note that none of the *properties* that are taken as essential are conferred properties; it is only the *essentiality* of these properties that is conferred. The things in the world have the properties they do quite independently of us, but the fact that some of these count as *essences* in the sense relevant here is due to our semantic dispositions and intentions.<sup>12</sup>

## References

- Bird, A. & Tobin, E. (2008). Natural Kinds. Stanford Encyclopedia of Philosophy. <http://plato.stanford.edu/entries/natural-kinds/>. Accessed 22 July 2014.
- Cohnitz, D. & Haukioja, J. (2013): Meta-Externalism vs. Meta-Internalism in the Study of Reference. *Australasian Journal of Philosophy* 91, 475-500.
- Devitt, M. (2005). Rigid Application. *Philosophical Studies* 125(2), 139-165.

---

11 For an interesting recent discussion of what such an argument would have to look like, see Tahko (forthcoming).

12 Earlier versions of this material were presented at conferences, seminars and workshops in Athens, Canberra, Dunedin, Auckland, Helsinki, and Tampere. I am grateful to audiences at these events for useful feedback. I also want to thank Daniel Cohnitz, Markku Keinänen, Erick Llamas, Daniel Nolan, Alexander Sandgren, and Tuomas Tahko, as well as an anonymous reviewer for the journal, for helpful comments on earlier written versions.

- Gelman, S. (2004). Psychological Essentialism in Children. *Trends in Cognitive Science* 8(9), 404-409.
- Gelman, S. & Ware, E. (2012). Conceptual Development: The Case of Essentialism. In S. Margolis, S. Stich & R. Samuels (Eds.) *Oxford Handbook of Philosophy of Cognitive Science* (pp. 454-479). New York: Oxford University Press.
- Haukioja, J. (2012). Rigidity and Actuality-Dependence. *Philosophical Studies* 157, 399-410.
- Jylkkä, J., Railo, H. & Haukioja, J. (2009). Psychological Essentialism and Semantic Externalism: Evidence for Externalism in Lay Speakers' Language Use. *Philosophical Psychology* 22, 37-60.
- Kripke, S. (1980). *Naming and Necessity*. Cambridge: Harvard University Press.
- LaPorte, J. (2013). *Rigid Designation and Theoretical Identities*. New York: Oxford University Press.
- Leslie, S.-J. (2013). Essence and Natural Kinds: When Science Meets Preschooler Intuition. *Oxford Studies in Epistemology* 4, 108-166.
- Mackie, P. (2006). *How Things Might Have Been*. New York: Oxford University Press.
- Malt, B. (1994). Water is not H<sub>2</sub>O. *Cognitive Psychology* 27, 41-70.
- Martí, G. (2004). Rigidity and General Terms. *Proceedings of the Aristotelian Society* 104, 129-146.
- Medin, D. & Ortony, A. (1989). Psychological Essentialism. In S. Vosniadou (Ed.), *Similarity and Analogical Reasoning* (pp. 179-196). Cambridge: Cambridge University Press.
- Putnam, H. (1975). The Meaning of 'Meaning'. In *Philosophical Papers vol. 2: Mind, Language, and Reality*. Cambridge: Cambridge University Press.
- Salmon, N. (1981). *Reference and Essence*. Princeton: Princeton University Press.
- Soames, S. (2002). *Beyond Rigidity: The Unfinished Agenda of Naming and Necessity*. New York: Oxford University Press.
- Strevens, M. (2000). The Essentialist Aspect of Naive Theories. *Cognition* 74(2), 149-175.
- Sveinsdottir, A. (2008). Essentiality Conferred. *Philosophical Studies* 140(1), 135-148.
- Sveinsdottir, A. (2013). Knowledge of Essence: The Conferralist Story. *Philosophical Studies* 166(1), 21-32.
- Tahko, T. (forthcoming). Natural Kind Essentialism Revisited. *Mind*.