

Christian Nimtz 2010  
Universität Erlangen

published: *Christian Nimtz: Thought Experiments as Exercises in Conceptual Analysis*,  
in: *Grazer Philosophische Studien 81*, 2010,  
189–214.

PHILOSOPHICAL THOUGHT EXPERIMENTS  
AS EXERCISES IN CONCEPTUAL ANALYSIS

Christian NIMTZ  
University of Erlangen-Nürnberg

*Summary*

In the paper, I defend the viability and importance of conceptual analysis to philosophical inquiry. My argument proceeds in two steps. In a first step, I argue that we rely on the notions guiding how we do and would apply our terms in order to evaluate the counterfactual conditionals we find at the heart of philosophical thought-experiments. In a second step, I argue that our notions determine what the relevant terms mean in our mouth. In order to defend the resulting neo-descriptivist semantics, I put forth an epistemic argument *for* descriptivism – the argument from communication. I conclude that philosophical thought experiments are exercises in conceptual analysis.

1. *Thought experiments and the fate of armchair philosophy*

Thought experiments<sup>1</sup> are widely acknowledged to be essential to our armchair ways of doing philosophy. It is hardly surprising, therefore, that philosophical thought experiments take centre stage in the current debate about the viability of armchair philosophy. (See e.g. Jackson, this volume, Williamson 2007, Nimtz 2007 and the papers in DePaul/Ramsey 1998,

---

<sup>1</sup>. This choice of label is not meant to imply that thought experiments literally are *experiments*, i.e. “controlled manipulations of events, designed to produce observations suited to confirm or disconfirm one or more rival theories or hypotheses” (Blackburn 1994, 131).

Gendler/Hawthorne 2002, and Knobe/Nichols 2008.) In this paper, I defend an epistemology of philosophical thought-experimenting. What entitles us to hold true the counterfactuals on which philosophical thought experiments pivot is tacit semantic knowledge we possess, or so I argue. On the account here defended, philosophical thought-experimenting is a priori in the sense in which determining what our words mean by reflection on their application is an a priori exercise.

My argument combines two lines of thought. The first is epistemic and focuses on the evaluation of counterfactuals we find at the heart of philosophical thought experiments. Having argued that philosophical thought experiments are best understood as pivoting on counterfactual conditionals (§§2–3), I argue that these counterfactuals – which I call *philosophical counterfactuals* – are epistemically distinctive. In contrast to mundane or scientific counterfactuals, our entitlement to hold philosophical counterfactuals true does not derive from empirical evidence, or from empirical theory (§§4–5). The second line of thought is semantic and focuses on how the conditions tacitly guiding how we do and would apply our terms – I sum these conditions up as *notions* – shape the semantic properties our terms have in our mouths. Having argued that we, in all plausibility, evaluate philosophical counterfactuals drawing on our notions (§6), I propose and defend a neo-descriptivist semantics (§§7–8). I explain that the key contention of neo-descriptivism consists in the idea that notions determine semantic properties, and I put forth the *argument from communication* in order to support the neo-descriptivism and reject the rival Kripkean externalism. Combining both lines of thought, I conclude that philosophical thought experiments are exercises in conceptual analysis (§9).

The account here presented runs counter to the ambitious rationalism espoused by Bealer (2000, 2002) or Katz (1998, 2002), who hold that we need a priori insight beyond mere semantic knowledge for philosophical thought-experimenting to work. It also puts me at variance with the armchair empiricism of Williamson (2007), who maintains that a priori knowledge plays no distinctive role in thought-experiments. I won't have much to say about ambitious rationalism. I will, however, take issue with Williamson's armchair empiricism (see §5).

## 2. *Thought experiments and counterfactual thinking*

Thought experiments are ubiquitous in philosophy. Paradigm examples include Davidson's Swampman (Davidson 1987), Gettier cases (Gettier 1963, Lehrer 1965), Putnam's Voyages to Twin Earth (Putnam 1975), and Jackson's Mary-case (Jackson 2004b). These thought experiments share some distinctive features. They all comprise a *narrative core* – a story inviting us to consider a possible situation where something happens, typically to some subject. In telltale fashion, Lehrer's case begins thus: "Imagine the following. I see two men enter my office whom I know to be Mr. Nogot and Mr. Havit (...)" (Lehrer 1965, 169). On this basis, we are expected to arrive at a specific upshot best read as being modal. In Lehrer's case, we are expected to conclude that our subject is in a possible situation where he has a justified true belief that someone in his office owns a Ford, but does not know this. All the thought-experiments listed are truly epistemic in that they aim to warrant (rather than just instil) such an upshot. They are not mere 'intuition pumps' in Dennett's (1995) sense, i.e. quasi-mechanical devices designed to inculcate convictions.

Thinkers devising philosophical thought experiments employ their upshots to draw general conclusions – say, that the JTB-analysis of knowledge is false, or that physicalism is flawed. Contrary to Häggqvist (2009), these conclusions are best considered extraneous to the thought experiment devised to establish them. Whereas a thought experiment fails if it does not establish its intended upshot, it can be considered sound even if the philosophical conclusion does not follow. Philosophers typically agree that Jackson's Mary-case establishes that a physically omniscient person inhabiting a world like ours could still learn something. Still, many of them deny that physicalism fails (see the contributions to Ludlow/Stoljar/Nagasawa 2004). This is easily explained: Jackson's thought-experiment is only part of his more encompassing knowledge argument, and it is the part that apparently works.

Our survey of paradigmatic philosophical thought experiments encourages the following general picture: In philosophical thought experiments, we are expected to come to the conclusion that a certain modal claim is true on the basis of considering a possible situation described in the experiment's narrative core. Going by this finding and taking our cue from Williamson (2007, ch. 6, 2004), we may understand thought experiments

to (paradigmatically) consist of three steps. In a first step, it is contended that a certain situation  $C$ , described by the experiment's narrative core and typically comprising a subject or other, is possible. Let me represent this *possibility premise* thus:

$$(1) \quad \Diamond(\exists s: s \text{ is in } C)^2$$

In Jackson's case, it is claimed that there could be some physically omniscient person by the name of 'Mary' who has never perceived something non black-or-white. On the basis of this possibility premise, we are expected to hold that our protagonist has some property  $F$  (that has not already been explicitly mentioned in characterizing  $C$ ). Contemplating Jackson's Mary, we conclude that she would learn something on seeing something red for the first time. Purging the anaphoric element for ease of exposition,<sup>3</sup> we can represent this *modal upshot* thus:

$$(2) \quad \Diamond(\exists s: s \text{ is } F)$$

How do we get from a thought experiment's possibility premise to its modal upshot? The required link needs to be modal, since it is employed to establish a possibility on the basis of another possibility. A natural as well as popular<sup>4</sup> idea for the modal link licensing us to hold (2) on the basis of (1) is the counterfactual statement that if some subject  $s$  *had* been in situation  $C$ , some subject *would* have property  $F$ :

$$(3) \quad \exists s: s \text{ is in } C \Box \rightarrow \exists s: s \text{ is } F$$

Read along these lines, Jackson's Mary case and Lehrer's Nogot/Havit-scenario are understood to pivot on the following counterfactuals:

---

<sup>2</sup>. Williamson assumes that the modality in question is metaphysical. I do not want to commit to that. Please read ' $\Diamond p$ ' and ' $p \Box \rightarrow q$ ' throughout as 'It is possible, that  $p$ ' and 'If  $p$  had been the case, then  $q$  would have been the case', respectively, with an unspecific modality.

<sup>3</sup>. If you want our representation to capture the intersentential anaphoric reference in "Imagine Mary to be in situation  $C$ . *She* would learn something", you could employ Berger's (2002) technique of indexed variables and represent it thus:  $\Diamond(\exists s^1: s^1 \text{ is in } C). \Diamond(s^1 \text{ is } F)$ .

<sup>4</sup>. See e.g. Cooper 2005, 329, Norton 2004, and Quine 1995, 98.

- (4) Some subject  $s$  is in a Mary type-situation  $\Box \rightarrow$  Some subject  $s$  is physically omniscient, yet learns something on seeing something red for the first time.
- (5) Some subject  $s$  is in a Nogot/Havit-situation  $\Box \rightarrow$  Some subject  $s$  has justified true belief that someone in her office owns a Ford, but no knowledge.

In order to have a convenient label, I call such counterfactuals we find at the heart of philosophical thought experiments *philosophical counterfactuals*.

The analysis offered takes philosophical thought experiments to be exercises in counterfactual thinking. In a philosophical thought experiment, we reason from a possibility premise to a modal upshot on the strength of a modal link provided by a philosophical counterfactual.<sup>5</sup> A survey of possible alternatives lends support to this account. On the one hand, one could point out that we apparently arrive at the modal upshot on the basis of the possibility premise without any discernible mediating step, and conclude that we should drop the modal link (3) from our analysis.<sup>6</sup> However, to do so would mean to misunderstand the task an analysis is set to accomplish. It does not aim to retrace the sequence of psychological steps we find ourselves making in thought-experimenting. It rather aims to uncover the (apparent) justificatory structure in thought experiments. We clearly feel entitled to an experiment's modal upshot because we judge it to follow from the possibility premise in a specific way. And a convincing way to capture this is to identify the modal link involved with a counterfactual.

On the other hand, one could point out that the strict implication

- (6)  $\Box(\text{Some subject } s \text{ is in } C \rightarrow \text{some subject } s \text{ is } F)$

would also allow us to infer the modal upshot from the possibility premise. But (6) arguably is too strong (see Williamson 2007, 185). The narrative core won't be specific enough to disallow all  $C$ -worlds where bizarre

---

<sup>5</sup>. I am prepared to argue that the thought experiments in the sciences can be analysed along similar lines.

<sup>6</sup>. I take Malmgren to take this position in her widely circulated, though still unpublished paper *Staying in the Armchair*.

going-ons prevent our subject to become *F*. Consider a world where, by a freakish law of nature, anyone seeing a colour evaporates milliseconds later. Mary would be hard pressed to learn something on seeing something red in such a world. Advocates of Jackson's case are unmoved by such scenarios, and rightly so. Everyone agrees that in evaluating thought-experiments, we may properly ignore such remote possibilities. Hence, the modal link cannot be a strict necessity such as (6).

A similar argument threatens to undermine the idea that the modal link is counterfactual.<sup>7</sup> I call it the *challenge from unintended instances*. Suppose that John actually satisfies the description of the subject as provided by the narrative core of the Nogot/Havit-case, but rightly distrusts his beliefs about who owns what car. (Say, he knows that he is frightfully forgetful about these things.) Clearly, John does not have the justified true belief that someone in his office owns a Ford. Since the situation described – call it the 'John-plot' – is by stipulation actual, the counterfactual (5) is false. For although the John-plot satisfies the counterfactual's antecedent, it fails to satisfy its consequent and, given centring,  $A \ \& \ \neg B$  entails that  $A \ \Box \rightarrow B$  is false. Hence, if the Nogot/Havit-case pivots on a counterfactual, learning that the John-plot is actual should make us forego the thought experiment's modal upshot. But it doesn't. In fact, we clearly feel that John's predicament, close to home though it is, is as irrelevant to the thought-experiment as some fancy going-on in a remote world. Do we therefore have to conclude that the modal link we rely on isn't a counterfactual after all?

We do not. A thought-experiment's narrative core will inevitably underspecify the intended situation. So whenever we are confronted with a situation satisfying the letter of a narrative core, we still need to determine whether it is a relevant instance, or whether it may be properly ignored. The John-plot arguably falls into the latter category. As a Gettier case, Lehrer's Nogot/Havit-thought experiment is exclusively concerned with scenarios where a subject will arrive at a justified belief, given that this epistemic principle as espoused by Gettier holds true (and given that one may justifiably believe falsehoods):

---

<sup>7</sup>. This objection has been made popular by Malmgren. See last footnote.

[F]or any proposition P, if S is justified in believing P, and P entails Q, and S deduces Q from P and accepts Q as a result of this deduction, then S is justified in believing Q. (Gettier 1963, 121).

The John-plot does not belong to these scenarios. John's belief will therefore be unjustified even if Gettier's principle holds true. The John-plot thus is an irrelevant instance of the Nogot/Havit-case's narrative core. Hence, (5) may well be true even if the John-plot happens to be actual.

There is a more general way to accommodate the challenge from unintended instances. We have seen that we do not rely on strict implications to arrive at modal upshots from possibility premises. We rather draw on some restricted necessary

(7)  $\Box_X(\text{Some subject } s \text{ is in } C \rightarrow \text{some subject } s \text{ is } F)$

with 'X' marking the restriction to worlds that fit with how we intend the description of *C* provided by the narrative core to be read, a restriction varying with the thought experiment in question. This fits well with the idea that what provides the link is a counterfactual; for on the standard Lewisian analysis, counterfactuals just are variable strict conditionals (see Lewis 1973\*). But we can take (7) as a proposal in its own right. Since this proposal nicely avoids the challenge from unintended instances, we might want to hold that our modal link is strictly speaking a restricted necessity, rather than a bona fide counterfactual. I do not mind. For none of this affects the arguments to come. Although I phrase them in the more familiar terms of 'counterfactuals', they equally apply to restricted necessities.

#### 4. *Counterfactuals and their epistemology*

The general recipe for evaluating a counterfactual seems simple enough. As Williamson puts it, "one supposes the antecedent and develops the supposition (...)" (2007, 153, see also Edgington 2008) in order to see whether one arrives at the consequent. However, how we judge counterfactuals to be true is one thing. What entitles us to make such judgments is quite another. I venture that we can ascertain the latter by asking how we justify such judgments. More precisely, we can ascertain what entitles us

to hold true a counterfactual  $A \square \rightarrow C$  if we ask what licenses our move from its antecedent to its consequent, a move that, by standard counterfactual theory (see Lewis 1973, Stalnaker 1968, Bennett 2003), amounts to the claim that any relevant possible A-situation is indeed a C-situation. Here we find marked differences between mundane and scientific counterfactuals on the one hand, and philosophical counterfactuals on the other.

Counterfactual thinking is ubiquitous in our everyday dealings with the world. We rely on *mundane counterfactuals* such as

- (8) If John had voiced his doubts at the board meeting, he would have been fired.

to assess choices, evaluate strategies, and learn from mistakes. What entitles us to move from the antecedent of a mundane counterfactual such as (8) to its consequent is well-confirmed empirical belief. We clearly take it for granted that these counterfactuals are subject to, as well as in need of, empirical confirmation.<sup>8</sup> And we patently expect anyone issuing a verdict on (8) to back her claim by empirical evidence establishing that any relevant possible situation where John raises his doubts is one where he gets fired. Someone asserting (8) might for instance point out that John's co-worker Bob did voice his doubts, and did get fired.

As a rule, the empirical evidence someone presents won't by itself allow us to pass verdict on her counterfactual claim. We typically need to invoke theoretical elements. For mundane counterfactuals, we habitually need to rely on our background theory of folk truths and folk laws, the latter comprising causal generalizations, social rules, economic presumptions, psychological principles, and the like. In short, what licenses our move from a mundane counterfactual's antecedent to its consequent typically is empirical evidence together with empirical theory.

Basically the same holds true for counterfactuals in the natural sciences. Consider this (rather uninspired) *scientific counterfactual*:

- (9) If someone formed a bare sphere of 7 kg Californium<sup>251</sup>, she would initiate a nuclear chain reaction.

---

<sup>8</sup>. The same holds good for empirical *disconfirmation*. For simplicity's sake, I throughout exclusively mention confirmation.



We manifestly take it for granted that such counterfactuals are subject to, and in need of, empirical confirmation. As before, we expect someone passing a judgment on (9) to provide empirical evidence in support of her verdict. In order to support our counterfactual, pointing out that Californium<sup>251</sup> is a fissile material with a bare-sphere critical mass of 7 kg or less would do fine. For given these empirical facts, one needs but a rough grasp of basic nuclear physics to see that (9) is indeed true. Our evaluation of the scientific counterfactual (9) thus is informed by general empirical evidence together with an empirical background theory. This time, however, both fall within the purview of science.

Our evaluation of counterfactuals such as (8) and (9) is critically informed by empirical evidence and by empirical theory, or so I have maintained. We take it for granted that these counterfactuals are subject to empirical (dis-)confirmation, we expect anyone issuing a verdict on them to provide empirical evidence, and we draw on empirical fact and/or empirical theory to move from their antecedents to their consequents. What in these cases licenses us to hold that any relevant possible A-situation is a C-situation is some body of beliefs *b*, where our entitlement to hold *b* is grounded in experience. This provides sufficient reason to hold that the counterfactuals themselves are to be classed as *a posteriori*, too.<sup>9</sup>

That does not hold true for all counterfactuals. Some counterfactuals are firmly to be classed as *a priori*, since the body of beliefs we draw on to move from their antecedents to their consequents is manifestly *not* grounded in experience. Just consider:

- (10) If twelve guests had come to dinner, more than ten guests would have come to dinner.

We might of course ask someone asserting (10) to justify her judgment. But we do neither expect nor require empirical evidence. We will be perfectly happy with her pointing out that there being twelve *F*s entails there being eleven, thus accepting a logico-mathematical truth as licensing her move from (10)'s antecedent to its consequent. Similarly, consider someone defending her judgment that

---

<sup>9</sup> I take it that *S* knows a priori that *p* iff *S* knows that *p* and *S*'s entitlement to believe that *p* is not grounded in experience; see Nimitz/Kompa/Suhm 2009. By this standard, a proposition *p* is *a priori* iff it can be known *a priori*.

- (11) If Paul had been a bachelor (as he claimed to be), he would have been unmarried.

This again requires no empirical evidence, since we straightaway see that the connections between ‘bachelor’ and ‘unmarried’, a connection we may for all that it’s worth dub ‘conceptual’, licenses her move from (11)’s antecedent to its consequent. In fact, if she were to defend (11) by stressing that all the bachelors she has heard of are unmarried, we would conclude that she had missed the point of her own claim.

The distinction between a priori and posteriori counterfactuals does not coincide with Bennett’s (2003) distinction between dependent and independent counterfactuals who explains independent conditionals to be “the ones where the route from A to C owes nothing to any particular contingent fact or belief” (Bennett 2006, 174, see 148f, 174f). In other words, independent counterfactuals are those where we can move from A to C on the strength of the respective theory *alone*. Since this theory might be empirical, not all independent counterfactuals are *a priori*. But all *a priori* counterfactuals are independent. What renders (10) and (11) *a priori* is the combination of two facts: first, their truths (and hence their evaluation) is independent of particular contingent fact, since it exclusively relies on a background body of beliefs, and, secondly, our entitlement to hold that body is *a priori*.

There is no doubt that philosophers habitually invoke *a posteriori* counterfactuals such as (8). But that is of no concern to us. What we need to ascertain is how philosophical counterfactuals such as (4) and (5) are to be classed. I hold that we have ample reasons to consider those *a priori*.

First of all, we do not treat these statements as in need of, or as subject to, empirical confirmation. In fact, it is hard to see what kind of empirical evidence could possibly settle the issue. Williamson (2007, 192f) rightly points out that we can create actual Gettier-type situations. But those won’t settle the issue. Since we don’t have a device telling us whether the epistemic state our subject is in is indeed a state of knowing, we have to rely on the very same resources we rely on in passing verdict in hypothetical cases.

Secondly, counterfactuals such as (4) and (5) are independent. Particular contingent facts do not bear on their truth; what licenses our move from their antecedents to their consequents are our background accounts. But by standard taxonomy, these accounts are non-empirical. Just as we

defend the likes of (10) by stressing logico-mathematical entailment, we defend the likes of (5) by pointing out that this is what we take knowledge to be.

Thirdly, please note that philosophical counterfactuals often are *counterlegals* in that they concern nomologically impossible situations (Goodman 1983, 7f). (4) is a case in point. But we can hardly draw on empirical theory and empirical laws to license the move from the antecedents to the consequents of counterfactuals of that variety (Bennett 2003, 227f). I conclude that the philosophical counterfactuals are often best classed as *a priori*.

Let me be clear about my argument. I have diagnosed a epistemic difference between mundane and scientific counterfactuals on the one hand, and logico-mathematical, conceptual and philosophical counterfactuals on the other; a difference I have characterized drawing on the *priori/a posteriori* distinction. Those with little faith in this distinction will mind me putting the diagnosed difference in its terms. But they still have to acknowledge that there is a marked epistemic difference between these counterfactuals. In the end, this is all I need.

##### 5. *Against Williamson's no-consequence assessment*

In arguing that philosophical counterfactuals are to be classed as *a priori*, I have presumed that such a classification would be of consequence. Williamson's takes this to be a mistake (see Williamson 2004, 2007b, 2007, ch. 5 and 6). He maintains that "the question '*A priori* or *A posteriori*?' is too crude to be of much epistemological use"(2007, 169). To back this claim, Williamson argues that once we have settled that we can reliably evaluate *a posteriori* counterfactuals, we can rest assured that the same holds true for counterfactuals best considered to be *a priori*:

We have a general cognitive ability to handle counterfactual conditionals. (...) But we have no good reason to expect that the evaluation of 'philosophical' counterfactuals such as [10] uses radically different cognitive capacities from the evaluation of 'unphilosophical' counterfactuals. (...) [W]e should not suppose [10] to involve fundamentally different cognitive capacities from evaluations of counterfactuals that we classify as a *posteriori*. *Consequently, we should not suppose it to raise fundamentally new questions of reliability.* (Williamson 2004, 13f, my italics)

Williamson argues in two steps. In a first step, he maintains that the very same cognitive capacities underlie our evaluation of all kinds of counterfactuals. He bolsters this claim by a general account of the cognitive procedures leading us to assert counterfactuals alluded to above. According to Williamson, we suppose the antecedent and then develop the supposition “adding further judgments within the supposition by reasoning, off-line predictive mechanisms and other off-line judgments” (Williamson 2007, 152f), and we assert the counterfactual if this development “eventually leads one to add the consequent”(ibid., 153). This process will often involve what Williamson terms ‘imaginative simulation’(ibid., 152).<sup>10</sup> In a second step, Williamson draws an epistemic moral from this cognitive sketch: Since one and the same cognitive procedure underlies evaluations of mundane and philosophical counterfactuals alike, the question of whether we can reliably evaluate counterfactuals of the latter kind is answered by our manifest ability to reliably evaluate counterfactuals of the former variety.

I hold that Williamson’s inference from a cognitive premise to an epistemic conclusion fails. To begin with, there is a reading of Williamson’s cognitive premise on which it is false. On a narrow individuation of cognitive capacities, evaluating the mundane counterfactual (8) and the logico-mathematical counterfactual (10) involves *different* cognitive capacities. These statements both are of the form  $A \Box \rightarrow C$ . But they comprise different contents and I need to draw on different cognitive resources to reason from their respective antecedent and to their respective consequent. I need to employ factual knowledge in combination with folk-theoretical projection in the one case, whereas I need to rely on logico-mathematical entailment in the other. Hence, in order for Williamson’s cognitive premise to be true, we need to read it as concerning our cognitive capacities *broadly individuated*.

But read thus, Williamson’s cognitive premise does not license his epistemic conclusion. From the fact that our evaluation of (8) and (10) relies on the same cognitive mechanisms broadly individuated, it does not follow that we are as reliable in evaluating (10) as we are in evaluating

---

<sup>10</sup> Williamson initially assigns the imagination a key role in the process described, see his 2004, 2007b. He later downplays that role, see his 2007, 152f. That sits well with my assessment of how we evaluate counterfactuals which does not assign imagination any special role.

(8), or that we rely on the same sort of evidence. These statements comprise different contents. But we have every reason to believe that how reliable we are at evaluating a counterfactual conditional does critically depend on how good we are at counterfactually relating the contents they comprise; the same holds good for what evidence we rely on. Sameness of cognitive capacity broadly construed thus does not ensure sameness – or even likeness – of epistemic status. In fact, I have taken some pains in the last section to argue that we find marked differences in what it takes to evaluate a counterfactual. These epistemic differences are of course compatible with the idea that the same cognitive capacity broadly individuated underlies all such evaluations. I therefore conclude that Williamson’s argument fails.

To support this conclusion, consider an analogy. On a broad individuation of cognitive capacities, we may agree that basically the same cognitive capacity underlies our evaluations of all statements of the form A & B. On this understanding of sameness of cognitive capacity, evaluating “Venus revolves around the sun & Pluto revolves around the sun” does not involve cognitive capacities radically different from those we bring to bear in evaluating “7 is an odd number & 11 is an odd number”. Still, it does not follow that our reliability in evaluating conjunctions concerning astronomic facts guarantees that we are likewise reliable to evaluate conjunctions concerning mathematical facts. The analogous holds true for (8) and (10).

#### 6. *Notions and how we evaluate philosophical counterfactuals*

My case has hitherto been negative. I have contended that neither our empirical folk theory of the world, nor our empirically established scientific theory licences us to hold true philosophical counterfactuals such as (4) and (5). In the remainder of this paper, I present a positive account of what philosophical thought experiments amount to. Introducing the notion of a ‘notion’ to sum up the conditions guiding how we do and would apply our terms, I maintain that our tacit knowledge of notions is an excellent candidate for what we draw on in evaluating philosophical counterfactuals (§6). I go on to defend a neo-descriptivist semantics that assigns notions centre stage in determining the semantic properties our terms have

in our mouth (§§7–8). I conclude that our tacit knowledge of notions is semantic knowledge. As it turns out, then, philosophical thought experiments simply are exercises in conceptual analysis (§9).

How we apply our terms is not arbitrary. Whether or not we, on reflection, class something as a planet or a sofa depends on the properties we take the object(s) in question to have. The same holds true of the way we would employ our terms. There should be no doubt that our considered application – I will exclusively be concerned with considered application and will henceforth drop this qualification – of ‘sofa’ or ‘planet’ across possible situations depends on the properties we take the relevant object to have. But if the application of terms by a speaker across possible situations is sensitive to the properties entities have, then she has to associate conditions with her terms determining how she applies them. Let me call the conditions a speaker *S* associates with an expression  $\phi$  that determine how *S* would apply  $\phi$  the *notion* *S* associates with  $\phi$ .<sup>11</sup> For example, the notion I happen to associate with ‘grandmother’ is *being a female parent of a parent*, since it is the female parents of parents that I, on reflection, apply my term ‘grandmother’ to across all possible situations.<sup>12</sup>

I take the idea that we associate notions with our terms governing how we do and would apply them to be non-contentious. To begin with, everyone agrees that there must be some cognitive structure on the speaker’s side determining how she does and would apply her terms. Given that application is sensitive to the properties we take the objects to have, taking this cognitive structure as encoding a condition objects may meet seems straightforward to me. Secondly, note that my claim is deliberately neutral on the specifics of the cognitive structures making it the case that I associate, say, *being a female parent of a parent* with the term ‘grandmother’. In fact, it is intended to be compatible with all ideas on offer. There is a reason for this. As might be clear from my characterization, notions are *con-*

---

<sup>11</sup>. Jackson, this volume, talks of ‘patterns in nature’ rather than of ‘conditions’.

<sup>12</sup>. Let me emphasize that although notions guide how we apply our terms, we do not have to think of those as meta-linguistic. Put into words, my  $\phi$ -notion will consist of material-mode sentences that, taken together, specify the conditions that someone has to meet whom I count as a  $\phi$ . Thus “Grandmothers are female parents of parent” exhaustively captures my ‘grandmother’-notion, whereas “Banana are fruit” partially characterize my ‘banana’-notion. This is in line with how we generally explain what our words mean, or apply to. Asked what ‘grandmother’ means, we simply say “Grandmothers are female parents of parents”.

*tents* (i.e. what does the representing). This is why I can leave issues concern representational *vehicles* (i.e. what bears these contents) to one side. Finally, the claim that there are notions is not a *semantic* claim about aspects of what our terms mean. It is a *cognitive* claim about how we apply them. In order not to beg the question against the Kripkean externalists, who maintains that how we do or would apply, say, ‘heat’ or ‘gold’ does not affect what these terms mean in our mouths, we need (at least for the time being) to keep the semantic and cognitive dimension neatly apart.

For many of our terms, we are hard pressed to specify the notions we associate with them without further ado. So it appears that we typically do not have explicit knowledge of which notions we associate with our terms. Yet the notion I associate with a term  $\phi$  must be manifest in how I do and would apply it, and this is something I can try to work out by considering possible situations and reflecting on how I apply  $\phi$  across those. By this *method of possible cases*, we aim to make our respective notion explicit. That is to say, we aim to determine the condition  $N$  such that, for any possible situations  $C$  (where  $\phi$  is applicable at all), we apply  $\phi$  to something if it is  $N$  in  $C$ .<sup>13</sup> There is little doubt that we can reliably attain some clarity about our notions in this manner. We rely on the method of possible cases in our everyday diagnoses of differences in use, and hence in notion. Barring incoherent use, we do not question someone’s judgment that she would not count, say, a Basque-style beret as a hat. We agree that a speaker may be mistaken about what a term means, and hence applies to. But we by default grant her authoritative insight into how she would apply it.<sup>14</sup>

As I have explained it, the method of possible cases is an indirect procedure. I aim to spell out my  $\phi$ -notion by determining how I would apply  $\phi$ , rather than by retrieving it from some mental manual. Any proposal as to what my  $\phi$ -notion is will therefore typically be projected from a partial

---

<sup>13</sup>. There is *no* presumption that  $N$  breaks down into a neat conjunction of properties. To the contrary. Our notions will often be complex or even disjunctive, they may comprise recognitional conditions (“ $\phi$ s look *thus*”), be indexed to specific items or events (“ $\phi$ s are the animals I have seen this one day in Berlin Zoo”), or comprise rigidifying elements making the application of the terms they govern in possible situations dependent upon their actual application.

<sup>14</sup>. As a case in point, note that champions of experimental philosophy (see Knobe/Nichols 2008) do not doubt that their interviewees can reliably tell how they apply their terms.

survey of possible cases, and a case not yet considered might force a correction. This familiar phenomenon explains why proposals are so much easier ruled out than confirmed (see Shope 1983). It might also explain why we are confident of having attained something approaching comprehensive clarity with respect to the notions of just a select few terms.

As already said, I take it to be uncontroversial that there are notions. What is controversial is whether our notions play a role in semantics. But let us for leave this issue until the next section. Semantic or not, notions seem well-suited to guide our evaluation of philosophical counterfactuals such as (5), and we in all plausibility draw on our notions to move from the antecedent of a philosophical counterfactual to its consequent.

First of all, suppose you find yourself, on reflection, withholding application of ‘knowledge’ to the epistemic state of our subject in Lehrer’s case. Suppose further that your most natural justification for this is quite simple: you point out that the subject does not count as a knower as you employ the term. This assessment will by default lead you to accept (5). Secondly, given the interplay of getting clear about our notions and the method of possible cases, evaluating the likes of (5) and getting clear about the notion you associate with ‘knowledge’ amounts to very much the same thing. More precisely, making up your mind on whether to count the protagonist as a knower just is part of determining which condition  $N$  you associate with ‘knowledge’ such that, for any possible situations  $C$  (where ‘knowledge’ is applicable at all), you apply ‘knowledge’ to something if it is  $N$  in  $C$ . Finally, your implicit knowledge of your notion can be classed as (weakly) *a priori*. You do not need to do empirical research to determine how you apply ‘knowledge’ across possible situation. Reflection works just fine.

However, guiding our evaluations is one thing, warranting them is quite another. The latter requires that our application tracks our terms’ semantic properties such that our application of  $\phi$  reliably indicates what the term  $\phi$  in fact applies to. Suppose your considered use of ‘hat’ is such that you accept “Basque-style berets aren’t hats”. This judgment is unlikely to be true unless how you apply ‘hat’ tracks what ‘hat’ in fact applies to. Consider an analogy. If you employ stick  $M$  as a yardstick for the length of one meter,  $M$  is likely to guide your evaluation of “Saint Peter’s Square is 240 meters wide”. But your judgment won’t be warranted unless  $M$  happens to be (roughly) one meter in length. Your stick-based



judgments need to track the length in question. Likewise, the considered application of your terms needs to track what they actually apply to. If your considered application is not a reliable guide to your terms' semantic properties, going by the former to pass judgment on the latter is not a sensible idea.

### 7. *Neo-descriptivism and the argument from communication*

Embracing the semantics I label *neo-descriptivism* renders it easy to argue that considered application tracks semantic properties.<sup>15</sup> Neo-descriptivism's key contention is that *notions determine semantic properties*. Put a bit more circumspect, the idea at the heart of this semantics is this:

**ND** The notion we associate with a term  $\phi$ , guiding the (considered) application of  $\phi$  across possible situations, determines the semantic properties  $\phi$  has in our mouths.

The idea is that the term 'hat' in our mouths applies to the kind of headgear it does apply to because these items best satisfy the notion guiding our considered use of 'hat'. This of course guarantees tracking. So if neo-descriptivism happens to be the right semantics, our evaluation of philosophical counterfactuals drawing on our considered application warrants us to move from their antecedents to their consequents. In much the same vein, our employment of stick *M* will warrant our meter-judgments if (we have grounds to hold that) *M* happens to be the standard meter.

Neo-descriptivism explains why our expressions mean what they do by embracing the descriptivist paradigm according to which the semantic properties of our expressions are in essence fixed by our cognitive states. In one respect, this is an uncontroversial stance to take. Barring the details, everyone agrees that many of our terms are notion-governed. Nobody thinks that what 'shareholder' applies to, or what 'here' refers to, is determined by factors beyond the conditions guiding their application. In another respect, neo-descriptivism is highly controversial. Kripke's (1980)

---

<sup>15</sup>. I take it that David Chalmers and Frank Jackson are committed to neo-descriptivism. See Jackson, this volume, as well as Chalmers & Jackson 2001, §3, Jackson 1998, 46–52, 2000, 2004, 2005, Chalmers 1996, 56–71, 2002, 2004. See also Nimtz 2004.

anti-descriptivist arguments have convinced most semanticists that the semantic properties of (at least) proper names and natural kind terms are externally determined. This is key idea of Kripkean externalism. Kripkean externalists hold that semantic properties are fixed by factors such as causal-historical chains leading back to initial baptisms – factors that are beyond the associated notions, and hence beyond what speakers may uncover by mere reflection.

I won't here take issue with Kripke's arguments.<sup>16</sup> Focussing on natural kind terms, and drawing on ideas to be found in Jackson (2007b, 2005, 2004, 266ff, 1998, 40, Fn. 16), I will rather turn the table on the Kripkean externalist. I will argue that it founders on an epistemic problem: it cannot account for knowledge competent speakers need to possess, and manifestly do possess, in order to participate in our communicatory practice.

Much of what we do aims at putting across information. We embellish our cheeks with small Swiss crosses to get across that we support the Swiss team, we put the name 'Christian Nimtz' on an office door to let people know that this is my office, and we put little red dots on boxes to convey that the article contained is red. But what is it to convey information – to get across an informational content? Here is a sensible answer:

**p1** To convey content *c* to an audience is to put the audience in a position to exclude possibilities in accordance with *c*.

This fits well the general idea that to bear information is to exclude possibilities (see Stalnaker 1984, ch.1, Jackson 2000, 331). It also fits with our examples. The Swiss emblem on the cheek allows you to rule out all other possible partialities, and the red dot allows you to rule out other possible colours for the contained article. There of course is no guarantee that things indeed are as the content conveyed presents them to be, and you have to be able to read the signs, as it were, for the content to get across. If you don't realize that national emblems are worn to indicate whom you support, you might well end up sitting with the wrong crowd.

Our most important means to convey information is language – we tell each other how things are (or rather: how we take them to be). Language is well-suited to that purpose. Asserting "Smoking kills" allows me to

---

<sup>16</sup>. See Spicer, this volume, for a statement and defence of Kripkean externalism as well as a summary of Kripke's famed arguments.

convey that smoking kills. It does so by and large irrespective of context. That is to say, I can convey such a content by asserting such a sentence to just about every competent speaker of my language, and doing so requires no background knowledge on the part of the audience beyond what realizing the pragmatic circumstances demands.<sup>17</sup> That is easily explained, for what we do is we simply exploit the sentences' truth-conditions and convey the semantic contents they bear.<sup>18</sup> This is always an option:

**p2** We can employ our declarative sentences to convey their semantic contents to competent speakers mostly irrespective of context.

You can convey to a perfect stranger that your wife just gave birth to twins asserting "My wife just gave birth to twins", thereby putting her in a position to exclude possibilities in accordance with the sentence's content. That is, you can do so given that she knows the sentence's semantic content. Asserting "Vaimoni synnytti juuri kaksoset" doesn't reliably bring the very same information across around here since most people don't speak Finnish and thus do not associate this sentence with that content. But if you happen to speak to a competent speaker of English, asserting an English sentence is a highly reliable means to convey its content.

Sentence containing natural kind terms – or names, for that matter – are no exception to that rule. Asserting "Iran does not possess any weapons-grade plutonium" puts competent speakers in a position to exclude a rather worrisome possibility, and asserting "There's water on Mars!" allows the audience to exclude a waterless Mars from the way they believe things to be. (As before, asserting the Finnish sentence "Marsissa on vettä!" won't work around here.) There should be no doubt that we can and do use such sentences to convey their semantic contents to competent speakers irrespective of context. In fact, part of the point of having those

---

<sup>17</sup>. You need to know enough about pragmatic circumstances to get clear about what Perry/Korta 2006 summarize as 'near-side pragmatics', i.e. ambiguity, context dependence and indexicality.

<sup>18</sup>. Those are equivalent. Suppose that sentence *S* is true iff smoking kills. Then the sentence's content is *that smoking kills*. Accepting an assertion of *S* puts you in a position to exclude possibilities in accordance with that content – you may rule out that our world is on in which smoking doesn't kill.

sentences in our repertoire is that they can be used without contextual support to get those very contents across.

Conveying contents by way of assertions requires knowledge on the part of the audience. In fact, it holds true that:

**p3** We could not employ our sentences as p2 says we do unless competent speakers knew which semantic contents our sentences bear (and knew this in a non-trivial way).

We all know who shot Kennedy: whoever shot Kennedy did. By the same token, we all know how someone asserting “Sataa” presents things as being: she presents things as being such that “Sataa” is true. Knowledge of this variety is trivial in that it does not put the knower in a position to actually exclude possibilities. Conveying contents by asserting sentences requires non-trivial knowledge of the sentences’ semantic contents on the part of the audience.<sup>19</sup> Unless I know that “Sataa” bears the semantic content that it’s raining, you cannot inform me about how things are (according to you) by asserting that sentence.

Our very praxis of conveying information by assertive utterances thus requires competent speakers to have non-trivial knowledge of semantic contents. This holds true for sentences that do not contain natural kind terms or proper names, as well as for those that do. Consequently, no semantics can be an adequate semantics for our language unless it accounts for this knowledge. This provides a good reason to embrace neo-descriptivism:

**con** The idea that the semantic properties of natural kind terms are notion-governed accounts for non-trivial knowledge of semantic contents. The idea that the semantic properties of natural kind terms are externally determined does not.

If the semantic properties of ‘water’, ‘plutonium’ etc. are externally determined, even competent speakers might well be ignorant of the semantic contents of our sentences comprising these terms. In fact, since on the orthodox causal-historic account initial baptism determines subsequent

---

<sup>19</sup> I feel that in his critique of my argument, Spicer, this volume, does not sufficiently appreciate that the knowledge conveyed needs to allow the audience to exclude possibilities in this manner.

meaning, this semantics allows that even all our experts might share the laymen's ignorance. If, however, these properties are determined by the notions we associate with 'water', 'plutonium' etc., the required knowledge is guaranteed. In other words, the neo-descriptivist picture accounts for the non-trivial knowledge on the part of competent speakers requisite for our communicative praxis, whereas the externalist picture doesn't. I conclude that we have every reason to hold ND. The notions we associate with our expressions do indeed determine the semantic properties our expressions have in our mouths across the board.

#### 8. *Objections and replies*

Kripkean externalism founders on an epistemic problem, or so I have argued. It cannot account for knowledge competent speakers need to possess, and manifestly do possess, in order to participate in our communicatory practice. Let me briefly deal with two objections to this argument and the neo-descriptivist stance it supports. Here is the first objection:

- Can't we hold that the contents we unvaryingly convey with our sentences have nothing to do with their *semantic* contents?

The proposal is that "There is water on Mars" has an externally determined semantic content (roughly: *that there is H<sub>2</sub>O on Mars*. Call this the 'secondary content') that is utterly independent from the non-semantic content unvaryingly conveyed with assertive utterances of that sentence (roughly: *that there is watery stuff on the Red Planet*. Call this the 'primary content'). There is something right about this. We should indeed acknowledge that our sentences have both contents here distinguished. The resulting two-dimensional semantic, assigning primary and secondary contents to our sentences and primary and secondary intensions to our subsentential terms, is of course congenial to neo-descriptivism (see Jackson 2004, Nimtz 2007, ch.8).

However, there is also something wrong about this. First, we have every reason to class the content conveyed with "There is water on Mars" as semantic. After all, bearing that content is a representational feature of the sentence, it is a feature any competent speaker has to be aware of, and it is a feature that we can and do exploit in communication irrespective of

the context. Secondly, the two contents in play are not independent of one another. That ‘water’ in our mouths picks out (i.e., secondarily designates) H<sub>2</sub>O depends on the fact that we use this very term to pick out (i.e., primarily designate) the watery stuff around here. Assuming otherwise doesn’t fit with what we find. What ‘elm’, ‘water’, ‘plutonium’ etc. pick out is constrained by what we use these terms to pick out. If our considered use of ‘elm’ picks out a variety of trees, the term’s semantic value cannot turn out to be a kind of bird. How we use our terms and what they apply to cannot deviate that much. Worse still, assuming otherwise means risking methodological self-defeat. Kripkean externalists rely on our intuitive judgments as to what ‘water’ designates to determine what the term in fact designates. But our intuitive judgments are informed by how we apply that term, and what we convey with it. But if the term’s designation wasn’t constrained by our application of it, we had no reason at all to hold that our intuitions track the term’s semantic properties.

The neo-descriptivist stance I have defended invites a more general worry. The objection here is this:

- Neo-descriptivists hold that what determines semantic properties are the notions individual speakers associate with their terms. But how can such an individualist account explain that we all speak *English*?

The driving force behind this objection is the lingering suspicion that neo-descriptivism can neither account for the collective nature of speaking a language, nor for the semantic homogeneity we find amongst speakers of the same language. However, to think so would be to overlook, first, the emphasis neo-descriptivists put on deference. They emphatically embrace Putnam’s ‘division of linguistic labour’ (1975, 227f) and acknowledge that many speakers defer to their co-speakers in their application of terms (Jackson 2004, 270–273, Chalmers & Jackson 2001, 327f, Chalmers 2002, 170–173). The resulting socio-linguistic picture distinguishes non-deferring experts from deferring laymen and thereby avoids the threat of rendering semantics a private matter.

To think so would, secondly, underrate the importance of conventions (see Jackson 2004). We mutually expect one another to associate notions with our expressions guaranteeing that communication proceeds smoothly. This requires that we convey essentially the same contents and

associate essentially the same notions with our expressions. (I'll come back to the 'essentially' in a moment.) This does not come about by itself. It rather requires social mechanisms and a collective effort. This is precisely what we find. We correct speakers whose notions differ from ours more than we think tolerable, we compile dictionaries and consult them to further consistency and settle controversies, we flag irreconcilable differences in notions if we think they matter, and we pass authoritative rulings on which conditions one should associate with a term, even if that means having only eight planets in the solar system instead of the traditional nine. Furthermore, whenever we teach someone a term, we do not rest content unless she applies it in line with our notions. In short, we employ reliable means to make people associate the right conditions with their terms.

To think so would, finally, be to succumb to what I dub the *homogeneity myth*. If we ignore context-dependence for the moment, it seems quite natural to hold that, since we all speak the natural language, *viz.* English, our words must mean the same. Embracing this is to presume that differences in meaning are incompatible with sameness of natural language. Natural though it may seem, this idea of semantic homogeneity is a myth. The natural languages we know *aren't* semantically homogenous. We find all kinds of semantic differences *within* the natural language we share. Consider once more the term 'hat'. According to the OED, there are two uses of this term. On the former, a hat is almost any covering for the head. On the latter, a hat is a specific kind of headgear that typically has "a more or less horizontal brim all round the hemispherical, conical, or cylindrical part which covers the head". If you follow the former use and I follow the latter, you will and I won't count a woollen bonnet as a hat. But there is no basis for holding that one of us must be making a mistake. And the fluency of our communication proves that both of us do, of course, speak English.<sup>20</sup>

Moreover, the natural languages we know *need not be* semantically homogenous. Reliable communication does not require that the content the audience takes up has to be precisely the one the speaker expresses. As long as the way I present things as being in uttering "Nice sofas are hard to find" and the way you take things to be in believing me are sufficiently

---

<sup>20</sup>. I develop this in more detail in Nimtz 2009.

similar, everything is fine. Generally speaking, there is no need for all the notions in a community to be exactly the same; it suffices if they overlap substantially. Our praxis reflects that. On the one hand, we often find us willing to tolerate deviations we deem harmless. We usually don't mind if how you apply 'sofa' or 'hat' slightly deviates from how I do. On the other hand, we habitually presuppose that our notions differ, and that we need to agree on our notions when something is at stake. It is no accident that our laws and contracts are full of meticulous classification.

9. *Thought experiments, conceptual analysis, and the limits of arm-chair philosophy*

In the paper, I have put forth two different lines of thought – one epistemic, the other semantic. As for the first line of thought, I have argued that there is a marked epistemic differences between mundane and scientific counterfactuals such as (8) and (9) on the one hand, and philosophical counterfactuals such as (4) or (5) on the other. Counterfactuals of the former variety are to be classed as *a posteriori*, since the bodies of belief licensing our move from their antecedents to their consequents comprise empirical evidence and/or empirical theory. By contrast, the counterfactuals we find at the heart of philosophical thought experiments are to be classed as *a priori*.<sup>21</sup> I have also ventured that we draw on the notions guiding how we do and would apply our terms to assess such philosophical counterfactuals. For example, we consider (5) to be true because we find that the subject in the Nogot/Havit-case does not satisfy the conditions guiding our application of 'knowing', and hence does not count as a knower.

As for the second line of thought, I have proposed and defended a semantic theory I label *neo-descriptivism*. The semantics is labelled *neo-descriptivism* since it maintains that what determines the semantic properties our terms have in our mouths are the notions we associate with these terms, guiding their (considered) application across possible situations. The semantics is labelled *neo-descriptivism* since by its two-dimensional structure, it incorporates key Kripkean insights such as that many of our

---

<sup>21</sup>. Mutatis mutandis the same holds true if we take modal links to be restricted necessities in line with (7), rather than bona fide counterfactuals.



terms designate rigidly, that there are necessary truths a posteriori, and that there are contingent truths a priori (see Kripke 1980, lecture I). For all that, the neo-descriptivist idea that notions determine semantic properties patently runs against Kripkean externalism. I have devised the argument from communication to defend neo-descriptivism against its externalist rival. I have argued that since neo-descriptivism can account for non-trivial knowledge competent speakers do possess, and need to possess in order to participate in our communicatory practice, whereas Kripkean externalism cannot, we have every reason to accept the neo-descriptivist account over the Kripkean alternative.

Combining the two lines of thought yields a lucid account of philosophical thought experiments. What we go on in evaluating the counterfactuals at the heart of philosophical thought experiments are the notions guiding our considered application of our terms. But on neo-descriptivist premises, these notions are semantic in that they determine the semantic properties our terms have in our mouth. What we do in contemplating the likes of the Mary-scenario or the Nogot/Havit-case, then, is that we get clear about what our terms apply to by inquiring into how we would apply them, thereby rendering explicit the notions we associate with them. Philosophical thought-experiments thus turn out to be instances of the method of possible cases, and philosophical thought-experimenting is found to be a variety what has traditionally been dubbed ‘conceptual analysis’ (see Grice 1958, Jackson 1998, ch.2) – a variety, that is, of the endeavour to ascertain what our words mean in our mouths.<sup>22</sup>

#### REFERENCES

- Bealer, George. 1998 “Intuition and the Autonomy of Philosophy.” In DePaul/Ramsey (eds.) 1998, 201–239.
- Bealer, George. 2000. “A Theory of the A Priori.” *Pacific Philosophical Quarterly* 81, 1–30.
- Bealer, George. 2002. “Modal Epistemology and the Rationalist Renaissance.” In Gendler/Hawthorne (eds.) 2002, 71–125.

---

<sup>22</sup>. I would like to thank Jessica Leech and Wolfgang Schwarz for valuable comments on earlier versions of this paper. This material has also greatly profited from discussions in Zürich and Cologne, for which I am very grateful.

- Bennett, Jonathan. 2003. *A Philosophical Guide to Conditionals*. Oxford: Oxford University Press.
- Berger, Alan. 2002. "A Formal Semantics for Plural Quantification, Intersentential Binding and Anaphoric Pronouns as Rigid Designators." *Nous* 36, 50–74.
- Chalmers, David. 1996. *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press.
- Chalmers, David. 2002. "On Sense and Intension." *Philosophical Perspectives* 16, 135–182.
- Chalmers, David. 2004. "Epistemic Two-Dimensional Semantics." *Philosophical Studies* 118, 153–226.
- Chalmers, David and Frank Jackson. 2001. "Conceptual Analysis and Reductive Explanation." *Philosophical Review* 110, 315–361.
- Cooper, Rachel. 2005: "Thought Experiments." *Metaphilosophy* 36, 328–347.
- Davidson, Donald 1987. "Knowing One's Own Mind." In *Self-Knowledge*. Cassam, Quassim (ed.). 1994. Oxford: Oxford University Press, 43–64.
- Dennett, Daniel. 1995. "Intuition Pumps." In *The Third Culture: Beyond the Scientific Revolution*. Brockman, John (ed.). 1995. New York: Simon & Schuster, 167–181.
- DePaul, Michael and William Ramsey (eds.). 2000. *Rethinking Intuition. The Psychology of Intuition and Its Role in Philosophical Inquiry*. Lanham/Boulder: Rowman & Littlefield
- Edginton, Dorothy. 2008. "Counterfactuals." *Proceedings of the Aristotelian Society* 108, 1–21.
- Gendler, Tamar Szabó and John Hawthorne (eds.). 2002. *Conceivability and Possibility*. Oxford: Clarendon Press.
- Gettier, E. L. 1963. "Is Justified True Belief Knowledge?" *Analysis* 23, 151–153.
- Goodman, Nelson. 1983. *Fact, Fiction, and Forecast*. Cambridge (Mass.): Harvard University Press.
- Grice, Paul. 1958. "Postwar Oxford Philosophy." In *Studies in the Ways of Words*. Grice, Paul. 1989. Cambridge (Mass.): Harvard University Press, 171–180.
- Häggqvist, Sören (2009). "Modal Knowledge and Thought Experiments." In Kompa/Nimtz/Suhm (eds.) 2009, 53–68.

- Jackson, Frank 1998. *From Metaphysics to Ethics. A Defense of Conceptual Analysis*. Oxford: Oxford University Press.
- Jackson, Frank. 2000. "Representation, Scepticism, and the A Priori." In *New Essays on the A Priori*. Boghossian, Paul and Christopher Peacocke (eds.). 2000. Oxford: Oxford University Press, 320–332.
- Jackson, Frank. 2004. "Why We Need A-Intensions." *Philosophical Studies* 118, 257–277
- Jackson, Frank. 2004b. "Foreword: Looking Back on the Knowledge Argument." In Ludlow/Stoljar/Nagasawa (eds.). 2004, xv-xix.
- Jackson, Frank. 2005. "What Are Proper Names for?" In *Experience and Analysis*. Reicher, M.E. and J.C. Marek (eds.). 2005. Kirchberg am Wechsel, 102–114.
- Jackson, Frank. 2007. "Reference and Description from the Descriptivists' Corner." *Philosophical Books* 48, 17–26.
- Jackson, Frank. 2007b. "On Not Forgetting the Epistemology of Names." *Grazer Philosophische Studien* 74, 239–250.
- Katz, Jerrold. 1998. *Realistic Rationalism*. Cambridge (Mass.): MIT Press
- Katz, Jerrold. 2002. "Mathematics and Metaphilosophy." *Journal of Philosophy* 99, 362–390
- Kripke, Saul. 1980. *Naming and Necessity*, second edition. Oxford: Blackwell.
- Knobe, Joshua and Shaun Nichols (eds.). 2008. *Experimental Philosophy*. Oxford. Oxford University Press.
- Lehrer, Keith. 1965. "Knowledge, Truth, and Evidence." *Analysis* 25,, 168–175.
- Lewis, David. 1973. *Counterfactuals*. Oxford: Blackwell.
- Ludlow, Peter, Daniel Stoljar and Yujin Nagasawa (eds.). 2004. *There's Something about Mary: Essays on Phenomenal Consciousness and Frank Jackson's Knowledge Argument*. Cambridge (Mass.): MIT Press.
- Miller, Richard. 2000. "Without Intuitions." *Metaphilosophy* 31, 231–250
- Kompa, Nikola, Christian Nimtzt and Christian Suhm (eds.). 2009. *The A Priori and Its Role in Philosophy*. Paderborn: mentis.
- Nimtzt, Christian. 2004. "Reference Fixing, Conceptual Analysis, and the A Priori." In *Selected Papers Contributed to the Sections of GAP.5*. Bluhm, Roland and Christian Nimtzt (eds.). 2004. Paderborn: mentis, 320–331.

- Nimtz, Christian. 2007. *A New Rationalism? Intuitions, Modal Knowledge, and the Analysis of Concepts*. Unpublished manuscript. (*Habilitationschrift*.)
- Nimtz, Christian. 2009. "Conceptual Truth Defended". In: Kompa/Nimtz/Suhm (eds.). 2009, 137–155.
- Norton, John 2004. "On Thought Experiments: Is There More to the Argument?" *Philosophy of Science* 71, 1139–1151.
- Putnam, Hilary. 1975. "The Meaning of 'Meaning'." *Mind, Language, and Reality. Philosophical Papers, Vol. 2*. Putnam, Hilary. 1975. Cambridge: Cambridge University Press, 215–271.
- Quine, Willard V.O. 1995. *From Stimulus to Science*. Cambridge (Mass.): Harvard University Press.
- Shope, Robert. 1983. *The Analysis of Knowing*. Princeton (NJ): Princeton University Press.
- Stalnaker, Robert. 1968. "A Theory of Conditionals." In *Studies in Logical Theory*. Rescher, Nicholas (ed.). Oxford: Oxford University Press, 98-112.
- Stalnaker, Robert. 1984. *Inquiry*. Cambridge (Mass.): MIT Press.
- Williamson, Timothy. 2004. "Armchair Philosophy and Counterfactual Thinking." *Proceedings of the Aristotelian Society* 105, 1–23
- Williamson, Timothy. 2007. *The Philosophy of Philosophy*. Oxford: Blackwell.
- Williamson, Timothy. 2007b. "Philosophical Knowledge and Knowledge of Counterfactuals." *Grazer Philosophische Studien* 74, 89-123.