## How to Use Thought Experiments

# Elijah Chudnoff

Thought experiments figure prominently in contemporary epistemology. Beyond that humdrum observation, controversy abounds. The aim of this paper is to make progress on two fronts. On the descriptive front, the aim is to illuminate what the practice of using thought experiments involves. On the normative front, the aim is to illuminate what the practice of using thought experiments should involve. I'll give priority to normative concerns. Since I'm of the optimistic opinion that epistemologists mostly do what they should, however, my hope is that light can be cast in both directions at once.

Thought experiments result in judgments that are passed on to further philosophical reasoning. What are these judgments? What is the point of making them? What basis do we have for making them? Let's say that a conception of thought experiments includes answers to these three questions. My thesis is that epistemologists should use thought experiments in accordance with a conception of them organized around the idea of explanation. This will take some spelling out, and that is the project pursued in the rest of the paper. My case for endorsing the thesis rests on the fruitfulness of the interventions in current debates about the content, point, and basis of thought experiment judgments that it suggests.

### I. The Content of Thought Experiment Judgments

Let's start with a classic.

Imagine that I have evidence that a Mr. Nogot, who is in my office, owns a Ford. The evidence consists of knowing that Mr. Nogot is in my office, that he drives a Ford, that he says he owns a Ford, that he has papers saying that he owns the Ford, and that he is an honest and reliable man. This evidence consists entirely of true statements. I conclude from it that someone who is in my office owns a Ford. The inference proceeds from true premises to a true conclusion.

As things are imagined, Mr. Nogot does not own a Ford. But Mr. Havit, who is in my office with Nogot, does own a Ford. This is entirely unknown to me. Thus my belief that someone in my office owns a Ford is completely justified and true, but I do not know that someone in my office owns a Ford. (Lehrer 1970)

The text presenting Lehrer's thought experiment naturally divides into two parts. First, there is his description of the imaginary scenario. This is the largest of the two parts and includes all the quoted text except the last sentence. I'll call the scenario, as Lehrer explicitly describes it, the Ford Case. Second, there is Lehrer's stated judgment about the Ford Case: "Thus my belief that someone in my office owns a Ford is completely justified and true, but I do not know that someone in my office owns a Ford." The debate considered here is about the content of this sort of judgment, what I'm calling thought experiment judgements.

There are two background assumptions I share with participants in the debate. First, Lehrer's thought experiment judgment does not wear its content on its sleeve. There is room for reasonable disagreement about how to interpret the judgment, so it is worth trying to express its content more explicitly. The second assumption states a condition of success on this project: the thought experiment judgment can serve as a premise in a straightforward argument against the thesis that knowledge is justified true belief. If knowledge is justified true belief, then necessarily someone knows that p if, and only if, they have a justified true belief that p. There should be a straightforward way to refute this modal claim using Lehrer's judgment as a premise.

Recent discussion of how to express thought experiment judgments traces to Williamson's (2005, 2007) criticism of and revision to the proposal that thought experiment judgments are strict conditionals. On this view, Lehrer's judgment is:

[N] Necessarily, anyone who stands to a proposition p as in the Ford Case has a justified true belief that p but does not know that p.

I will not rehearse the details for all the proposals mentioned here, but to make clearer what counts as a straightforward refutation of the view that knowledge is justified true belief, let's examine how [N] can be invoked as a premise in such a refutation.

- Step 1. Possibly, someone stands to a proposition p as in the Ford Case.
- Step 2. [N] Necessarily, anyone who stands to a proposition p as in the Ford Case has a justified true belief that p but does not know that p.
- Step 3. Possibly, someone has a justified true belief that p but does not know that p.
- Step 4. So, it is not necessary that someone knows that p if, and only if, they have a justified true belief that p.

The problem Williamson notes is that [N] is false because too strong. Consider a scenario that satisfies Lehrer's description of the Ford Case and in which the subject receives, but ignores, evidence that Mr. Nogot doesn't own a Ford. In such a case someone stands to a proposition p as in the Ford Case but does not have a justified true belief that p. Williamson's response to the problem is to replace the strict conditional [N] with a counterfactual conditional:

[C] If someone were to stand to a proposition p as in the Ford Case, then they would have a justified true belief that p but not know that p.

One worry about taking [C] to express Lehrer's thought experiment judgment is that if the augmented Ford Case just described is actually realized, then [C] is false but Lehrer's judgment isn't (cf. Ichikawa and Jarvis 2009, Malmgren 2011). Several variations on and alternatives to the counterfactual proposal have now joined it in the literature, and debate over them remains vigorous (Horvath 2015 and Saint-Germier 2019 consider types of proposal I do not discuss).

The proposal I focus on here is Malmgren's (2011) suggestion to express thought experiment judgments as claims about possibility:

[P] It is possible that someone stands to a proposition p as in the Ford Case and that they have a justified true belief that p but do not know that p.

[N] and [C] run into problems because they are too strong. [P] runs into problems because it is too weak (cf. Gardiner 2015, Geddes 2018). Consider the thought experiment judgment of someone who disagrees with Lehrer. Suppose Lehrer's disputant thinks the subject in the Ford Case lacks justification for believing that someone in their office owns a Ford. Following Malmgren's suggestion, this judgment should be expressed as:

[P\*] It is possible that someone stands to a proposition p as in the Ford Case and that they lack a justified true belief that p and do not know that p.

The problem is that [P] and [P\*] are compatible, so taking them to express the two thought experiment judgments fails to capture the conflict.

Here is where I suggest invoking the idea of explanation. Thought experiment judgments should be expressed as claims about possible explanation. Consider the following:

[B] It is possible that: someone has a justified true belief that p but does not know that p because they stand to a proposition p as in the Ford Case.

[B\*] It is possible that: someone has an unjustified true belief that p and does not know that p because they stand to a proposition p as in the Ford Case.

[B] and [B\*] do conflict. To see this, consider a simpler pair:

[U] It is possible that: someone is a U.S. citizen because they were born in the U.S.

[U\*] It is possible that: someone is not a U.S. citizen because they were born in the U.S.

One cannot consistently endorse [U] and [U\*]. If someone can be a U.S. citizen in virtue of being born in the U.S., then it is not the case that someone can fail to be a U.S. citizen in virtue of being born in the U.S. Of course, it is possible that someone is not a U.S. citizen *even though* they were born in the U.S. For example, they might renounce their citizenship or be convicted of treason. But [U] does not rule these possibilities out, and [U\*] does not invoke them. This point can be put in better relief by elaborating a bit on "because."

I'm using "because" to make statements of non-causal explanatory dependence. There is nothing esoteric about such explanations. If I say, "My car is illegally parked because I couldn't find a spot," then I am giving a causal explanation of why my car is illegally parked. If I say, "My car is illegally parked because it is parked next to a fire hydrant," then I am giving a non-causal explanation of why my car is illegally parked. I might also say, "My car is illegally parked in virtue of being parked next to a fire hydrant." Such explanations are familiar from ordinary communication and philosophical tradition, though it is only recently that they have drawn a dedicated research literature in the way that causal explanations have for some time now.<sup>2</sup>

<sup>&</sup>lt;sup>1</sup> For the purposes of the illustration, I'm assuming the laws of U.S. citizenship do not change from world to world.

<sup>&</sup>lt;sup>2</sup> I mean the literature on grounding, reviewed in (Bliss and Trogden 2014). As discussed in Section 4 of their article, there are different views about the relation between grounding and non-causal explanation. Here I take the notion of non-casual explanation as basic and independently intelligible. In Section III, I consider but do not commit to one possible grounding theoretic approach to it.

Epistemic facts obtain in virtue of non-epistemic facts. It follows that if someone has a justified true belief that p but does not know that p, then this is because of other facts about them. What [B] says is that those other facts might consist in standing to the proposition p as in the Ford Case. This is supposed to be a sufficient explanation. If their standing to the proposition p as in the Ford Case is a sufficient explanation of why someone has a justified true belief that p but does not know that p, then it cannot also be a sufficient explanation of why someone has an *unjustified* true belief that p and does not know that p. This is so even if there is some other sufficient explanation of why someone has an unjustified true belief and does not know that p, where this other explanation includes standing to the proposition p as in the Ford Case as a proper part. Hence [B] is incompatible with [B\*].

The general point that underwrites the foregoing is that giving a sufficient explanation is not the same as giving a metaphysically sufficient condition. Let's approach this point by considering the following objection to [B] as an expression of Lehrer's thought experiment judgment:

A sufficient explanation is a complete explanation. P completely explains Q only if P is a metaphysically sufficient condition for Q. It follows that [B] implies [N]: for suppose [B] is true; then standing to a proposition as in the Ford Case is a metaphysically sufficient condition for having a justified true belief that p without knowing that p; so [N] is also true. Since [N] is objectionable because too strong, it follows that [B] is objectionable because too strong.

We should resist this line of reasoning. There are two options. The first option is to deny that a sufficient explanation must be a complete explanation. Explanations are often taken to be sufficient in context even though they leave out some details. An account of thought experiment judgments need not also be an account of the pragmatics of explanation. The second option is to deny that P completely explains Q only if P is a metaphysically sufficient condition for Q.<sup>3</sup> It is trivial that P completely explains Q only if P includes all the explanatorily relevant factors in virtue of which Q. But it is not trivial that P completely explains Q only if P includes a set of factors that are metaphysically sufficient for Q. It could very well be that some of the factors required to form a set of factors metaphysically sufficient for Q do not play any role in explaining why Q. Such factors might include the presence of enabling conditions or the absence of disabling conditions. I think both options just reviewed are correct. Sufficient explanations might leave out some details, and even when they do not, they might still fail to constitute metaphysically sufficient conditions for what they completely explain.

# II. The Point of Thought Experiment Judgments

Why should epistemologists occupy themselves with thought experiments? One familiar answer connects thought experiments to the project of analysis—semantic, conceptual, or

<sup>&</sup>lt;sup>3</sup> When combined with the assumption of a tight connection between grounding and non-causal explanation, this is a minority view in the literature (Bliss and Trogden 2014, Section 5). I defend it and explore other aspects of its significance for epistemological theorizing in (Chudnoff 2011, 2013), some ideas from which are developed further in the next section.

metaphysical. I favor an answer that connects them to the project of explanation instead. Thought experiments can be used to test the explanatory ambitions of epistemic theories.<sup>4</sup>

I am assuming that epistemic theories have explanatory ambitions, even though these are not always made explicit. Consider virtue theoretic accounts of knowledge. One might try to formulate such an account using a strict biconditional:

 $[V_N]$  Necessarily, someone knows that p if, and only if, they believe truly that p due to their intellectual virtue.

But I think that  $[V_N]$  is both expressively too weak and expressively too strong. It is too weak because it fails to express the idea that knowledge depends on virtuous true belief rather than the other way around. Once we repair this weakness, then we'll be positioned to see that  $[V_N]$  is also too strong because it saddles virtue epistemology with a modal commitment that is incidental to its explanatory ambitions. I suggest the following:

[V<sub>B</sub>] Necessarily, if someone knows that p, then they know that p because they believe truly that p due to their intellectual virtue.

The difference between  $[V_N]$  and  $[V_B]$  opens the following dialectical possibility.<sup>5</sup> Consider Fake Barn cases:

Barney is driving through the country and happens to look out of the window into a field that he is driving past. In doing so, he gets to have a good look at a barn-shaped object, whereupon he forms the belief that there is barn in the field. This belief is true, since what he is looking at really is a barn. Unbeknownst to Barney, however, he is presently in 'barn façade county' where every other object that looks like a barn is actually a clever fake. Had Barney looked at one of the fake barns, then he would not have noticed the difference. Quite by chance, however, Barney just happened to look at the one real barn in the vicinity (Pritchard 2016).

Plausibly, Barney has a virtuous true belief that doesn't amount to knowledge, so  $[V_N]$  is false. There are several familiar directions one might go from here: deny that Barney's belief is really virtuous; argue that Barney's belief does amount to knowledge; modify  $[V_N]$  either by adding

[V<sub>IFF&B</sub>] Necessarily, someone knows that p if, and only if, and because they believe truly that p due to their intellectual virtue.

 $[V_{IFFRB}]$  retains the biconditional form of  $[V_N]$  while adding explanatory content. Once the explanatory content is added, however, then the point of retaining the biconditional form of  $[V_N]$  needs defending. In some special cases such a defense might be mounted. My claim is that there is no obvious interest in giving sufficient conditions, but there is obvious interest in giving all possible explanatory conditions.

<sup>&</sup>lt;sup>4</sup> In this section I continue to focus on the contents associated with thought experiments. Another approach to rationalizing the use of thought experiments that I find plausible focuses on the formats and processes in which their associated contents are represented and manipulated (cf. Aronowitz and Lombrozo 2020, Gendler 2004, 2007, Nersessian 1992).

<sup>&</sup>lt;sup>5</sup> The same cannot be said about the following alternative to both  $[V_N]$  and  $[V_B]$ :

complexities, or by starting over (cf. section 5 of Turri et al 2017). Here is an alternative: take  $[V_B]$  to be the proper expression of the virtue theoretic account of knowledge, concede that Barney has a virtuous true belief that doesn't amount to knowledge, but argue that Barney's belief doesn't amount to knowledge because of factors that are irrelevant to explaining why someone has knowledge when they do. Such factors might be the absence of merely enabling conditions or, perhaps more likely in this case, the presence of special disabling conditions (Blouw et al 2018 nicely tease apart various options).

It follows that there is no straightforward argument against  $[V_B]$  from the judgment that Barney has a virtuous true belief that doesn't amount to knowledge. "Straightforward" here means approximately as involved as the illustration discussed in Section I. It doesn't follow, however, that there is no more involved line of reasoning that recruits the judgment into an argument against  $[V_B]$ . For example, it could be that there is an alternative account of knowledge that has something explanatory to say about Fake Barn cases and that scores better than  $[V_B]$  when evaluated on various theoretical virtues, such as simplicity, scope, integration with other commitments, etc. The methodological lesson I draw is that thought experiments aimed at criticizing the explanatory sufficiency of an epistemic theory should be supplemented by more theoretical considerations.

There are other thought experiment judgments that do have the right form to be used in straightforward arguments against [V<sub>B</sub>]. Consider Jennifer Lackey's Chicago Visitor case.

Having just arrived at the train station in Chicago, Morris wishes to obtain directions to the Sears Tower. He looks around, approaches the first adult passerby that he sees, and asks how to get to his desired destination. The passerby, who happens to be a lifelong resident of Chicago and knows the city extraordinarily well, provides Morris with impeccable directions to the Sears Tower by telling him that it is located two blocks east of the train station. Morris unhesitatingly forms the corresponding true belief (Lackey 2009).

Suppose one judges that Morris knows but does not have a virtuous true belief that the Sears Tower is located two blocks east of the train station. Here is a straightforward argument against  $[V_B]$ .

- Step 1. It is possible that: someone knows that p but does not have a virtuous true belief that p because they stand to a proposition p as in the Chicago Visitor case.
- Step 2. It is possible that: someone knows that p but not because they believe truly that p due to their intellectual virtue.
- Step 3. So, it is not necessary that if someone knows that p, then they know that p because they believe truly that p due to their intellectual virtue.

Step 1 is the thought experiment judgment. Steps 2 and 3 clearly follow, and Step 3 is the denial of  $[V_B]$ . If successful, Lackey's thought experiment undermines the explanatory ambitions of  $[V_B]$ . Whether the thought experiment is indeed successful is a further question, which I am not able to address here. It is possible, however, to draw another general methodological lesson, namely that thought experiments aimed at criticizing the explanatory necessity of an epistemic

theory are less theoretically fraught than thought experiments aimed at criticizing its explanatory sufficiency.

Overall, thought experiment judgments best serve their purpose when their uptake is appropriately sensitive to the explanatory ambitions of epistemic theories. This may diminish their role relative to their outsized treatment in methodological discussions. But it doesn't imply that thought experiments should be relegated to non-evidential uses, such as illustrating ideas or posing problems. In my view, thought experiment judgments are in good epistemic standing, the issue to which I now turn.

# III. The Basis for Thought Experiment Judgments

Thought experiment judgments can be made for all sorts of reasons—e.g., intuitions, arguments, hunches, and testimony. I focus on the basis that I think should be central to discussions of their proper use. With this qualification understood throughout, the view I defend is that thought experiment judgments are based on impressions shaped by the target scenario, an epistemic theory, and deliberate cognition. Here is a picture:



On my preferred version of the view, impressions are intuitions, and intuitions are sui generis experiences that in certain conditions immediately justify making a judgment all on their own and in virtue of their phenomenology. I'm setting that aside for present purposes. I continue to assume that judgments are based on things other than themselves, namely impressions.

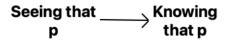
Thought experiment judgments are based on impressions because they are formed by endorsing the contents of those impressions. The diagram shows factors that determine the contents of such impressions, and awareness of which should be useful for methodology. So, returning to the Ford Case, the judgment about it expressed in [B] above will be based on a certain impression because it is formed by endorsing that impression's content, namely that it is possible that: someone has a justified true belief that p but does not know that p because they stand to a proposition p as in the Ford Case. Clearly, the Ford Case scenario plays a role in determining the content of this impression, since the impression is about the scenario.

Epistemic theories are another influence on thought experiment impressions. A large body of empirical research converges on the idea that psychological processing of theories about various domains factors into learning, judgment, and action with respect to those domains. These are often called intuitive theories, and among the most studied are intuitive physics, intuitive biology, and intuitive psychology. Researchers adopt a spectrum of views about the origins of intuitive theories, the contents of intuitive theories at various stages of development, and the formats in which those contents are represented and manipulated. There is broad agreement, however, that mature cognition is subserved by psychological structures that are appropriately conceived of as intuitive theories. Gopnik and Wellman (2012) helpfully illuminate the force of this conception by distinguishing three aspects of intuitive theories, namely their structure,

function, and dynamics. Structurally, intuitive theories represent causal information about their domains at various levels of abstraction. They include concepts and principles for entities whose existence and behavior are not open to direct observation. Functionally, theories are used to make predictions, formulate explanations, plan interventions, and draw counterfactual inferences. With respect to dynamics, intuitive theories change in ways that are rationally sensitive to evidence. Work on causal modeling and Bayesian inference have significantly improved the tractability of such changes. In all these ways, intuitive theories are like scientific theories. The main difference is that intuitive theories are psychological structures operative in individual minds, for the most part implicitly. Their structure, function, and dynamics are due to nature, not choice. It follows that if there are intuitive epistemic theories, then they are different from the epistemic theories considered in the previous section. The epistemic theories considered in the previous section are publicly debated cultural artifacts. Maybe the minds of some virtue epistemologists contain intuitive epistemic theories that match one or another virtue epistemology in content, but this is hardly a foregone conclusion.

The same kind of evidence adduced in favor of intuitive physical, biological, and psychological theories can be adduced in favor intuitive epistemic theories. Jennifer Nagel's work (2012) suggests one direct line of thought: intuitive psychological theories are, in part, theories of knowledge; so, the evidence for intuitive psychological theories is also evidence for intuitive epistemic theories. Fedyk, Kushnir, and Xu (2019) draw on multiple lines of research in support of the view that by the age of four children possess intuitive epistemic theories that include what they call a theory of evidence. The function of a theory of evidence is to make "accurate judgments about what sorts of events, effects, and occurrences do and do not count as evidence." One line of research they discuss focuses on children's abilities to identify, evaluate, and selectively rely on testimony (see also Harris et al 2018). This research demonstrates children's abilities to make epistemic evaluations of themselves in addition to others. For example, children make decisions about when to defer to and when to discount testimony that conflicts with their prior beliefs, and these judgments are sensitive to epistemically relevant factors. They discount statements that are obviously false, but they are willing to accept surprising statements (the Earth is round) when such statements are given a plausible backstory. If four-year-old children have intuitive epistemic theories, then likely mature philosophers do too.

The question remains, however, whether intuitive epistemic theories have the right content for shaping thought experiment impressions. Research on intuitive theories supports attributing causal-explanatory content to them, but I've suggested that thought experiment impressions have non-causal-explanatory content. The worry is that there is a gap between theory and impression. Suppose one's intuitive epistemic theory includes the following causal diagram, maybe as part of a larger structure:



It is clear how the information in this diagram can shape the impression that seeing Mr. Nogot in the office causes Lehrer to know that Mr. Nogot is in the office. But how can it shape the impression that Lehrer knows that Mr. Nogot is in the office in virtue of seeing him there, where this is a claim of non-causal dependence? A detailed answer must wait on empirical investigation of the matter. Here I'll give two reasons to think such an answer is possible.

First, it might be that intuitive epistemic theories explicitly represent non-causal-explanatory content in addition to causal-explanatory content. Suppose Jonathan Schaffer (2016) is correct that non-causal explanation is backed by a relation of grounding that is structurally like causation. Then the additional representational demands on intuitive epistemic theories would be minimal.

Second, thought experiment impressions are shaped by more than just scenarios and intuitive epistemic theories. Deliberate cognition plays a role too. Suppose one's intuitive epistemic theory represents the simple epistemic principle that if S sees that p, then S knows that p. The additional cognitive processes that go into generating a thought experiment impression might enrich this to the principle that if S sees that p, then S knows that p because S sees that p. Such an enrichment could be rationalized by explicitly held views that are not part of one's intuitive epistemic theory, such as general views about determinables, determinates, and non-causal dependence.

Deliberate cognition must play *some* role in determining the content of thought experiment impressions. Which epistemic properties of a scenario are represented in a thought experiment impression will be determined by the larger cognitive enterprise within which the thought experiment occurs. However, it is likely that deliberate cognition plays a larger role in determining the contents of impressions than merely picking out the properties to be applied in them. Deliberate cognition includes activities such as drawing distinctions, considering similarities and differences with other thought experiments, attending to relevant theoretical principles, maintaining coherence in one's commitments, etc. All of these can and should make a difference to the impressions one has in a thought experiment. For example, the quoted Ford Case is Lehrer's response to Charles Pailthorp's (1969) criticisms of Lehrer's (1965) earlier presentation of a similar thought experiment. The dispute between them comes down to whether you can have complete justification for believing that someone is F without inferring this from a proposition that says of some particular person, a, that a is F. Lehrer's discussion of the thought experiment includes an explanation of why you can, and in general fussing over this sort of thing is epistemological bread and butter. In my view, experimentally driven skepticism about "the method of cases" is largely founded on inadequate understating of the role deliberate cognition plays in shaping thought experiment impressions.<sup>6</sup>

As promised, the conception of thought experiments I've elaborated is organized around the idea of explanation. Thought experiment judgments are judgments about possible explanations, based on impressions shaped by explanatory content distributed between intuitive theories and deliberate cognition, and used to test the explanatory ambitions of publicly debated theories. Such use is inseparable from reasoning about broader theoretical considerations. This is especially so when thought experiments are marshaled in criticizing the explanatory sufficiency of an epistemic theory, less so when marshaled in criticizing the explanatory necessity of an epistemic theory. This suggests to me that the difficult questions about methodology in epistemology are less about our ability to discern the patterning of epistemic properties over non-epistemic properties, and more about our ability to make philosophical sense of those patterns.

#### References

<sup>&</sup>lt;sup>6</sup> Unfortunately, there isn't space to explore this issue here. I pursue the project, from a somewhat different angle, in (Chudnoff 2020).

Aronowitz, S., & Lombrozo, T. (2020). Learning through simulation. *Philosophers Imprint*, 20(1), 1-18.

Bliss, Ricki & Trogdon, Kelly. (2014). Metaphysical grounding. *Stanford Encyclopedia of Philosophy*.

Blouw, P., Buckwalter, W., & Turri, J. (2018). Gettier cases: A taxonomy. In *Explaining knowledge: New essays on the gettier problem* (pp. 242-252). Oxford University Press.

Chudnoff, Elijah. (2011). What Should a Theory of Knowledge Do?. Dialectica, 561-579.

Chudnoff, Elijah. (2013). *Intuition*. Oxford University Press, Oxford.

Chudnoff, Elijah. (2020). Forming Impressions: Expertise in Perception and Intuition. Oxford University Press, Oxford.

Fedyk, M., Kushnir, T., & Xu, F. (2019). Intuitive epistemology: Children's theory of evidence. In D. Wilkenfeld & R. Samuels (Eds.), *Advances in experimental philosophy of science*. New York, NY: Bloomsbury Academic Publishing.

Gardiner, Georgi. (2015). Normalcy and the Contents of Philosophical Judgements. *Inquiry*, 58(7–8), pp. 700–40.

Geddes, Alexander. (2018). Judgements about thought experiments. Mind, 127(505), 35-67.

Gendler, Tamar. S. (2004). Thought experiments rethought—and reperceived. *Philosophy of Science*, 71(5), 1152-1163.

Gendler, Tamar. S. (2007). Philosophical thought experiments, intuitions, and cognitive equilibrium. *Midwest studies in philosophy*, *31*, 68-89.

Gopnik, Alison., & Wellman, Henry. M. (2012). Reconstructing constructivism: Causal models, Bayesian learning mechanisms, and the theory theory. *Psychological bulletin*, *138*(6), 1085.

Harris, P. L., Koenig, M. A., Corriveau, K. H., & Jaswal, V. K. (2018). Cognitive foundations of learning from testimony. *Annual Review of Psychology*, 69, 251-273.

Horvath, Joachim. (2015). Thought experiments and experimental philosophy. In *The palgrave handbook of philosophical methods* (pp. 386-418). Palgrave Macmillan, London.

Ichikawa, Jonathan and Benjamin Jarvis. (2009): Thought-Experiment Intuitions and Truth in Fiction. *Philosophical Studies*, 142(2), pp. 221–46.

Lackey, Jennifer. (2009). Knowledge and credit. Philosophical studies, 142(1), 27-42.

Lehrer, Keith (1965). Knowledge, Truth and Evidence. Analysis 25 (5):168 - 175.

Lehrer, Keith. (1970). The fourth condition of knowledge: A defense. *The Review of Metaphysics*, 24(1), 122-128.

Malmgren, Anna-Sara. (2011). Rationalism and the Content of Intuitive Judgements. *Mind*, 120, pp. 263–327.

Nagel, Jennifer. (2012). Intuitions and experiments: A defense of the case method in epistemology. *Philosophy and Phenomenological Research*, 85(3), 495-527.

Nersessian, Nancy. J. (1992). In the theoretician's laboratory: Thought experimenting as mental modeling. In *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* (Vol. 1992, No. 2, pp. 291-301). Philosophy of Science Association.

Pailthorp, Charles. (1969). Knowledge as justified, true belief. *The Review of Metaphysics*, 23(1), 25-47.

Pritchard, Duncan. (2016). Epistemology: Second Edition, New York: Palgrave Macmillan.

Schaffer, Jonathan. (2016). Grounding in the Image of Causation. *Philosophical studies*, 173(1), 49-100.

Saint-Germier, Pierre. (2019). Getting Gettier straight: thought experiments, deviant realizations and default interpretations. *Synthese* 198 (2):1783-1806.

Turri, J., Alfano, M., & Greco, J. (2017). Virtue Epistemology. *Stanford Encyclopedia of Philosophy*.

Williamson, Timothy. (2005). Armchair Philosophy, Metaphysical Modality and Counterfactual Thinking. *Proceedings of the Aristotelian Society*, 105(1), pp. 1–23.

Williamson, Timothy. (2007). The Philosophy of Philosophy. Wiley-Blackwell