Philosophica 72 (2003) pp. 61-78

THE DUAL NATURE VIEW OF THOUGHT EXPERIMENTS

Tim De Mey[1]

ABSTRACT

It is often assumed that thought experiments are *either* experiments or arguments. In this paper, I argue that this disjunction is not an exclusive one and that we can only account for the demonstrative force or evidential significance of thought experiments by conceiving of them as *both* experiments and arguments. First, I distinguish between three related, but nevertheless distinct problems that thought experiments in physics pose. In this context, I also show that the choice between an *experiment view* and an *argument view* on the nature of thought experiments pops up in connection with two of those problems: (1) the problem of the source of thought-experimental knowledge and (2) the problem of the evidential significance of thought experiments. Subsequently, I argue that as far as the issue of *evidential significance* is concerned, we should at least recognise that thought experiments have a *dual structure*: an experiment-like set up and an argument-like winding up. Then I introduce Hans Radder's conceptual framework for the analysis of "real" experiments and I apply it to thought experiments. Finally, I argue that Radder's distinction between the aspects of 'theoretical description or interpretation" and "material realization" allows us to move on from the recognition of the dual structure of thought experiments to a full-blown *dual nature* view of them.

1. Introduction

One should always be careful with talk about "the nature of ...", because more often than not, quite different issues are lumped together into that

[1] The author is Postdoctoral Fellow of the Fund for Scientific Research - Flanders (Belgium) and works at the Centre for Logic and Philosophy of Science, Ghent University. Thanks to Michael Bishop, Sören Häggqvist and Hans Radder for informal discussions about the topic and to Maarten Van Dyck and Erik Weber for comments on earlier drafts of this paper.

broad category. The debate about thought experiments in physics nicely illustrates the risks. According to contemporary wisdom, one either holds that thought experiments are, basically, experiments or one subscribes to the rather deflationary view that they are, deep down, arguments.[2] But thought experiments in physics raise at least three distinct problems: (1) the source of thought-experimental knowledge, (2) the heuristic value of thought experiments and (3) the evidential significance of thought experiments. It is quite important to distinguish carefully between these issues.[3] For one often finds in the relevant literature that another view on thought experiments is taken to be a *rival* view and is then criticized accordingly, whereas it is in fact simply an answer to a different question.[4] But that is exactly what one risks if one claims to be able to reveal "the nature of thought experiments".

In this paper, I will develop a view on the nature of thought experiments in function of a solution to what I consider to be the most pressing problem among the three mentioned above, i.e. the problem of evidential significance.[5] I'm not going to solve that problem, I will just show that a first step towards a solution is to stop conceiving of thought experiments as *either* experiments or arguments. To that end, I will draw from Hans Radder's conceptual framework for the analysis of experiments (Radder 1998). I will use, more specifically, Radder's distinction between the two aspects of "material realization" and "theoretical description or interpretation" to make sense of the idea that thought experiments have a dual nature, i.e. that they are *both* experiments and arguments.

---

[2] See, e.g., McAllister (1996) and Bishop (1998).

[3] The fact that I distinguish between these issues doesn't imply that I consider them equally momentous. My aims are, firstly, to give some structure to the muddled impressions that the relevant literature might evoke, and, secondly, to give the views on the nature of thought experiments that differ from mine a fair trial.

[4] See, e.g., Miscevic's (1992) criticism of Brown (1991), Häggqvist's (1996) of Nersessian (1993) and Gendler's (1998) of Norton (1991; 1996).

[5] How does my view on the nature of thought experiments relate to the two other problems thought experiments in physics pose? Although I cannot argue for this in the present context, I believe (1) that the dual nature view *dissolves* the problem of the source of thought-experimental knowledge and (2) that it is *compatible* with Nersessian's (1993) solution to the problem of the heuristic value of thought experiments.

I will develop my argument as follows. First I distinguish between the three philosophical problems thought experiments in physics pose (section 2). Subsequently, I qualify an argument by Michael Bishop against the *argument view* of thought experiments (section 3). At that point, my aim is not to rescue the argument view, but rather to show that we should at least recognise that thought experiments have a *dual structure*: an experiment-like set up and an argument-like winding up (section 4). Finally, I move, on the basis of Radder's distinction (section 5), from the recognition of that dual structure to a full-blown dual nature view of thought experiments (section 6).

2. The three problems of thought experiments in physics

2.1 The problem of the source of thought-experimental knowledge.

A first problem is to offer a viable account of the *source* of the knowledge we obtain by thinking through thought experiments. To someone unfamiliar with epistemological problems, this might sound a bit odd. What else could be the source of thought experimental knowledge than thought experiments themselves? However, given any kind of knowledge, epistemologists always wonder whether, and if so, how, that kind of knowledge can be traced back to experience. Such questions are residues of the old, and partially cooked up debate between empiricists and rationalists: whereas empiricists take *experience* to be the ultimate basis of all of our knowledge, rationalists take *reason* to be far more important in grounding our knowledge than empiricism allows. In the contemporary form of this debate, the burden of proof is typically on the empiricist: it is up to her to trace the kind of knowledge at investigation back to experience. Rationalists, by contrast, confine themselves to pinpoint possible pitfalls or even irremediable lacunae in empiricist analyses.

Similarly, the very possibility of acquiring knowledge by means of thought experiments is generally taken to be a problem for empiricists.[6]

[6] See, e.g., the interaction between Brown (2004) and Norton (2004) on whether or not thought experiments "transcend" empiricism.

Obviously, one need not agree that such knowledge is possible.[7] But as soon as an empiricist accepts that possibility, she seems to be bound to find some way to trace thought experimental knowledge back to experience. Norton (1991) formulates the problem of the source of thought experimental knowledge and his empiricist solution to it as follows:

> Thought experiments in physics provide or purport to provide us information about the physical world. Since they are *thought* experiments rather than *physical* experiments, this information does not come from the reporting of new empirical data. Thus there is only one non-controversial source from which this information can come: it is elicited from information we already have by an identifiable argument, although that argument might not be laid out in detail in the statement of the thought experiment. The alternative to this view is to suppose that thought experiments provide some new and even mysterious route to knowledge of the physical world (Norton 1991: 129).

The alternative view to which Norton alludes is the rationalist one developed and defended by James Robert Brown.[8] By performing *some* thought experiments in physics, Brown claims, thought experimenters detect the relevant laws of nature through a kind of "non-sensory perception." He argues that a small number of thought experiments, which simultaneously destroy an old theory and generate a new one, are *a priori* in that they are not based on new empirical evidence, nor merely logically derived from old data (Brown 1991: 77). According to McAllister (1996), the view Brown develops is an *experiment* view.

Norton's, by contrast, is an *argument* view.[9] Thought experiments

---

[7] Alternatively, one can question whether there are *many* thought experiments that produce knowledge. E.g., according to Bunzl (1996), knowledge production is only under quite limited circumstances a goal of thought experiments.

[8] See Koyré (1939) for another rationalist view of thought experiments. Basically, Koyré argues that Galileo's experimental accounts in the *Two New Sciences* must be thought experiments, since their degree of precision looks unbelievable: they almost provide perfect confirmation of the theory of motion (Naylor 1989: 121).

[9] See Rescher (1991: 31), Irvine (1991: 150) and Forge (1991: 210) for similar views on the nature of thought experiments, as far as the problem of the source of thought experimental knowledge is concerned.

are not typically presented as straightforward arguments; many of them seem to have more in common with experiments than with arguments. But according to Norton, the features of thought experiments that make them experiment-like, are, strictly speaking, redundant. So, basically, the empiricist view of thought experiments, boils down to a kind of "elimination thesis:"

> Thought experiments are arguments which contain particulars irrelevant to the generality of the conclusion. Thus any conclusion reached by a good thought experiment will also be demonstrable by an argument which does not contain these particulars and therefore is not a thought experiment (Norton 1991: 131).[10]

2.2 The problem of the heuristic value of thought experiments.

Whatever the merits of Norton's argument view are, it does not explain why thought experiments can be "psychologically helpful" (and thereby rhetorically effective). One might maintain that it suffices for an empiricist to show that, in principle, every piece of thought experimental knowledge is also obtainable by means of a straightforward argument. But even then, we are still in need of theories that explain why the very same results can be obtained more *quickly* and *easily* by means of thought experiments.

The burden of accounts of the *heuristic value* of thought experiments is to explain our ability to think purposefully and fruitfully through them. That such philosophical theories draw heavily from adjacent evolutionary and psychological theories shouldn't surprise us.[11] After all, whereas the problem of the source of thought experimental knowledge is an issue in theoretical epistemology, the problem of the heuristic value of thought

---

[10] Gendler (2000: 34-9) notes that the elimination thesis is compatible with both a weaker reading and a stronger reading. On the former, a good scientific thought experiment can be replaced by an argument without loss of demonstrative force. On the latter, if a thought experiment has demonstrative force, it is because, deep down, the thought experiment is an argument.

[11] See, e.g., Miscevic (1992) and Nersessian (1993) for the *mental model* view. Bishop (1998) also invokes mental models, but he is not really dealing with the problem of the heuristic value of thought experiments. See also Sorensen (1992).

experiments is rather one in *naturalised epistemology*.

2.3 The problem of the evidential significance of thought experiments.

Thought experiments can play various roles and one legitimate way to approach them is to investigate the evolutionary and cognitive basis (or bases) of goal-directed modal reasoning. However, when we think about thought experiments, most of us spontaneously think of their most spectacular capacity, i.e., the role they can play in *theory choice*. Somehow, what would happen (or, for that matter, what we think would happen) if some imaginary state of affairs were to obtain, can speak for or against a thesis or theory. How on earth is that possible?

Norton's empiricist solution to the problem of the source of thought experimental knowledge can give us a first hint. Nobody doubts that *arguments* can play a role in theory choice. So, as far as thought experiments *are* arguments, their evidential significance seems fairly unproblematic.

Häggqvist (1996) adopts this view, at least an amended version of it.[12] McAllister (1996) holds, by contrast, that the most promising way to regard thought experiment is as a species of experiment, alongside concrete experiment. However, he remarks that where the discussion turns to the evidential significance of thought experiments, experimentalist accounts, such as those of Brown and Sorensen, diverge, strangely enough, from the most persuasive present-day accounts of concrete experiment:

> Brown and Sorensen hold to a logicist notion of evidential significance, according to which evidential significance is an intrinsic property of thought experiment. In contrast, the most convincing accounts of concrete experiments available today hold to a historicist notion of

[12] It is only fair to indicate that Häggvist's view is actually quite subtle and that what I call "the dual nature view" is inspired by it. On the one hand, Häggqvist insists that one cannot properly call thought experiments arguments, since thought experiments are not composed of truth-valued entities and cannot have the properties, such as validity, that arguments have. On the other hand, he points out that thought experiments can only have evidential significance through their connection with arguments, since "only arguments can matter when the truth-value of a scientific or philosophical theory or hypothesis is to be assessed" (1996: 87).

> evidential significance, according to which the evidential significance of concrete experiment is the outcome of historical and local accomplishments (McAllister 1996: 234).

Subsequently, McAllister suggests that the persuasive force of thought experiments is a conceptually and historically localised attainment. He argues that practitioners of Aristotelian mechanics couldn't accept Galileo's thought experiments as having relevance to establishing or discrediting claims in mechanics, since Aristotelian mechanics is a science of natural occurrences.[13]

Tamar Szabó Gendler (1998) argues, by contrast, that Galileo's famous thought experiment on the rate of fall of bodies of different weights is precisely designed to convince Aristotelians, thereby suggesting that thought experiments, or at least this one, possess evidential significance intrinsically. Gendler's argument is directed against Norton's elimination thesis. By analyzing Galileo's thought experiment she tries to show that it is impossible to replace the thought experiment by a straightforward argument without loss of *demonstrative force*. She contends, more specifically, that, confronted with the naked argument, the Aristotelian still disposes of several "ways out," all revolving around the special properties of bodies that are strapped together. However, Gendler argues, these ways out do not seem available when the thought experiment is presented in its original form. Contemplation of the case Galileo describes brings the Aristotelian to see, among other things, that "whether we consider a strapped-together body to be a single object, or two objects held together by a strap, or indefinitely many objects held together by internal forces, ... is a question about our words, not a question about the world" (Gendler 1998: 407).

3. Qualifying Bishop's argument against the argument view

In 'An Epistemological Role for Thought Experiments,' Michael Bishop (1998) raises an interesting objection to Norton's argument view of thought experiments. He discusses 'the clock-in-the-box episode': a failed

---

[13] McAllister's claim seems to be backed up by the fact that Aristotelians cited actual falls against Galileo's case (Shea 1972: 11).

attempt by Einstein to confront Bohr with a counterexample to the uncertainty principle. Now, according to Bishop, the clock-in-the-box episode shows that thought experiments cannot be arguments:

> The quick and dirty argument for why this conception of thought experiments cannot properly account for the clock-in-the-box episode is that Bohr and Einstein were analyzing a single thought experiment (the clock-in-the-box) but proposing distinct arguments, arguments with contradictory conclusions. The slowed and cleaner version of the argument begins with the commonplace that both real and thought experiments can be, and sometimes are, repeated. For this to be so, it must be possible for there to be different tokens of the same thought experiment type. Now, if thought experiments are arguments, then the distinction between thought experiment types and tokens will be made in terms of argument types and tokens. And here's the problem. Einstein and Bohr proposed different argument types; their argument had contradictory conclusions. But they were discussing and analyzing a single thought experiment type. So thought experiment types cannot be identified with argument types (Bishop 1998: 22).

After having developed this argument, Bishop specifies what the defender of the argument view should show to make sense of the clock-in-the-box episode in terms of her account of thought experiments:

> Either Einstein and Bohr were dealing with only one argument/thought experiment type or they were dealing with two argument/thought experiment types. Neither of these options is very plausible (Bishop 1998: 22-3).

At this point, I would like to make a qualification, since I happen to believe that *in a sense both* options are plausible. If you conceive of thought experiments as arguments, you should (and could) analyse the clock-in-the-box episode in terms of two argument/thought experiment types. However, if you conceive of thought experiments as experiments, you should (and could) describe this episode in terms of a single thought experiment that gives rise to a single argument. To clarify and establish this, it suffices to tell the points on which the antagonists in the episode, i.e. Einstein and Bohr, agree, from the points on which they disagree.

3.1 The argument view of the clock-in-the-box

Let us first subscribe, for the sake of the argument, to the *argument view* of thought experiments. I agree with Bishop that we have then two arguments here. On the one hand, Einstein argues that the clock-in-the-box speaks against the uncertainty principle. On the other hand, Bohr shows that it doesn't. Now it is up to me to show that these two argument types can be taken to correspond with two thought experiment types. As a matter of fact, that's not so difficult. The only thing the two thought experiments (i.e., the one performed by Einstein and the other performed by Bohr) share is the set up, i.e. the initial imaginary situation: a box full of photons and a shutter mechanism on one of its walls that opens for a brief interval at which time a single photon escapes. However, Einstein's and Bohr's respective thought experiments differ in the settlement or winding up of that initial imaginary situation. Einstein reasons: (1) we weigh the box before and after a single photon escapes, (2) the change in the weight gives us the weight of the photon, thereby (3) its mass, and using $E = mc^2$ also (4) its energy. Bohr, by contrast, notes (1) that weighing the box will involve the motion of the clock-in-the-box apparatus in a gravitational field and thereby shows (2) that there is a fundamental limit to the accuracy to which any clock-in-the-box apparatus can measure a photon's weight.

So when performing his thought experiment, i.e. the one supporting the premises of an argument *against* the uncertainty principle, Einstein simply failed to imagine or to take into account all the relevant factors. And when thinking his thought experiment through, i.e. the one which dismisses Einstein's counterexample to the uncertainty principle, Bohr was successful in pinpointing the factor that would make a difference in the settlement or winding up of the initial imaginary situation. To put it "quick and dirty": just as we have a successful and an unsuccessful argument, we have a successful and an unsuccessful thought experiment here.

3.2 The experiment view of the clock-in-the-box

Nevertheless, in another sense we are dealing with one argument/thought experiment type. Let us consider the points on which the antagonists in the clock-in-the-box episode, agree. In his description of the episode,

Bishop quotes the following description of Bohr's reaction when Einstein presented him the clock-in-the-box counterexample to the uncertainty principle during the 1930 Solvay Conference on magnetism:

> It was quite a shock for Bohr ... he did not see the solution at once. During the whole evening he was extremely unhappy, going from one to the other and trying to persuade them that it couldn't be true, that it would be the end of physics if Einstein were right; but he couldn't produce any refutation. I shall never forget the vision of the two antagonists leaving the club: Einstein a tall majestic figure, walking quietly, with a somewhat ironical smile, and Bohr trotting near him, very excited ... The next morning came Bohr's triumph (Rosenfeld, quoted in Bishop 1998: 21).

So Bohr was extremely distressed by Einstein's putative counterexample. Now, it wouldn't make sense to be distressed by a putative counterexample, let alone to work all night to «refute» it, unless you agree that the experiment is properly designed, i.e. that its results will speak for or against the hypothesis or theory at issue. What the thought experiments of Einstein and Bohr have in common then is the set-up, i.e. the initial imaginary situation (which is agreed to be relevant for the theory at issue). So there is a consensus about experimental design and in that sense we are dealing with one experiment type, albeit one that is performed in "the laboratory of the mind."

Moreover, there is also one argument type in a sense, because both Einstein and Bohr clearly presume that *there is a fact of the matter*, which determines how the initial imaginary situation at consideration will settle or wind up. Again, without this assumption there would be no point for Bohr to work all night to refute it, nor for Einstein to give in as soon as Bohr had spelled out the factor which Einstein had overlooked.

4. The dual structure of thought experiments

Perhaps my argument will have inconvenienced the reader by now. I have analysed, *pace* Bishop, the clock-in-the-box episode as involving one argument/thought experiment type and as involving two argument/thought experiment types. Moreover, I used an expression that is bound to inconvenience readers with some background in philosophy, namely "in a sense". However, the underlying ideas are quite simple. Firstly,

thought experiments like that of the clock-in-the-box have a dual structure: they involve (1) the description of an imaginary situation and (2) the description of its settlement or winding up. Secondly, on the experiment view of thought experiments, sameness of or difference between thought experiments is identified on the basis of (1). On the argument view of thought experiments, by contrast, sameness or difference between thought experiments is identified on the basis of (2). (And Bishop's analysis blurs (1) and (2), which is okay with me, but only makes sense on what I will call the dual nature view of thought experiments.) And thirdly, a consensus with respect to (1) doesn't guarantee, a consensus with respect to (2). Although there is (supposed to be) a fact of the matter determining what would happen, thought experimenters are not necessarily successful in finding out what would happen.

At this point, we need to differentiate between two meanings of "evidential significance". A thought experiment with a dual structure acquires *evidential significance*$_1$ as soon as the antagonists in a debate agree that the settlement of the imaginary situation will speak for or against the hypothesis or theory at issue. So in that sense thought experiments are *experiment-like*: like many "real" or "ordinary" experiments, they are designed and used to test specific hypotheses or theories. Moreover, the results will only effectively speak for or against the hypothesis or theory at issue on the condition that there was, prior to its performance, a consensus about its design.

However, the *evidential significance*$_2$ that a thought experiment has, i.e. whether it *does* speak for or rather against a hypothesis or theory at issue, depends on what would happen if the imaginary situation were to obtain. Now any thought experimenter can be successful or unsuccessful in finding out what would happen. It requires one to take all the relevant factors into account, and it often happens that a thought experimenter overlooks such a factor when thinking through a thought experiment.

So my suggestion is that the settlement or winding up of the imaginary situation is what is "argument-like" about thought experiments with a dual structure that play a role in theory choice. After all, it corresponds to the reading of and interpretation of the results of an ordinary experiment.

In the next section, I introduce Hans Radder's conceptual framework for the analysis of experiments. This will not only allow me to make

more subtle distinctions, but also to move from the dual structure view for which I have been arguing up to now, to a full-blown dual nature view.

5. Hans Radder on experimentation

According to a somewhat outdated conception, experimenting is a matter of asking nature a question, which, if formulated aptly, nature then conclusively answers. On this conception "ordinary" or "real" experiments have a dual *structure* too, i.e. question and answer.

Sophisticated conceptions of experimentation, typically invoke more structural dimensions. Radder (1996), e.g., differentiates between three "phases" of experimentation: (1) preparation, (2) interaction and (3) detection. Firstly, during *preparation* the object and the apparatus are prepared in agreement with the plan of the experiment. Subsequently, *interaction* results in the transfer of information from the object to the apparatus. Finally, *detection* involves obtaining the information by measuring or observing the relevant property of the apparatus (Radder 1996: 11).

Furthermore, Radder tells (a) the *theoretical description or interpretation* of each phase from (b) its *material realization*. Although one can thus differentiate between the *interpretations* of each phase taken separately, the overall description of the experimental process consists out of *an argument* by which the (intended) theoretical result q is inferred from the premises p. Radder assumes that although such argumentation is not always *explicit* in experimental practice, "if asked the experimenter will or should be able to come up with a plausible story about how the experimental result is produced" (1996: 12).

For my purposes, two points that Radder makes with respect to the aspect of theoretical description are crucial. Firstly, he stresses that the theoretical description not only refers to preparation, interaction and detection, but also "to the screening and control of potential disturbances from the outside, that is, to the closedness of the experimental system" (1996: 12). At face value, it is redundant to single this aspect out as an extra component of the theoretical description of an experimental process. After all, ruling out artefacts is one of the main tasks during preparation. However, even if an experiment is carefully prepared, *in practice*

artefacts might pop up during interaction and even detection. Since the closedness of the experimental system should be checked (or, for that matter, is presumed) during all three phases of the experimental process, it deserves to be mentioned separately among the components of the theoretical description.

Secondly, Radder insists that the theoretical description is not a complete account of the experimental process, but only considers "those aspects of the experimental process that are deemed *relevant* to obtaining the intended result" (1996: 12).

For my purposes, it is also useful to discuss how Radder introduces the notion of "material realization". As a complement of "theoretical descriptions", material realization is supposed to be *theory independent*. Here is the "thought experiment" on the basis of which Radder argues that we need such an abstraction:

> Suppose we want to determine experimentally the mass of an object that is at rest in relation to the measurement equipment. Two scientists each carry out such an experiment in the same way. Nevertheless, one interprets the actions performed as a measurement of the Newtonian mass; and the other, as a determination of the Einsteinian mass. But both performed "the same" actions and thus –in a certain sense- the same experiment. Therefore, if we want to describe experimental action unambiguously, we have to find some sort of abstraction of these various specific theoretical interpretations. Yet, it is an indisputable fact that concrete experimental action is always action on the basis of certain theoretical ideas: without theoretical ideas there can be no experiments (Radder 1996: 13).

So Radder introduces the theoretical-philosophical notion of "material realization" to describe sameness of action in spite of difference of interpretation, which allows him, in turn, to define the *reproducibility* of the material realization of an experiment under different theoretical interpretations. What the notion does then is "to explicate the mechanism through which theoretical interpretations are underdetermined, not by data or empirical statements, but by the process of experimental action and production" (1996: 20).

6. Applying Radder's framework to thought experiments

Apart from the reproducibility of the material realization of an experiment, Radder (1996) also discusses the reproducibility of an experiment under a fixed theoretical interpretation and the reproducibility of the result of an experiment. The latter is more commonly known as a *replication* and doesn't seem very relevant for our purposes. The former, however, almost sounds as a criterion for good thought experiments in physics. After all, as Radder also notes (1996: 16, 18), obtaining reproducibility is always a significant achievement and it is traditionally assumed to be the hallmark of successful experimentation. Now since thought experiments are not realised materially, what else could their reproducibility involve than reproducibility under a fixed theoretical interpretation?

Such a suggestion, however, would be premature. Let us first look at the premise that thought experiments are not realised materially. There is, of course, a sense (sorry again for this terminology) of material realisation in which this claim is trivially false: even a dualist would agree that thought experiments are materially realized "in the head" of the thought experimenter. But there is also a sense in which the claim is trivially true: we would take the analogy between experiments and thought experiments much too far, if we say that in the case of the latter there is "suitable interaction" between objects and apparatuses.

The question we have to answer, however, is whether we can make sense of Radder's theoretical-philosophical notion of "material realization" in the case of thought experiments. So is sameness of action in spite of difference of interpretation possible in the case of thought experiments?

Obviously, our answer to this question depends, in part, on how exactly we tell "action" from "interpretation" in the case of thought experiments. We might be tempted, more specifically, to consider the possibility of sameness of action and difference of interpretation for at least two classes of thought experiments, i.e. in Gendler's terminology, conceptual and valuational thought experiments (2000: 25). In the former case, the question is how *we should describe* what would happen were some imaginary state of affairs to obtain. Valuational thought experiments in turn ask *how we should evaluate* what would happen were some imaginary state of affairs to obtain. In both cases, then, it is possible in

principle (and in practice) that there is agreement on the thought-experimental facts, i.e. on *what would happen* (note that in many cases of conceptual and valuational thought experiments *what would happen* is even stipulated), but disagreement on the proper application of concepts or the proper moral or aesthetic response. So, as far as conceiving of some imaginary situation (and of its winding up) can be properly called the "action" involved in performing such a thought experiment, and as far as our answers to conceptual or valuational questions rather belongs to its "interpretation", it makes sense to speak of "material realization" in Radder's sense.

However, the same idea can also be expressed in terms of the phases of an experimental process. If a conceptual or valuational thought experiment is to be reproduced, there should at least be a consensus about what Radder calls *preparation* (the set up), but also about *interaction*.

Now the question is whether there is such a way to differentiate between action and interpretation in the case of what Gendler calls factive thought experiments (Gendler 2000: 25), i.e. thought experiments, which really put the question (and not merely stipulate) what would happen were some imaginary state of affairs to obtain. I think it is fair to say that what would happen were some imaginary state of affairs to obtain corresponds to what Radder calls the "interaction" phase of an experimental process. The clock-in-the-box episode nicely illustrates that. The consensus was restricted to the experimental design, or, in Radder's terminology, to the *preparation*. Einstein and Bohr differed though on what would happen. Or better: since there is a fact of the matter which determines what would happen were such an imaginary situation to obtain, they differed in what they *thought* or how they *interpreted* what would happen. So in the case of factive thought experiments it seems that the description of the material realization is restricted to the preparation-phase of the experimental process and that the theoretical description or interpretation concerns both interaction and detection.[14]

[14] Here is, perhaps, the ground, for Norton's elimination thesis: you can have the argument, i.e., the theoretical description or interpretation of the phases of interaction and detection, without the fancy stuff that goes into the set up of the thought experiment, i.e. the description of the material realization of the preparation.

7. The dual nature view of thought experiments

According to Hans Radder, there are two aspects to each phase of an experimental process: (1) "material realization" and (2) "interpretation". Material realization is an abstraction: it refers to *theory independent* actions. Its description is phrased in common language, so that even a layperson in the field can "realize" the experiment "materially". Interpretation, by contrast, refers to a *theory dependent* description of the experimental process, i.e. an *argument* in which the outcome of the experiment is inferred from the premises (which involve subdescriptions of the object, the apparatus, the interaction between them, the interaction between the experimental system and its external setting, etc.).

Similarly, factual, conceptual and valuational thought experiments *necessarily* have two aspects: (1) "material realization" and (2) "interpretation". Typically the actions to be performed by the thought experimenter are phrased in common language. More importantly, a thought experiment will only acquire *evidential significance*$_1$ if at least its set up (Radder's *preparation*) and in the case of conceptual and valuational thought experiments also what Radder calls its *interaction*, can be described *independently of theory* (or at the very least in a *theory-neutral* way). The *detection* of the outcome, however, the *evidential significance*$_2$ a thought experiment turns out to have, can only by described theoretically, i.e. in terms of an argument.

So there are two ways to describe a thought-experimental process, i.e. in terms of its material realization and in terms of its interpretation. To do full justice to the evidential significance of a thought experiment then, of whatever kind it is, we need both. As far as the material realization of the thought-experimental process is what adherents of the experiment view on the nature of thought experiments stress (as I believe they do) and as far as the theoretical interpretation of the thought-experimental process is what adherents of the argument view have in mind (as I believe they do), we can safely conclude that it doesn't make sense to say that thought experiments are, basically, either experiments are arguments. They are, deep down, both experiments and arguments.

Universiteit Gent

REFERENCES

Bishop, Michael A. (1998), 'An Epistemological Role for Thought Experiments' in Shanks (1998), pp. 19-33.

Brown, James Robert (1991), *The Laboratory of the Mind: Thought Experiments in the Natural Sciences*, New York and London: Routledge.

Brown, James Robert (2004), 'Why Thought Experiments Transcend Empirism' in Hitchcock (2004), pp. 23-43.

Bunzl, Martin (1996), 'The Logic of Thought Experiments', *Synthese* **106**, pp. 227-240.

Gendler, Tamar Szabó (1998), 'Galileo and the Indispensability of Scientific Thought Experiment', *British Journal for the Philosophy of Science* **49**, pp. 397-424.

Gendler, Tamar Szabó (2000), *Thought Experiment: On the Powers and Limits of Imaginary Cases*, New York and London: Garland Publishing, Inc.

Gooding, David C., Pinch, T. and Schaffer, S. (eds.) (1989), *The Uses of Experiment: Studies in the Natural Sciences*, Cambridge: Cambridge University Press.

Häggqvist, Sören (1996), *Thought Experiments in Philosophy*, Stockholm: Almqvist & Wiksell International.

Hitchcock, Christopher (ed.) (2004), *Contemporary Debates in Philosophy of Science*, Oxford: Blackwell Publishing.

Horowitz, Tamara and Gerald Massey (eds.) (1991), *Thought Experiments in Science and Philosophy*, Savage, MD: Rowman and Littlefield Publishers.

Irvine, Andrew D. (1991), 'On the Nature of Thought Experiments in Scientific Reasoning', in Horowitz and Massey (1991), pp. 149-165.

Janis, Allen I. (1991), 'Can Thought Experiments Fail?', in Horowitz and Massey (1991), pp. 113-118.

Koyré, Alexandre (1968), *Metaphysics and Measurement*, London: Chapman and Hall.

McAllister, James W. (1996), 'The Evidential Significance of Thought Experiments in Science', *Studies in History and Philosophy of Science* **27**, pp. 233-250.

Miscevic, Nenad (1992), 'Mental Models and Thought Experiments', *International Studies in the Philosophy of Science* **6**, pp. 215-226.

Naylor, R.H. (1989), 'Galileo's Experimental Discourse', in Gooding, Pinch and Schaffer (1989), pp. 117-134.

Nersessian, Nancy (1993), 'In the Theoretician's Laboratory: Thought Experiment as Mental Modeling', *Proceedings of the Philosophy of Science Association*, Vol. 2, pp. 291-301.

Norton, John (1991), 'Thought Experiments in Einstein's Work', in Horowitz

and Massey (1991), pp. 129-148.
Norton, John (1996), 'Are Thought Experiments Just What You Thought?', *Canadian Journal of Philosophy* **26**, pp. 333-366.
Norton, John (2004), 'Why Thought Experiments do not Transcend Empiricism', in Hitchcock (2004), pp. 44-66.
Radder, Hans (1996), *In and About the World: Philosophical Studies of Science and Technology*, Albany, N.Y.: State University of New York Press.
Rescher, Nicholas (1991), 'Thought Experiment in Pre-Socratic Philosophy', in Horowitz and Massey (1991), pp. 31-41.
Shanks, Niall (ed.) (1998), *Idealization IX: Idealization in Contemporary Physics*, Amsterdam: Rodopi.
Shea, W.R. (1972), *Galileo's Intellectual Revolution*, London: Macmillan.
Sorensen, Roy (1991), *Thought Experiments*, New York and Oxford: Oxford University Press.