Poznań Studies in the Philosophy
of the Sciences and the Humanities
1998, vol. 63, pp. 19-33

Michael Bishop

AN EPISTEMOLOGICAL ROLE FOR THOUGHT EXPERIMENTS

ABSTRACT. This paper has three main sections: (1) A thought experiment is a mental representations (a mental model) of an experiment in which a result is derived by a process of reasoning that employs substantive assumptions about how the world works (Nersissian 1993). I defend this account on the grounds of utility – it fits smoothly into our best understandings of the historical trajectory of actual thought experiments and it offers insights about how thought experiments work. (2) Thought experiments can test how unified a scientific theory is. They can play a role in rational theory choice if unification is a theoretical virtue. (3) Ironically, the use of unactualized thought experiments to test a theory typically betrays a commitment to a kind of realism which is incompatible with van Fraassen's (1980) views about the aim of science.

1. Introduction

Why should a thought experiment, an experiment that only exists in people's minds, alter our fundamental beliefs about reality? After all, isn't reasoning from the imaginary to the real a sign of psychosis? A historical survey of how thought experiments have shaped our physical laws might lead one to believe that it's not the case that the laws of physics lie – it's that they don't even pretend to tell the truth.

My aim in this paper is to defend an account of thought experiments that fits smoothly into our understandings of the historical trajectory of actual thought experiments and that explains how any *rational* person could allow an imagined, unrealized (or unrealizable) situation to change their conception of the universe. In section 2, I will argue that thought experiments are mental representations (or models) of experiments. But unlike the results in real experiments, thought experimental results are derived by a process of reasoning about the mental model that employs substantive assumptions about how the world works (see Nersessian 1993). In section 3, I address the epistemological quandary raised by the prevalence of thought experiments in science. Thought experiments can play a role in rational theory choice by testing how unified a scientific theory is. (For a similar view, see Lennox 1991, p. 223. We part ways because Lennox believes a thought experiment cannot confirm a scientific law.) So if *unification* (or explanatory power) is an important ingredient in rational theory choice, then it will sometimes be rational for a thought experiment to convince people to believe or disbelieve a scientific theory. In the final section, I will attempt to draw a surprising lesson from the prevalence of idealized thought experiments in 20th century physics.

2. The nature of thought experiments

A philosophical account of thought experiments is worthwhile only insofar as it is useful. A useful account of thought experiments must at the very least fit smoothly into our best understandings of how thought experiments actually work. If it can offer insight about how or why they work the way they do, so much the better. If this is the goal, then our best account of thought experiments is disappointingly prosaic: A thought experiment is an experiment that is represented in someone's mind. It is a *thought* (*about*) *experiment*. Nancy Nersessian describes the view nicely: "The original thought experiment is the construction of a dynamical model in the mind by the scientist who imagines a sequence of events and processes and infers outcomes" (1993, p. 292). So thought experiments are not unique kinds of creatures. They are simply a subset of a much broader class of mental events in which mental models play significant roles. To see whether this account of thought experiments performs as advertised, let's briefly sketch a real thought experiment in action.

2.1. The clock-in-the-box thought experiment

In 1927, Werner Heisenberg introduced the uncertainty principle that bears his name. The principle says that there is a fundamental limit to the accuracy to which pairs of conjugate variables, p and q (variable pairs like position – momentum, and energy – time), can be measured (and hence known). This relation is described by the following equation,

$$\Delta p \times \Delta q > h$$

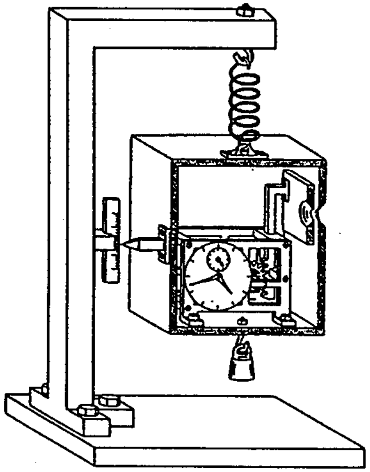
where Δp and Δq stand for the uncertainty in the measurement of the respective variables and h is a constant (Plank's constant divided by 2π).

During the 1930 Solvay Conference on magnetism, Einstein presented Neils Bohr with the clock-in-the-box counterexample to the uncertainty principle: Suppose we have a box full of photons that has, on one of its walls, a shutter that is controlled by a clock. Weigh the box. Now set up the shutter mechanism so that it opens for a brief interval at which time a

single photon escapes. Weigh the box again. The change in the weight of the box gives us the weight of the photon, which gives us its mass. And using Einstein's famous equation, $E = mc^2$, we can determine the photon's energy. In principle, therefore, we can measure the photon's energy and its time of passage to any arbitrary degree of accuracy. So on the basis of this thought experiment, Einstein concluded that Heisenberg's uncertainty principle is false. L. Rosenfeld describes Bohr's reaction.

> 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 Pais, pp. 446-447).

Bohr's "triumph" consisted of focusing on the practical instruments and procedures that one would have to use in order to measure the photon's energy at a particular time. This is why Bohr (1949, p. 227) illustrated the thought experiment (figure 1, reproduced with permission) in such a realistic way – note particularly that weighing the box will involve the motion of the clock-in-the-box apparatus in a gravitational field.



Bohr showed that there was a fundamental limit to the accuracy to which any clock-in-the-box apparatus could measure a photon's weight (and hence energy) at a particular time, given by Heisenberg's formula.

$$\Delta E \times \Delta T > h$$

I'll leave the details of Bohr's argument for the *Appendix*. But it is clear that as a result of Bohr's argument, Einstein abandoned this attempt to undermine the uncertainty principle.

2.2. *Troubles with the argument view of thought experiments*

A number of contemporary philosophers hold that thought experiments are simply *arguments*, e.g., Nicholas Rescher (1991, p. 31), Andrew D. Irvine (1991, p. 150), John Forge (1991, p. 210). John Norton presents a sophisticated account and defense of this sort of view.

> Thought experiments are arguments which: (i) posit hypothetical or counterfactual states of affairs, and (ii) invoke particulars irrelevant to the generality of the conclusion . . . 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 (1991, p. 129).

Of course, it is possible to construct an argument on the basis of anything real or imagined, including experiments. Both Lennox (1991) and Nersessian (1993) have pointed out that thought experiments are no more arguments than are real experiments. While this objection might have intuitive appeal, I want to argue that the real problem with this view is that it has trouble making sense of the historical trajectory of thought experiments, like the clock-in-the-box.

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 slower 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 arguments had contradictory conclusions. But they were discussing and analyzing a single thought experiment type. So thought experiment types cannot be identified with argument types.

In order to make sense of the clock-in-the-box episode, the defender of the argument view of thought experiments must show one of two things. 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. Rather than pursue these options, however, I want to set in sharp relief some of the benefits of the thought-about-experiment view of thought experiments.

2.3. *The virtues of the mental model view of thought experiments*

A thought experiment is a mental representation of an experimental process, in which the outcome of the process is inferred. Those inferences will be based, implicitly or explicitly, on general views about how nature operates. So a reasoning process is essential to get thought experimental results. (This may be why it is natural to identify thought experiments with arguments.) But on occasion, there are quite different kinds of argument associated with thought experiments. When a thought experiment is supposed to convince people to believe or disbelieve a particular theory, there is an argument (perhaps implicit) from the thought experimental result to a conclusion regarding the status of that scientific theory. So while arguments should not be identified with thought experiments, arguments do play at least two important roles in how they work.

One reason to accept this view is that it provides us with a coherent way to understand the actual historical trajectory of thought experiments. It offers a smooth account of the clock-in-the-box episode. Einstein and Bohr were analyzing the same thought experiment because they had the same experiment (type) in mind. Where they differed was in the arguments they proposed. Einstein and Bohr reasoned to different experimental results, and as a consequence, they argued to very different conclusions about the status of the uncertainty principle. Notice that by treating thought experiments *as* experiments, this view illuminates and explains the fact that thought experiments share certain characteristic features with real experiments. These similarities include:

(1) Different tokens of an experiment type will differ in their particulars, sometimes in significant ways. For example, Bohr's clock-in-the-box thought experiment included prosaic details (see figure 1 and *Appendix*) about the ways in which the time and energy measurements would be made that Einstein seems to have ignored. The history of the Michelson-Morley experiment is a fine example of a real experiment-type whose tokens differed considerably. This experiment used an interferometer to find the relative speed of the Earth moving through the optical ether, the medium in which light was thought to propagate. Michelson first tried this experiment in 1881, using an interferometer with 1,000 mm. brass arms (Swenson 1972, pp. 71-73). The famous 1887 trial by Michelson and Morley used a much larger interferometer, placed on a massive sandstone

block which floated in 200 pounds of mercury (Swenson 1972, pp. 90-94). In 1926, the Frenchmen Piccard and Stahel repeated the experiment by releasing a small interferometer in a balloon (Swenson 1972, pp. 214-215). In Jena four years later, a large, automated interferometer was run for a full year (Swenson 1972, p. 225). In each of these experiments, the very small fringe shifts were explained by experimental error.

(2) Different tokens of the same experiment type can have different results. As we have seen, the clock-in-the-box experiments of Einstein and Bohr had very different results. And in 1925, Dayton Miller "equipped with a much more highly refined [interferometer] than ever before" (Swenson 1972, p. 205), announced a result that was at odds with almost every other well-documented token of the Michelson-Morley experiment: that the Earth moves at about 9 km/sec relative to the ether, which is about 1/3 of its orbital velocity (Swenson 1972, pp. 209-12). Of course, thought and real experimental results are obtained in very different ways. Einstein and Bohr, unlike Michelson, Morley, *et al.*, did not have to set up equipment and "run" their experiment.

(3) Different tokens of an experiment can lead scientists to draw different conclusions about the status of scientific theories or laws, especially when those tokens have different results. As we have seen, Einstein (initially) and Bohr drew different conclusions about the uncertainty principle from their different thought experimental results. And through the 1930's, Dayton Miller continued to adduce his 1925 experiment in defense of the existence of the optical ether, even though by this time, the vast majority of physicists had accepted the "null result" and rejected the ether (Swenson 1972, p. 226).

There are surely other ways in which thought and real experiments are similar. For example, both play important pedagogical roles in science. Given these and other likely similarities, we might wonder why thought experiments function in many ways like real experiments. The mentalistic view, unlike its competitors, has a ready answer: Because they are a lot like real experiments – except that they happen in the mind.

2.4. Potential objections

One might object that this account of thought experiments is too broad because it does not draw conceptual boundaries for intuitively difficult cases. For example, suppose a scientist develops a perfectly practicable experiment but has not yet performed it. Is that a thought experiment? Or what if the experiment could be performed but never actually is? Or what if the experiment the scientist has in mind is not practicable at the time, but can be approximated closely? Or somewhat closely? Or not particularly closely? Or what if the experiment can be performed, but only in the future – near, far, or somewhere in between? I propose to ignore all these tricky questions for three reasons. First, I have no settled pretheoretical intuitions about the hard cases. Second, trying to legislate conceptual boundaries at this point in our thinking about thought experiments is a sucker's game (for an argument, see Bishop 1992). And third, I doubt that anything of philosophical interest will turn on whether we decide to call a particular tricky case a "thought experiment" or not. In order to get a sensible discussion started, all we need is to agree about the paradigm cases, e.g., Galileo on falling bodies, Newton's bucket, Einstein's train, Schroedinger's cat.

A more serious objection is that this view counts mental models of *actualized* experiments as thought experiments. Many believe that this is a mistake (Lennox 1991, Rescher 1991, Irvine 1991, Forge 1991, Norton 1991). Why might they believe this? When we ponder thought experiments, what comes to mind are those thought experiments that played dramatic roles in the history of science. They are dramatic and mysterious because they moved (or were meant to move) people to accept, reject or revise theories on the basis of an unactualized, imagined experiment. Mental models of actualized experiments can't be dramatic in the same way. But my goal here is to give an account of thought experiments, not *dramatic* thought experiments. Others might have different goals, and I am already on record as being disinclined to waste energy fighting about how to use the term "thought experiment." But there is a theoretical reason to not restrict our conception of thought experiments to just the dramatic cases. To see why, let's ask why Einstein didn't just begin a thought experiment as follows, "Imagine that we have an instrument that can measure a particle's precise energy at a precise time..." Of course, this thought experiment would not have been taken seriously. In order for a thought experiment to have some chance of moving people, it must observe reasonably well-defined (and probably largely implicit) conventions. These conventions govern how legitimate idealizations work, precisely which real world complexities can be altered (and how) and which can be ignored. Where do these conventions come from? The answer, I think, is that scientists digest them while working through those common, undramatic idealized representations that confront them as soon as they begin learning and working with a theory. These idealizations – e.g., the problems that fill textbooks – perform their pedagogical service whether they are idealizations of real or just hypothetical situations. The reason to not restrict our account of thought experiments only to the dramatic cases is that we run the risk of type-casting thought experiments as rare, anomalous, hopeful monsters. But they aren't. They're perfectly conventional, everyday sorts of things. It's just that sometimes they can be used in stunning and spectacular ways.

Niall Shanks (in correspondence) has raised a very different objection against this liberal account of thought experiments: It is actually not liberal enough. On the plausible assumption that computers can't think (or at least not yet), this view neglects the possibility that a thought experiment might be represented by a computer. A computer simulation could play much the same kind of role in a scientific episode as an experiment represented in someone's mind. I think this is reason enough to broaden the account to potentially include computer simulations: If something is a thought experiment, it is an experiment that is represented by some sort of representational system, human or otherwise. So certain computer representations might be thought experiments, even if it is a mistake to suppose that a computer has a mind or that it thinks. Those who are squeamish about thought experiments without thoughts may use scare quotes: Certain computer simulations would be "thought experiments."

3. *An epistemological role for thought experiments*

To be told that thought experiments are a certain kind of imaginary fantasy might be informative, but it makes their epistemological mystery just that much more pressing. Some thought experiments convince (or are meant to convince) scientists to believe or disbelieve some theory about the real world. But why should an *imaginary fantasy* of *any* sort alter our fundamental beliefs about reality? Paul Feyerabend argues that thought experiments (or at least Galileo's thought experiments) convince people *via* propaganda and psychological trickery (1978). One of my goals in this section is to argue that Feyerabend is wrong. While thought experiments can be powerful rhetorical and persuasive devices, this is consistent with their having important roles in the rational assessment of scientific theories.

Is there some recognizable notion of rationality that would respect and account for the use of thought experiments in coming to believe theories about the world? If *unification* is an important ingredient in rational theory choice, then it is sometimes rational for thought experiments to convince people to believe or disbelieve a scientific theory. Of course, other standards of rational theory choice might make room for thought experiments as well. But to solve the central mystery about thought experiments, it is enough to show that *some* recognizable standard of rational theory choice could make room for them. While I am not making any historical claims about whether anyone actually embraced unification as a scientific virtue (so the "*A priori* history of science!" objections are premature), I do think that some historical figures have embraced something like the ideal of unification (see Lennox 1991). Nonetheless, my contention is a conditional one: *if* unification is a theoretical virtue, then thought experiments can sometimes play a legitimate role in rational theory choice.

What do thought experiments do? A thought experiment, like a real experiment, is a test of a theory or law. But thought experiments do not confirm or disconfirm a theory by testing its predictions against the world. Instead, a thought experiment tests the explanatory power of a theory. A theory has "explanatory power" to the extent it can explain a lot of phenomena with a small number of basic principles. One particularly elegant way to characterize this is in terms of *unification* (Friedman 1974, Kitcher 1981, 1989). Suppose that a scientific theory consists of a set of specific explanatory (or problem-solving) strategies; the unification of a theory is inversely proportional to how many specific explanatory strategies it employs and directly proportional to the number of phenomena it is capable of explaining (Kitcher 1989). Assume for the sake of argument that other things being equal, it is rational to accept the most unified scientific theory. Given this assumption, we can see how thought experiments could play a role in rational scientific theory choice. Thought experiments test the extent to which a theory is unified, and if we rationally choose theories, in part at least, because they are unified, then thought experiments can play a role in rational theory choice.

To show how well these views of thought experiments work, let's consider the Einstein-Podolsky-Rosen thought experiment (1935). (For a fascinating discussion of the history of this experiment, see Jammer 1974, chapter 6.) The experimental set-up is straightforward: Suppose there are two quantum systems that interact so that their properties are correlated by conservation laws. For convenience, consider a case in which a pair of electrons, *e1* and *e2*, rebound from a collision in opposite directions and at the same speed. The tricky part of the EPR thought experiment is the derivation of the following result: electron *e2* has a precise position and momentum at a specific time. EPR derived this result as follows.

> After the electrons have travelled some arbitrary distance, it is possible to measure the precise position of *e1* and thus infer, using conservation laws, the precise position of *e2*. Alternatively, it is possible to measure the precise momentum of *e1* and thus infer, using conservation laws, the precise momentum of *e2*. The uncertainty principle says that it is not possible to measure precisely both the position and momentum of *e1*, and hence infer both the precise position and momentum of *e2*. But no matter. EPR assumed the following "reasonable" principle: "If, without in any way disturbing a system, we can predict with certainty (i.e., with probability equal to unity) the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity" (p. 777). If the measurement of *e1* occurs far enough away from *e2*, and in principle it could occur on the

other side of the universe, EPR assumed it would not disturb *e2*. Since we can predict with certainty *e2*'s precise position at time t without disturbing it, then *e2* must have a precise position at *t*. And since we can predict with certainty *e2*'s precise momentum at *t* without disturbing it, then *e2* must have a precise momentum at *t*.

So far, this example has the features sketched in section 1: an imagined experimental set-up and a process of reasoning that leads to a result.

EPR proposed a second argument, the nature of which is the subject of this section. On the basis of their result, EPR concluded that quantum mechanics is not an adequate theory. "While we have thus shown that the wave function does not provide a complete description of the physical reality, we left open the question of whether or not such a description exists. We believe, however, that such a theory is possible" (1935, p. 780). My contention is that given the thought experimental result (which Bohr did not accept), it is rational to draw conclusions about the epistemic status of quantum mechanics if unification is a theoretical virtue. The reasoning is straightforward. Given the EPR result, there are facts which are part of the explanatory base of the theory (the EPR result) but which cannot be explained by the theory. And this is a failure of quantum mechanics because the theory is not as unified as it could be, *and unification is a theoretical virtue*.

One might think that any reasonable theoretical virtue would leave a role for thought experiments in rational theory choice. This is not true. Consider a group of scientists who are radical instrumentalists, for whom accurate prediction is the primary theoretical virtue. They believe that the purpose of their theories is to make accurate predictions and postdictions in a computationally simple fashion, though whether this generates greater theoretical unification is immaterial to them. We would not expect them to develop or even take seriously unactualized thought experiments that are meant to undermine confidence in their theories. This philosophical thesis has falsifiable implications: instrumentalists (Ptolematic astronomers, for example) will not be moved by thought experiments.

4. Thought experiments: truth and idealization

Thought experiments are idealizations. They abstract away from the messy complexities and practical impossibilities inherent in the real world. What does the prevalence of thought experiments in contemporary physics tell us about the nature of contemporary physics, as well as other branches of science that traffic in thought experiments? Philosophers have extracted varying lessons. For Feyerabend (1978), thought experiments show that science is much more interesting than most philosophers make it out to be, oft consisting of propaganda and psychological tricks. For J.R. Brown (1991), thought experiments show that Plato was broadly correct about the nature of empirical knowledge. I want to suggest that the prevalence of thought experiments in scientific theory choice typically betrays an implicit commitment to a kind of *realism* about scientific theories. Their prevalence forces us to reject van Fraassen's view that the aim of science is the modest one of accounting for observables rather than the more ambitious one of accounting for both observables and unobservables (1980, chapter 1).

It is ironic that scientists' commitment to truth could be betrayed by their use of untrue, imaginary, unactualized (or even unactualizable) experiments in testing their theories. My argument for this contention begins with the premise that the aims of science are implicit in the theoretical practices of successful scientists. If they are not to be found in the behavior of those who do successful science, then where ever could they be? Unless we naively assume that a scientist has some special access to the standards implicit in her practices, this suggests that when it comes to figuring out what the aims of science are, we should pay attention to what scientists do, not what they say.

(1) The epistemological aims of science are implicit in the theoretical practices of successful scientists.
(2) Many of the most successful 20th century physicists have used thought experiments to test scientific theories.
(3) Many of those thought experiments require that a theory be able to account for unobservable phenomena.

---

(4) Therefore, one aim of science is to account for unobservable phenomena.

Premise 3 requires some justification. According to van Fraassen, "*X* is observable if there are circumstances which are such that, if *X* is present to us under those circumstances, then we observe it" (1983, p. 16). One might argue that by definition, unactualized thought experiments do not have observable results. If they did, they would be actualized experiments. But this objection misunderstands what van Fraassen means by observable. According to van Fraassen, non-existing things are observ*able* (*Ibid.*, p. 15). Under the right sorts of conditions, including of course the presence of a unicorn, we *could* observe a unicorn. So some possible or imaginary findings are observ*able*. The real problem for van Fraassen is that the results of both the clock-in-the-box and EPR thought experiments are not observable in his sense.

Recall that in the clock-in-the-box episode, Einstein and Bohr disagreed about whether it was possible to measure a photon's weight (and hence energy) at a particular time. No normal human observer put in front of the clock-in-the-box apparatus (figure 1) could observe any change in the apparatus after the photon exited the box. The thought experimental findings of both Bohr and Einstein are, in van Fraassen's sense, unobservable. And yet both Bohr and Einstein behaved as if there were just no question about whether quantum mechanics should account truly and literally for this unobservable phenomenon. Einstein, after all, thought that the result showed quantum mechanics is false; and recall Rosenfeld's touching description of the befuddled Bohr, "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." EPR is even more trouble for van Fraassen (which is ironic, since he uses EPR to oppose arguments for scientific realism, 1983, pp. 28-31). Einstein and Bohr agreed completely about what observations could be obtained from the EPR experiment. *What they disagreed about was a non-observable result of the thought experiment*: whether the unmeasured particle has a precise position and momentum. And once again, Bohr and Einstein behaved as if there were just no question about whether quantum mechanics should account for this unobservable phenomenon. "Einstein never abandoned the view that quantum mechanics, as presently formulated, is an incomplete description of physical reality" (Jammer, p. 188). And Bohr does not reject the EPR challenge as outside the scope of what science aims to do. Instead, he launches head-first into a discussion of what quantum mechanics implies about unobservable reality: it entails "the necessity of a final renunciation of the classical ideal of causality and a radical revision of our attitude towards the problem of physical reality" (Bohr, p. 233). The main thrust of Bohr's proposed revision is that we are "dealing with phenomena where no sharp distinction can be made between the behavior of the objects themselves and their interaction with the measuring instruments" (*Ibid.*, p. 234).

Once again, the behavior of Bohr and Einstein implicitly presupposed that one of the aims of a theory (quantum mechanics) is to account for something unobservable (the EPR result). What is more, their behavior implicitly presupposed that one aim of a successful scientific theory is to explain unobservable phenomena *even if that phenomena might have no observational effect*. Of course, it turned out that the tests for Bell's inequalities showed that the Bohr-Einstein disagreement did have observational implications. But they didn't know that. If the epistemological aims of science are to be found in the epistemological aims implicit in the behavior of successful scientists doing science, then it would appear that at least sometimes, the aim of science is to account for phenomena that are not only unobservable but that potentially have no observable effects!

I have argued that the use of thought experiments by figures like Bohr and Einstein reveals something about the aims of contemporary physics. One of its aims is to account for the truth about unobservables. Whether this aim can be realized by appeal to imaginary goings-on is another issue entirely.

Appendix

Bohr proposes a method for measuring the box's weight change after the photon's exit: add a weight to the bottom of the box in order to bring the pointer back to zero. It is possible to bring the pointer's position back to zero to within any desired degree of accuracy, Δq. But given the uncertainty principle, this quantity, Δq, will imply a minimum latitude, Δp, in the uncertainty of the box's momentum,

$$\Delta p \approx h \,/\, \Delta q$$

where h is a constant (Plank's constant divided by 4π). Now, the total impulse (change in momentum $\Delta(v\ '\ m)$) of the box during the weighing procedure is this,

$$I = T \times g \times \Delta m$$

where T is the duration of the weighing procedure and g is the gravitational constant. Clearly, Δp (the uncertainty in the box's momentum) will be smaller than the box's total momentum change (impulse).

$$\Delta p \approx h \,/\, \Delta q < T \times g \times \Delta m \tag{1}$$

At this point, Bohr appeals to Einstein's own theory of general relativity. According to the red shift formula, a clock displaced by Δq in the direction of a gravitational force implies a change, Δt, in the clock's reading.

$$\Delta t = T \times g \times \Delta q \,/\, c^2 \tag{2}$$

Now come some very simple manipulations of formulas (1) and (2). We can isolate T in formula (2) on the left side of the equation.

(3) $$T = \Delta t \times c^2 / g \times \Delta q$$

Replace T in formula (1) with the right side of (3).

(4) $$h / \Delta q < \Delta t \times c^2 \times g \times \Delta m / g \times \Delta q$$

After canceling, we get the following.

(5) $$h < \Delta t \times \Delta m \times c^2$$

Given that $E = mc^2$ and hence that $\Delta E = \Delta m \times c^2$, Bohr derives Heisenberg's uncertainty principle:

(6) $$h < \Delta t \times \Delta E$$

Michael A. Bishop
Department of Philosophy
Iowa State University
USA

REFERENCES

Bishop, M. (1992). The Possibility of Conceptual Clarity in Philosophy. *American Philosophical Quarterly* 29, 267-78.

Bohr, N. (1935). Can Quantum-Mechanical Description of Physical Reality be Considered Complete? *Physical Review* 48, 696-702.

Bohr, N. (1949) Discussion with Einstein on Epistemological Problems in Atomic Physics. In: P. Schilpp (Ed.), *Albert Einstein: Philosopher-Scientist*. Evanston: The Library of Living Philosophers, pp. 199-242.

Brown, J.R. (1991). *The Laboratory of the Mind*. New York/London: Routledge.

Einstein, A., B. Podolsky, N. Rosen (1935). Can Quantum-Mechanical Description of Physical Reality be Considered Complete? *Physical Review* 47, 777-80.

Feyerabend, P. (1978). *Against Method*. London: Verso.

Forge, J. (1991). Thought Experiments in the Philosophy of Physical Science. In: T. Horowitz and G. Massey (Eds.), *Thought Experiments in Science and Philosophy*. Maryland: Rowman & Littlefield, pp. 209-22.

Friedman, M. (1974). Explanation and Scientific Understanding. *Journal of Philosophy* 71, 5-19.

Gooding, D.C. (1993). What is *Experimental* about Thought Experiments? In: D. Hull and M. Forbes (Eds.), *PSA 1992, Volume Two*. East Lansing: Philosophy of Science Association, pp. 280-90.

Irvine, A.D. (1991). On the Nature of Thought Experiments in Scientific Reasoning. In: T. Horowitz and G. Massey (Eds.), *Thought Experiments in Science and Philosophy*. Maryland: Rowman & Littlefield, pp. 149-66.

Jammer, M. (1974). *The Philosophy of Quantum Mechanics*. New York: John Wiley & Sons.

Kitcher, P. (1981). Explanatory Unification. *Philosophy of Science* 48, 507-31.

Kitcher, P. (1989). Explanatory Unification and the Causal Structure of the World. In: P. Kitcher, W. Salmon (Eds.), *Scientific Explanation*. Minneapolis: University of Minnesota, pp. 410-505.

Lennox, J. (1991). Darwinian Thought Experiments: A Function for Just-So Stories. In: T. Horowitz and G. Massey (Eds.), *Thought Experiments in Science and Philosophy*. Maryland: Rowman & Littlefield, pp. 223-46.

Nersessian, N. (1993). Thought Experimenting as Mental Modelling. In: D. Hull and M. Forbes (Eds.), *PSA 1992, Volume Two*. East Lansing: Philosophy of Science Association, pp. 291-301.

Norton, J. (1991). Thought Experiments in Einstein's Work. In: T. Horowitz and G. Massey (Eds.), *Thought Experiments in Science and Philosophy*. Maryland: Rowman & Littlefield, pp. 129-48.

Pais, A. (1982). *"Subtle is the Lord..."* Oxford: Clarendon Press.

Rescher, N. (1991). Thought Experimentation in Presocratic Philosophy. In: T. Horowitz and G. Massey (Eds.), *Thought Experiments in Science and Philosophy*. Maryland: Rowman & Littlefield, pp. 31-42.

Swenson, L.S. (1972). *The Ethereal Aether: A History of the Michelson-Morley-Miller Aether-Drift Experiments, 1880-1930*. Austin/London: University of Texas Press.

van Fraassen, B.C. (1980). *The Scientific Image*. Oxford: Clarendon Press.