Do Thought Experiments Have a Life of Their Own? Comments on James Brown, Nancy Nersessian and David Gooding
Author(s): Ian Hacking
Reviewed work(s):
Source: *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, Vol. 1992, Volume Two: Symposia and Invited Papers (1992), pp. 302-308
Published by: The University of Chicago Press on behalf of the Philosophy of Science Association
Stable URL: http://www.jstor.org/stable/192844
Accessed: 02/08/2012 21:22

Do Thought Experiments have a Life of Their Own? Comments on James Brown, Nancy Nersessian and David Gooding

Ian Hacking

University of Toronto

I believe that thought experiments do not have a life of their own. The tide of opinion is going the other way. That is evident from the ample citations provided by our three speakers. They include for example the outstanding collection of papers, edited by Tamara Horowitz and Gerald Massey (1991)with its wonderful editorial introduction that really does say what is in each paper. There are David Gooding's chapters (1990, chs. 8-10) on the relationships between Faraday's thought experiments and real experiments. There is James Brown's book (1991). It is sad that John Norton had to return on family business to Australia, and was unable to chair our meeting, because his views are in such absolute opposition to Brown. The "empiricism" of Brown's title today is none other than that of John Norton (1991). There is also the forthcoming book by Roy Sorensen, for which he has written an excellent popular prolegomenon (1991). In addition there are numerous forthcoming papers that I have not seen, and which are cited by our speakers. Finally this session itself confirms that the discussion of thought experiments is going through a substantial boom.

This enthusiasm for thought experiments is all the more remarkable because our forbears have not written a great deal on the topic. As Nersessian reminds us, our philosophical predecessors—as opposed to scientists who used thought experiments in their argumentation or reflection—are Mach and Koyré. We might add a mention of Karl Popper, whose comments in *Logic of Scientific Discovery* (1959, appendix XI) were singled out for praise by Koyré. And of course we have Kuhn's famous paper (1977 [1964]). He tells us that he almost called it "The" Function for Thought experiments—not just "A" Function as the final version of the title has it.

We have heard three very distinct approaches today. People have been struck by Brown's platonism. He uses a lower case "p", and his platonism is fallibilist. For me the important message in his paper may be that there are different kinds of thought experiments, a case argued at greater length in chapter 2 of his book. Sometimes Brown and I differ in point of nomenclature. I resist calling Maxwell's demon part of a thought experiment. It is part of a fantasy. Here I agree with Nersessian, who has remarked to me that it is hard to see what is experimental about the demon. Perhaps it is only a rhetorical device to reinforce Maxwell's statistical analysis. The demon does not, for me, prove even the possibility of anything. Perhaps it shows an aspect of a scene in statistical mechanics.

PSA 1992, Volume 2, pp. 302-308

Brown tells me that people readily call the demon a thought experiment. That certainly helps his case. He claims that Norton's analysis is inadequate, and I certainly don't see the demon fitting into Norton's framework. But is it a thought experiment? Nersessian says that thought experiments characteristically are so compelling that there is no need to conduct the experiment. The problem with Maxwell's demon as an experiment is that you can't conduct it at all, no matter how much fantasy and idealization you allow yourself. Nor can you conduct another fantasy noticed by Gooding—that of Faraday considering whether the electrical charges felt by mites on a sphere are positively or negatively charged. (Were the mites ancestors of Maxwell's demon?) Perhaps Gooding would say that Maxwell's demon is not part of a thought *experiment* because you cannot put yourself in the place of that demon, or of Faraday's mites either. I do not deny, of course, that there is something of a continuum here and understand Brown's wish to count the demon as part of the litany of thought experiments.

At any rate Brown's position must be seen in part as a contrast with that of the absent Norton, who maintains that every thought experiment is sound insofar as it can be cast into a valid deductive form (Norton 1991). Brown grants this of some, but not of all thought experiments. In particular he wants to insist on thought experiments that resemble intuitive insight in mathematics, rather than the combinatorial checking of proof. In my opinion, he grants too much to Norton by ceding to him the word "argument". There are many kinds of argument than are regimented by symbolic logic. Norton claims that all thought experiments can be cast into the form of deductions. The ideal is that each step is an instance of a rule of inference. That is how we have come to define the word "argument" in elementary logic. An argument is a sequence of sentences; it is valid if every member of the sequence is a premise or follows from preceding sentences by one application of a rule of inference. It is sound if the premises are all true. That is a powerful logistical idea. It captures one of the strands in mathematical thinking. That strand is, to put it too crudely, algorithmic and Arabic. Brown's favoured examples are geometrical and Greek. There is no reason for him to grant that all arguments using imaginary experiments, if they deserve to be compelling, must be capable of being turned into valid proofs. At most he needs to grant that if they cannot be turned into valid proofs, they have shown themselves to be fallible. And Brown is an avowedly fallibilist platonist.

Nancy Nersessian's and David Gooding's approaches at first seem as different from each other as from Brown's, yet all three are compatible—and anti-Nortonian. All look at thought experiments as something other than concealed derivations. Nersessian's approach is from the mental modelling wing of cognitive science. "Simulated model based reasoning" is her catchphrase. David Gooding's language only looks different from hers. His keyword is "embodiment." We could see Nersessian and Gooding combined after a reading of Mark Johnson's recent book *The Body in the Mind* (1987). Nersessian has confirmed in correspondence that she takes embodiment to be a central feature of her analysis. I notice that both speakers have today favoured the word "narrative", and Brown has used it too. All three have an essentially non-propositional attitude to thought experiments. Stories come in sentences, but for these philosophers, we tell a story so that we can put ourselves into an imagined situation. All three might reply to Norton that to turn a thought experiment into a derivation is to transform it from an experiment into something else. On this account, Norton does not have a theory of thought experiments at all! He has a theory about what can be constructed out of thought experiments. That in turn may provide a criterion of correctness for thought experiments, but is not the experiment itself.

Thus all three speakers follow routes importantly different from Norton's. It may be true that there is a formal derivation of every conclusion reached in a sound thought ex-

periment. Certainly Norton's own work is a marvellous way of unpacking the deductions in Einstein's stories. And we have to regret once again that Norton is not here to reply against this common front that I have just read into the three papers we have just heard.

Let us go back to first, or at any rate early, principles. What is the problem about thought experiments? Mach, who seems to have invented the term *Gedankenexperiment*, seems to have seen no problem at all. Koyré, who took over the idea under the name of *expérience imaginaire*, was equally untroubled. Koyré thought that imaginary experiments were ideal abstractions, in which he revelled. Mach believed in contrast that there is something of a continuum between real experiments and thought experiments. David Gooding is a Machian. Indeed he shows us a continuum in Faraday's work, a chain of experiments ending with an impossible experiment, a thought experiment. But we still can see a problem about thought experiments. Kuhn stated it with characteristic éclat.

> If we have to do with a real thought experiment, the empirical data upon which it rests must have been both well-known and generally accepted before the experiment was even conceived. How then, relying exclusively upon familiar data, can a thought experiment lead to new knowledge or to new understanding of nature? (1977 [1964], p. 241).

This statement has become almost classical, quoted for instance at the beginning of Mark Wilson's paper on thought experiments (1991). And yet there is something not quite right about Kuhn's statement. How on earth, he makes us wonder, could we find out something about the world "just by thinking"? But pause a moment. We do it all the time! What, after all, is applied mathematics, but the deduction of what we don't know from what we know? There is of course an ancient tradition that says deduction merely makes us aware of what was contained in our premises. Very well, but it still leads us to new awareness, new understanding of the world. I shall return to Kuhn because I do think he has left a permanent mark on our topic. Thought experiments are valuable when they bring out, in a succinct way, a conceptual tension between two ways of thinking, and force us to come to grips with it. Thought experiments are not simply a matter of probing conceptual confusion. Rather it is a tension between two ways of describing the world, both internally consistent, but one in some way simpler than the other. The simpler way is forced to yield to the more rigorous. (Wilson has an example of the more rigorous yielding to the simple).

Without in any way calling in question that aspect of Kuhn's insight, I would like to add a rather pedestrian way of understanding the quotation from Kuhn. It drives us along Brown's path, and leads us to compare thought experiments to intuitive mathematical reasoning. What is the philosophical problem involved? There is a famous problem whose most eloquent statement has been passed down to us for almost 2500 years. That is the problem about the slave boy in *Meno*, asked to construct a square double the size of a given square. Is that mathematics? It might be better to call it rational geometry, by analogy with rational mechanics. Plato was developing the idea of mathematical knowledge as pertaining to a higher plane than knowledge of material things, but his picture was not our simplistic one of an analytic-synthetic distinction, used in the twentieth century by logical positivists and others to combat platonism. Let us take Plato in his own terms, without forcing on him our distinctions between the mathematical and the empirical.

When we do that, it is far more natural to call Plato's story a thought experiment than to call Maxwell's story a thought experiment. The fact that we call it mathematics does not put me off. The argument presented in *Meno* conforms to Nersessian's list of five characteristic features of thought experiments:

1. It is a narrative—and what a narrative!

2. We follow it through as in the real world.

3. The experiment works. Both Gooding and Nersessian note that we tend not to be shown the tinkering that led up to a polished thought experiment. Plato shows us the tinkering, and the initial failures.

4. The thought experiment works by abstractions. At one juncture Socrates is made to say that the figures drawn on sand could be larger or smaller.

5. The thought experiment is so compelling that there is no need actually to perform the experiment. Indeed someone who started measuring, after seeing the boy's final proof, would not have understood the argument (or would be in the wrong millennium, wondering about non-Euclidean geometry).

The example given in *Meno* is part of a philosophical argument. It led Socrates to the most astonishing philosophical theory. The boy remembers from a pre-existing soul, and is helped by the process of dialectic. The theory of anamnesis is a powerful reminder of how deeply a profound thinker can be moved by the experience of proof—the very experience to which Brown's platonism constantly directs us.

But if that is the problem ("the hardness of the logical *must*") then it is not a problem peculiar to thought experiments. It is a problem about mathematical knowledge. It is hardly surprising that Brown gives us a picture for grasping an arithmetical identity, and calls it a thought experiment. Note that his arithmetical picture-example does not obviously satisfy Nersessian's fifth feature. I don't well understand what performing the experiment would be, for the general proposition. I can perhaps see that it works for any numeral but I don't see that it works for every numeral. But I have no doubt about the cogency of Brown's general aim. It is always an essential task, to remind us how strange mathematics is.

I nevertheless find a problem with Brown that I find with the other two speakers as well. The *explanations* of how thought experiments work, that are offered by our three speakers, do extraordinarily little for me. We have been inoculated against Brown's variety of platonism for too long. Most of us see at once that his idea, of immediate acquaintance with universals, abstract objects, just does not do any explanatory work for us. But oddly the same is true for the ideas presented by our other speakers too.

Before urging that claim, I should acknowledge that Gooding's talk of embodiment is deeply important. If we look at the picture of Stevinus's inclined plane (drawn by Brown), and think through the experiment, we want not just one figure, as in Brown's talk. We need at least a before and after picture of the plane and necklace, and we need someone snipping the necklace—simultaneously—at both ends of the block of wood. We see ourselves getting in there and cutting it. In real life, we would have a terrible time making the two cuts simultaneously—I might in fact go for the lowest point of the necklace and make just one cut. But at any rate there is embodied action. Just as pictures of real life experiments usually omit the experimenters and what they literally do, so we too often forget that it is the bodily feel of the thought experiment that convinces us. Norton is right to produce formal derivations arising from Einstein's thought experiments. But Gooding is right to emphasize the man in the elevator. On his view, I believe, it is that man who makes the story into a thought experiment rather than a proof; even the man's being drugged, as in one version of the experiment, is not irrelevant.

Having spoken up for Gooding's idea of embodiment, I must repeat that I find that all three speakers do little for me when I look at a thought experiment. I will take as an example a particularly easy one. In March 1686 Leibniz published a little anti-Cartesian broadside in the *Acta Eruditorum*. It had a trenchant title: "A Brief Demonstration of a Notable Error of Descartes and Others Concerning a Natural Law." It was concerned with perplexing concepts like momentum, kinetic energy and work. This was a Kuhnian scene *par excellence*, in which conceptual clarification was at issue. There is a certain resemblance of Leibniz's argument to the famous objects in free fall that Galileo tied together. But the reasoning is part of a process of sophistication leading up to Leibniz's *Essaie de dynamique* (1693) and *Specimen Dynamicum* (1695). But rather than considering Leibniz's first experiment, let us consider a supplement to it, dating from a later period. that is connected with the idea of work. The project was to establish something we almost take for granted. "It is to be shown that the power (*potentiae*) required to lift 1 pound 2 feet is the same as the power required to lift 2 pounds 1 foot." (Leibniz 1969, p. 298).

"I postulate," wrote Leibniz, "that I am permitted to assume various connections of the heavy bodies with each either, and their separation again, and to introduce any other changes which do not make a change of force. I also make us of threads, axes, levers and other mechanisms lacking in weight and resistance." Some abstraction.

Leibniz then proved that a body of one pound, falling through two feet, will have exactly the "power" (*potentiae* which the translator here renders "force", although a page earlier as "power") required to lift a two pound body through one foot. And he does this by dividing the two-pounder into two, and then drawing up each half successively by letting the one-pound weight fall one foot in succession, for a total of two feet.

That, I suggest, is a paragon of a thought experiment. As I contemplate it, none of our speakers does anything for me. Yes, there is narrative. Yes, there is embodiment. Simulated model based reasoning? It seems easier to think about the experiment, perhaps draw a figure, than to take me down to the level of mental models. And as for Brown's universals with which we are acquainted, once again that does no work for me. This is a very unphilosophical type of objection to my colleagues, but it must be one that the lay person feels quite strongly. It certainly is at the core of the general resistance to Brown's platonism. And if cognitive science were not so fashionable, I think many of us would feel the same resistance to Nersessian's conceptual models. They are presented as if they were explanatory, but in fact explain nothing.

I am in fact more moved by reading Wittgenstein, for example one remark in the last part of his remarks on mathematics. He draws a picture of a mechanism, a sort of piston in which a wheel is driven round by a rod fixed to it at point A, and passing through a slide at point B.

> While the point A describes a circle, B describes a figure eight. Now we write this down as a proposition of kinematics.
>
> When I work the mechanism its movement proves the proposition to me; as would a construction on paper. The proposition corresponds e.g. to a picture of the mechanism with the paths of the points A and B drawn in. Thus it is in a certain respect a picture of that movement. It holds fast what the *proof* shews me. Or—what it persuades me of. (Wittgenstein 1978, p. 434.)

Wittgenstein's word "picture" does more for me than the fashionable word "narrative". It is essential to recall his much earlier thought:

> The experimental character disappears when one looks at the process as a memorable picture (p. 68).

I do not deny that a mental modeller may in some distant day provide an account of how Wittgenstein's pictures work. In the meanwhile, you will find many aphoristic insights about pictures and experiments in the *Remarks on the Foundations of Mathematics*. I commend consulting the index to the third edition (1978).

Wittgenstein is among other things a great comic author, in a certain Austrian tradition that we dowdy philosophers of science tend to ignore. And when we look at thought experiments, especially the ones that are shown us by Brown, we may experience a touch of iconoclasm. What do these pictures remind me of? They remind me of jokes, that is, very trenchant well-worked out items with a punch line. They also remind me of optical illusions. That is, they are experiences that it is hard or impossible to exorcise. They are snappy and, like optical illusions or probability paradoxes, they will get you every time, no matter how well prepared you are.

Over a decade ago I wrote that experiments have a life of their own. I intended partly to convey the fact that experiments are organic, develop, change, and yet retain a certain long-term development which makes us talk about repeating and replicating experiments—see Hans Radder's paper at this conference (See PSA 1992, Vol. 1). I think of experiments as having a life: maturing, evolving, adapting, being not only recycled but also, quite literally, being retooled. But thought experiments are rather fixed, largely immutable. That is yet another respect that they are like mathematical proofs, but good proofs have proof ideas that can be used over and over in new contexts—which is not, in general the case with thought experiments. They have just one tension to expose. Of course there are false starts, and the exposition gets neater over time. And here the prescience of Kuhn's paper comes to the fore. The reason that people wrestle with thought experiments, use them for exposition and put-down argument, is that they can reveal tensions between one vision of the world and another. They can dislodge a person from a certain way of describing the world. They can replace one picture by another. That is their job, their once and future job.

Kuhn, in the discussion following our presentations, urged against me that thought experiments do have a life of their own. But he was making a point that is consistent with mine. He was urging that a thought experiment is constantly repeated, and not interfered with. People do rethink through the thought experiment, even from generation to generation. But what they think is what was once thought, and continues to use the very same diagrams and dialogue—just as Plato's account of Socrates and the slave boy leads to us to draw the same diagrams, from generation to generation, once upon a time in the sand, now on paper or a transparency. But is this life? Is it not like acting a part in a play? Olivier brought new life to Othello, but no matter how profound the power of Othello or Lear, those characters do not have a life of their own, Tom Stoppard notwithstanding. Once the thought experiment is written out in perfection it is an icon. Icons, to reiterate, do not have a life of their own.

References

Brown, J. R. (1991), *The Laboratory of the Mind: Thought Experiments in Natural Science*. London and New York: Routledge.

Gooding, D. (1990), *Experiment and the Making of Meaning*. Dordrecht: Kluwer.

Horowitz, T. and Massey G.J. (1991), *Thought Experiments in Science and Philosophy.* Savage, Md.: Rowman & Littlefield.

Johnson, M. (1987), *The Body in the Mind: the Bodily Basis of Meaning, Imagination, and Reason.* Chicago: University of Chicago Press.

Kuhn, T.S. (1977 [1964]), "A Function for Thought Experiments", reprinted in *The Essential Tension: Selected Studies in Scientific Tradition and Change.* Chicago: Chicago University Press, 1977, 240-265.

Leibniz, G.W. (1969), in *Leibniz: Philosophical Papers and Letters*, L. E. Loemker (trans. and ed.). Dordrecht: Reidel.

Norton, J. (1991), "Thought Experiments in Einstein's Work" in Horowitz and Massey, 129-148.

Popper, K.R. (1959), *The Logic of Scientific Discovery.* London: Hutchinson.

Sorensen, R. (1991), "Thought Experiments", *American Scientist* **79**, 149-63.

Wilson, M. (1991), "Reflections on Strings", in Horowitz and Massey, 193-207.

Wittgenstein, L. (1978), *Remarks on the Foundations of Mathematics.* 3rd edition, Oxford: Blackwell.