# THOUGHT EXPERIMENTS IN SCIENCE EDUCATION

Dr Maurizio Toscano

The University of Melbourne, Australia

*Mere contemplation of imaginary scenarios can have profound effects on our beliefs about contingent features of the natural world. Yet the question of how such knowledge is derived from experiments that are conducted in thought alone remains unanswered. The empiricist argues that thought experiments are more indebted to empirical evidence than their name would suggest, the Platonist argues that thought experiments allow access to a priori knowledge while the logician contends that thought experiments are merely picturesque arguments. The lack of theoretical consensus on the nature of thought experiments is a major hurdle to researching their educational utility. This paper presents existing accounts of thought experiments while putting forward an alternative account based on scientific modelling and metaphor, in the hope of illuminating the place of thought experiments in science education.*

## A MODEL-BASED ACCOUNT OF THOUGHT EXPERIMENTS

In this paper, I claim that thought experiments are models. This identification makes possible a new way of looking at students' and educators' understanding of the nature of scientific thinking. For the philosopher of science this paper strengthens the arguments in favour of a constructivist, non-aprioristic and non-argument-based interpretation of thought experiments. For the educator it demonstrates the usefulness of introducing thought experiments into science lessons, not as specimens of esoteric scientific reasoning, but rather as exemplars of how scientists construct their theories.

Model-based accounts of thought experimentation have been suggested most recently by Cooper (2005), Miščević (1992), and Nersessian (1992). Although my interpretation certainly resonates with their work, it differs significantly in approach and intention. For my aim is not just to offer a possible mechanism by which to explain how thought experiments might work, it also seeks to reveal those aspects of thought experimentation that could be successfully employed in science education, or more precisely in teaching the nature of science to students. Consequently, instead of looking at modelling from the perspective of cognitive science or cognitive psychology as Miščević and Nersessian do, I deliberately place modelling firmly in the context of scientific theory construction. I argue that this is an important and hitherto neglected stage towards incorporating the implications of the lively philosophical debate about thought experiments into mainstream science education

Much of the momentum enjoyed by the thought experiment revival in the early 1990s was generated by the ongoing polemic between James Brown and John Norton (Brown 2004, Norton 2004a). While John Norton put forward the idea that thought experiments were arguments in disguise and nothing more (Norton 1991, 2004b), Brown suggested that thought experiments worked by steering the experimenter towards intuiting some platonic realm in which the laws of nature were manifestly self-evident (Brown 1991). As philosophers posited their own theories in opposition to the most striking limitations of these interpretations, less extreme positions on thought experiments began to emerge; ones that awarded greater significance to the agency of the thought experimenter in constructing their own meaning of the thought experiment. As an example, Gendler writes: '…thought experiments rely on a certain sort of constructive participation on the part of the reader, and that the justificatory force of the thought experiment actually comes from the fact that it calls upon the reader to perform what I will call an experiment-in-thought' (Gendler 1998: p 413). Gendler takes experiments-in-thought to mean 'actual experiments' performed by actual persons in real time.

In light of Gendler's insight, (see also Gendler 2004) it would be difficult to articulate thought experimentation simply in terms of a passive reception of the experimental claims that the originator of the thought experiment intended to convey. As in the interpretation of any text, a thought experiment requires action on the part of the though experimenter. Put another way, we are called upon to avoid the often-made tacit assumption that there is a true experimenter-independent thought experiment out there; presumably some idealization of the experiment that was developed and publicized by the original author in the interests of supporting her scientific or philosophical claims. I believe that such an assumption operates in Norton and Brown's respective interpretations of thought experiments. Taking their treatment of Galileo's tethered spheres thought experiment as an example (Brown 2004, Norton 2004a), Norton's reconstruction seems to suggest that Galileo's claim that all bodies fall at the same rate irrespective of their masses is just sitting there ready to be discovered via a formal chain of argumentation, whilst Brown's platonic account has the presupposition of a experimenter-independent reality built into it from the beginning. I suggest removing this assumption and re-conceiving thought experiments as examples of dynamic model making.

Just as it is unwise to undervalue the constructive participation of the thought experimenter or to assume that the content of thought experiments is somehow independent of the experimenter, I argue that we must also avoid underestimating the social, cultural and historical forces that shape a thought experiment long before the instant when we are expected to conduct (or construct) the experiment using our own imaginations. I have already discussed the third and final stage of thought experimentation; namely the active construction of meaning on the part of the experimenter. The other two stages involve (1) the initial phase of thought experiment construction, and (2) the stage in which refined thought experiments are used to defend established theories. Stage (1) begins with a real person making a scientific claim. In the absence of suitable sense-extending apparatus, she may be forced to support her claims with results drawn from experiments-in-thought instead of empirical data gathered from her own sensory experience, or from an instrument-aided experience of the world. I believe the dearth of such direct empirical data may be nothing more than the absence of an appropriate language with which to describe and explain the underlying mechanisms of new phenomena. Hence, thought experimentation is about constructing a new language out of ordinary language for rhetorical purposes. This is equivalent to Sutton's figurative stage in the development of scientific language (Sutton 1996). In stage (2), thought experiments are presented to an audience with the aim of buttressing the scientific status quo. This stage differs from the first in that its intention is to present things as they are rather than as they might be. That is, thought experiments are utilized to give a kind of folkloristic account of well-established scientific theories in order to 'normalize' them. There is no question that the narrative plot line sanctioned by the scientific community is fixed. In Sutton's terminology, the scientific language of thought experiments has become literal. Here we have the first hint that thought experiments, at least in the early incarnations, apply analogic reasoning rather than logico-deductive or propositional logic. Furthermore, we can see how thought experiments parallel the development of scientific language.

It should be said from the outset that long tradition of the philosophy of science has to date not produced a consensus view on the nature of science: its ontology, its claims to knowledge (both theoretical and empirical), and of course its practices and methodologies. Nevertheless, few would dispute the special place of modelling in the natural sciences. Consequently, I have taken as my starting point the realist account of science developed by Rom Harré, since he sees modelling as one of the primary tools by which scientists describe and explain phenomena in the world (Harré 2002, 2004). According to Harré, science explores and expands the human *umwelt*; the environment to which humans have access. This human *umwelt* consists firstly of that region of the world that is readily accessible through naked sense perception; secondly, that region of the world that is, or is potentially, accessible via the use of sense-extending instrumentation; and finally, that region that will always remain inaccessible to the aided or unaided 'eye'. It is this final region – the imaginary world – that is explored and expanded using thought experiments.

Of course the scientific method of exploring the human *umwelt* relies firstly on the processes of describing and classifying. This is achieved with a taxonomy by which knowledge is arranged according to class, type and kind. In particular, a type-hierarchical taxonomy 'stores knowledge vertically, in the inheritance relation. To discover what is presupposed about a lower type one runs up the hierarchy through the nodes to the apex. Thus the species 'cat' is vertebrate, animal, living thing' (Harré, 2002: p 41). We shall see later how this taxonomical system is essential to understanding how models function, how scientific theories are formulated and how thought experiments operate.

Scientific models are real or imagined representations of an object or process in the world that scientists utilize for the purpose of describing or explaining phenomena in the human *umwelt*. A model is an iconic representation or analogue of its subject: it displays sufficient likeness to its subject to make it a useful device for thinking about unknown mechanisms by which the phenomena associated with the subject in question operate. If a model is constructed for the purpose of describing a complex or remote process in a more accessible way, then the subject of the model (what the model is of) will also provide its source (what the model is modelled on); e.g. a model car in a wind tunnel. However, if a model is constructed for the purpose of hypothesizing about unobservable processes with the aim of explaining observed phenomena, then the source of the model will usually differ from its subject; e.g. the use of hydrodynamics to model electrical phenomena.

Models are useful because they rely on an analogic mode of reasoning. That is, they work by bringing to the foreground the likenesses and differences between the model, its subject and/or its source. More importantly, 'the use of analogy presupposes that model, source and subject are subtypes of the same supertype within a type hierarchy. They are related to one another via the inheritance relation' (Harré, 2002: p 54). It is the inheritance relation that allows us to explain why some models are better than others in much the same way that it explains why we can excise the infinite number of trivial and negative analogies that arise when only comparing entities using structural isomorphism (Aronson, Harré & Way 1995).

The analogies or models that are available to the scientist are dependent on extant type-hierarchies which in turn, must be determined empirically since the properties of any type within the hierarchy can change places between definitions and accidental attributes. Type-hierarchies are, therefore, dynamic representations of the human *umwelt*. In this picture, scientific theories are taken as segments of a type-hierarchy, with the inheritance relation doing all the work in ensuring the internal consistency characteristic of successful theories: this is the consistency that Cooper (2005) identifies in successful thought experiments. Theories, as sub-sets of type-hierarchies, can emerge or change as new empirical data makes it necessary to form new connections between types within the hierarchy.

But new theories, models (and in particular thought experiments) are possible even without new empirical data, since we can re-interpret a type-hierarchy through a metaphoric lens as it were. That is, we take the types and inheritances in an existing 'literal' type-hierarchy and imagine what other inheritance relations are possible between the same types. This is precisely what the new language generating capacity of thought experiments is evidence of. It is evident that Gendler's 'constructive participation' is the action on the part of the experimenter to metaphorically re-connect the types within the type-hierarchies or their personal *umwelten*. It should also be clear that contrary to Brown, no Platonic realm is necessary. That non-empirical entities seem to emerge from thought experiments is a direct consequence of the metaphoric lens through which we see the underlying, non-apriori type-hierarchy. Perhaps Brown has mistaken the existence of inheritance in type-hierarchies for a priori laws of nature. Also as Aronson, Harré and Way have pointed out, it is possible to describe anything (models, theories or thought experiments) in terms of propositions and deductions once the underlying type-hierarchical structure with its inheritance relation is in place. This is something Norton has demonstrated well. Nonetheless, although thought experiments can be described in terms of arguments they cannot be reduced to arguments.

What I have demonstrated is that thought experiments are well described in terms of the structure and function of scientific modelling. The figurative language they employ and their reliance on analogy stems directly from the model-making and theory-making processes they exemplify. Thought experiments are useful to scientists and students of science because they fulfil the need to explore and expand that region of the human *umwelt* only accessible through the human imagination. Whilst such experiments-in-thought do not result in a prior knowledge, they do allow for a personal re-interpretation of empirically derived knowledge that is arranged hierarchically according to type in an inheritance relation. Since the internal organization of these underlying type-hierarchies are subject to change in the light of new empirical data thought experiments are necessarily dynamic and not reducible to fixed arguments. As examples of scientific thinking they are a useful tool for teaching students about the nature of science.

## References

Aronson, J., Harré, R., & Way, E. (1995). *Realism Rescued.* London, UK: Open Court.

Brown, J. (1991). *The Laboratory of the Mind: Thought experiments in the Natural Sciences*. London, UK: Routledge.

Brown, J. (2004). Why Thought Experiments Transcend Empiricism. In C. Hitchcock (Ed.), *Contemporary Debates in Philosophy of Science* (pp. 23-43) Malden, USA: Blackwell.

Cooper, R. (2005). Thought Experiments. *Metaphilosophy*, *36*, 328-347.

Gendler, T. S. (1998). Galileo and the Indispensability of Though Experiments. *The British Journal of the Philosophy of Science*, *49*, 397-424.

Gendler, T. S. (2004). Though Experiments Rethought – and Reperceived. *Philosophy of Science*, *71*, 1152-1163.

Harré, R. (2002). *Cognitive Science: A Philosophical Introduction.* London, UK: Sage.

Harré, R. (2004). *Modeling: Gateway to the Unknown.* Amsterdam, The Netherlands: Elsevier.

Miščević, N. (1992). Mental Models and Thought Experiments. *International Studies in the Philosophy of Science*, *6*, 215-226.

Nersessian, N. (1992). In the Theoretician's Laboratory: Thought Experimentation as Mental Modelling. *Proceedings of the Biennial Meeting of the Philosophy of Science Association*, *6*, 291-301.

Norton, J. (1991). Thought Experiments in Einstein's Work. In T. Horowitz, & G. J. Massey (Eds.), *Thought Experiments in Science and Philosophy* (pp. 129-148). Savage, USA: Rowman and Littlefield.

Norton, J. (2004a). Why Thought Experiments do not Transcend Empiricism. In C. Hitchcock (Ed.), *Contemporary Debates in Philosophy of Science* (pp. 44-66) Malden, USA: Blackwell.

Norton, J. (2004b). On Though Experiments: Is There More to the Argument. *Philosophy of Science*, *71*, 1139-1151.

Sutton, C. (1996). Beliefs about Science and Beliefs about Language. *International Journal of Science Education*, *18*, 1-18.