Title
Dueling Theories: Thought Experiments in Cognitive Science

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 22(22)

Author
Scott, Sam

Publication Date
2000

Peer reviewed

Dueling Theories: Thought Experiments in Cognitive Science

Sam Scott (sscott@ccs.carleton.ca)
Department of Cognitive Science
Carleton University
Ottawa, ON K1S 5B6

Abstract

Brook (1999) identified thought experiments as one of the key elements of philosophy's contribution to the cognitive sciences. In this paper, I tackle the question of how and why thought experiments work, and what exactly it is they do for us when they do work. I propose that thought experiments almost always involve two different theories of the world being compared to show that they do, or more often do not, fit together. Sometimes both theories are clearly articulated in the narrative of the thought experiment, but more often one of the two goes unarticulated - the thought experimenter instead relies on our shared folk theories of the world. The strength of some of the more famous and persuasive thought experiments lies in their ability to show that a given theory runs afoul of these deeply held folk intuitions. I will compare the "Dueling Theories" account of thought experiments to both Brook's "empirical" account and Brown's (1991) platonic account.

Introduction

In Hilary Putnam's famous thought experiment, we are asked to imagine a possible world called "Twin Earth" identical in all respects to our own, but for the chemical composition of what is called "water" on both worlds. On our world, water is H_2O, while on twin earth it is something else, say XYZ. Now we are asked to consider what happens when Adam on our world and his counterpart, "Twadam" on Twin Earth use the word "water". Do they mean the same thing? Apparently not. Even though their mental states are the same, the external referent is different. Hence, according to Putnam, "meanings just ain't in the head". (Putnam, 1975)

In Jackson's (1991) story of Mary the Colorblind Scientist, we are asked to imagine that Mary knows everything there is to know about color. That is to say, she knows everything that can be measured, described, and communicated about what color is and the process by which we perceive it. But she has never actually seen color before. Then one day she sees a color, say red, for the first time and learns something about color that she did not know before, namely what red looks like. The moral we are asked to draw is usually something about the uniqueness and indescribability of phenomenal experience.

In introducing these two famous thought experiments, what I am first interested in is how well they "work". Reactions tend to vary, but if you're like me, you will be immediately convinced by the story of Mary, but quite skeptical (at least at first) about the Twin Earth story. Perhaps your reactions differ, but the interesting question is why we have the reactions we do. What makes a thought experiment "work" or "not work"? The thesis that I want to defend is that thought experimentation is a meta-activity - a duel between conflicting theories in which one appears to be a clear winner, thus challenging anyone who holds both theories. Thought experiments can never tell us something new about the world, because the world doesn't participate in the experiments. The objects of evaluation are theories of the world. On the other hand, thought experiments are a perfect device not only for revealing problems with various theories of the world, but in some cases, for making clear to us what our theories of the world actually are.

It may not be obvious that the two thought experiments described above fit this "Dueling Theories" story, but if we take the word "theory" to encompass both scientific and folk theory, then the story not only works but can be quite revealing. Consider Mary first. In one corner, we have some theory in which everything that is part of the physical world can be fully described in scientific terms. In the other corner, we have a folk intuition, based on our own experience of perception, that the phenomenal experience of color is indescribable and would therefore be unknowable to someone not directly acquainted with it. What makes this thought experiment so stunning to so many people is that they may not have realized that they held the folk theory, or that the folk theory was so difficult to give up, until it was put into direct conflict with the other theory. In the Twin Earth experiment, any theory that imparts semantic properties solely to brain states is supposed to lose to a reference theory of semantics. What makes someone embrace or reject Putnam's conclusion will be their pre-existing acceptance or skepticism about whether such reference theories of semantics can be made to work.

In what follows, I shall attempt to develop this Dueling Theories view in the context of two alternative views. The first, from Brown's *The Laboratory of the Mind* (Brown 1991), states that thought experiments can reveal a priori truths about the world through their investigation of Platonic universals. I hope to show that this idea should be rejected, mostly because it actually has very little to offer in aid of our understanding of the nature of thought experiments. The second view, from Andrew Brook's "Does philosophy offer cognitive science distinctive methods?" published last year in this forum (Brook, 1999), states that thought experiments get part of their value from their empirical content. This empirical content may make thought experiments capable of testing hypotheses against the real world. I believe that with careful exegesis, Brook's view is actually quite close to my own. The main difference is that my conception allows in principle that there may or may not be empirical content to thought experiments, and in fact, a thought experiment need not actually have directly empirical content in order to make a useful contribution to science or philosophy.

Platonic Thought Experiments

Brown's two paradigm cases of a priori reasoning about the world through thought experiments are Galileo's Coupled Falling Bodies and the photon-decay thought experiment by Einstein, Podolsky, and Rosen, which has become known as the "EPR Paradox". For the benefit of the uninitiated, both these experiments will be explained shortly.[1] But first, a short discussion of Brown's brand of Platonism is in order. According to Brown:

> A *platonic thought experiment* is a single thought experiment which destroys an old or existing theory and simultaneously generates a new one; it is *a priori* in that it is not based on new empirical evidence nor is it merely logically derived from old data; and it is an advance in that the resulting theory is better than the predecessor theory. (p.77, Brown's italics)

Brown thinks that this kind of a priori thought experimentation is possible due to the existence of natural laws as abstract platonic universals, similar to the abstract objects that he supposes to exist in mathematics and logic.

Postulating abstract objects in a Platonic heaven is certainly a controversial move, and one that I am not sympathetic to. But rather than attempting to destroy the monolith of Platonism entirely, I will concentrate on trying to demonstrate that the Dueling Theories story does a better job of accounting for Brown's two favorite thought experiments than does his own Platonic account. But first I will digress briefly to point out two aspects of Brown's Platonism that are clearly problematic from a cognitive perspective.

Cognitive Science and Abstract Objects

A central feature of Brown's defense of Platonism is its heavy reliance on the notion of "obviousness". According to Brown, without Platonism it is "an utter mystery why '3 > 2' seems intuitively obvious" (p. 56). Later, he continues, "if there were no abstract objects, then we wouldn't have intuitions concerning them; '2 + 2 = 4' would not seem intuitively obvious" (p. 64). Here we can take a lesson from research on mathematical cognition. What are we to make of the *non*-obviousness of mathematical facts like:

8329273847592 < 78374223847532, and
89652 + 15265 = 104917 ?

These mathematical facts are of the same order of complexity as '3 > 2' and '2 + 2 = 4'. That is, they use the same number of operators and the same number of arguments. Only the magnitudes differ. But the facts expressed are certainly not intuitively obvious to most people. The reason for the difference is easily accounted for in symbol-processing terms. We have memorized the order of the 10 digits, and this makes ordering judgements easy for single digit numbers, and more complicated for larger ones. Similarly, we have memorized all possible single-digit sums. Thus the obviousness of facts about small numbers and the *non*-obviousness of facts about large numbers can be accounted for in cognitive terms, whereas this difference is difficult to account for in Platonic terms. Perhaps some abstract objects are bigger than others? Or maybe some are further away than others? But what do "bigger" and "further away" mean in the abstract realm?

Brown also makes use of an analogy between sense perception of objects in the real world and our intuitions about objects in the abstract world. He defends this analogy by asserting "the perception of abstract laws of nature is certainly no more mysterious than [ordinary sense perception]." He justifies this by stating that,

> at best we understand part [of visual perception] - the *physical* process starting with photons emitted by an object and ending with neural activity in the visual cortex. From there to *belief* about the object seen is still a complete mystery. (p. 87, Brown's italics)

Fair enough, but at least we do have: 1) an account of the links between an object and the visual cortex and 2) a research program capable of making progress in understanding how the visual system processes information from this link and communicates it to the rest of the brain. As far as I know, we have no account of the link between an abstract object and whatever organ of abstract sense perception we use to perceive it. Nor do we have an account of how these abstract perceptions would be processed in the brain.

The moral of this digression? Putting the logical neatness of Platonism aside and ignoring the natural fascination and attraction that many people feel towards abstract objects, there is much here to be suspicious of from a cognitive perspective.

Coupled Falling Bodies and the EPR Paradox

Back to thought experiments. Galileo's Coupled Falling Bodies is one of Brown's favorite examples of a platonic thought experiment, and is at the very least an impressive piece of reasoning. Aristotle held the view that heavier bodies must fall faster than lighter bodies. This has now been refuted experimentally - we know that on earth, when air friction is removed, bodies of different masses fall at the same rate of acceleration. But before Aristotle's view was empirically refuted, Galileo supposedly refuted it with an act of pure thought. He did so by asking us to consider a

[1] Unless otherwise referenced, more extensive descriptions of the thought experiments described here can be found in (Brown, 1991).

cannon ball and a musket ball attached by a string and asking what will happen when this entire assembly is dropped from the top of a tower.

According to Galileo, Aristotle would be forced into a contradiction. On the one hand, the combined system of musket ball, cannon ball, and string is heavier than the cannon ball alone. Therefore, the whole should fall faster than its parts. On the other hand, the lighter musket ball would try to fall more slowly than the heavier cannon ball, and would act as a drag on the entire system. Therefore the assembly should fall slower than its heaviest component, the cannon ball. So we have a contradiction. If C is the falling rate of the cannon ball, and S is the falling rate of the combined system, then both $S > C$ and $S < C$ are true. So Aristotle must be wrong. But the whole problem goes away if we propose that everything falls at the same rate ($S=C$). Brilliant. But what is justifying this line of reasoning?

The theory that is clearly on trial here is Aristotle's. But another theory is being brought to bear here as well - our folk theory (or theories) of how objects behave in the real world. To see this, we have to poke around for the assumptions Galileo makes. The celebrated contradiction is derived from two inferences. First, Galileo assumes that if Aristotle is right, then the whole assembly should fall faster than the cannon ball alone. Why? Because if you put the whole assembly on a scale, it would weight more than the cannon ball. So far so good. The second inference is that the musket ball must act as drag on the cannon ball. But what justifies this inference? Why does the small ball retain its autonomy as a lightweight when it becomes part of a heavier whole?

The answer to the above seems to be that we just *know* that it does! Imagine holding the two balls attached with string. If you put the musket ball in your left hand, and the cannon ball in your right, with the string stretched between the two, you just *know* that your right arm will tire more quickly. Our folk theory of the world, based on real experience, says that parts of wholes can in some cases be experienced as if they were autonomous. It is this folk theory that justifies the crucial second inference and allows Galileo to complete the contradiction. But what if our folk theory had been wrong? What if, despite our shared experience with lifting assemblies of objects, such assemblies in free fall actually do behave as integrated wholes? Well then, Galileo's conclusion would have been wrong, too. The moral of the story is that looking past the first theory on trial to the second (in this case unarticulated) theory that it is confronting leads us away from the idea that Galileo actually proved something through pure thought, and towards the more productive idea that he demonstrated that two theories were at odds and therefore one of them had to be discarded in favour of the other. The obvious choice for most people was to let go of Aristotle.

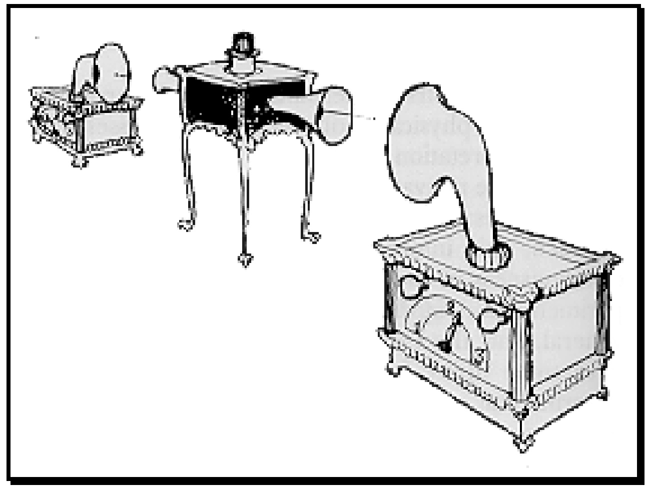

Figure 1: The apparatus for the EPR Paradox thought experiment. Picture from www.reed.edu/~rsavage/epr.html

The EPR Paradox[2] is another of Brown's favorite platonic thought experiments, although this is a slightly less impressive story because in the end the conclusion of the thought experiment was tested and refuted in a real experiment. The target of the EPR Paradox was Heisenberg's uncertainty principle in particular, and the Copenhagen interpretation of quantum mechanics in general. Heisenberg's uncertainty principle states that we can never know the complete trajectory of a quantum particle. "Trajectory" here means position plus momentum, where momentum is a vector that encodes mass, as well as direction and velocity. The problem is that in order to observe a quantum particle, we have to bounce another quantum particle off it, and this interaction will necessarily change one of the two aspects of its trajectory. The Copenhagen interpretation of this result stated that the indeterminacy existed not in our knowledge of reality, but in reality itself. That is, the trajectory of a quantum particle is not actually fully determined until it is observed.

Einstein, Podolsky and Rosen did not accept the Copenhagen interpretation - they believed that the theory was incomplete, not that reality was indeterminate. To prove their point, they published a 1935 *Physical Review* article in which they proposed the following thought experiment. Imagine two quantum particles that interact and then fly apart. We could later measure one of the particles for either momentum or position and automatically, through conservation laws, deduce the corresponding property in the other particle, even though the particles were far apart and could no longer interact. (Einstein et al., 1935)

The crucial step in the EPR reasoning was to notice that we could make *either* measurement of the first particle, and deduce the corresponding measurement of the second

[2] The name of this thought experiment is rather misleading. "EPR" stands for Einstein, Podolsky and Rosen, but don't expect to find an actual paradox anywhere.

particle. Since the two particles are now far apart, and could not be influencing each other, the second particle must have predetermined real values for both position and momentum. It must not be indeterminate with respect to any property. Einstein et al. were aware that they were making a crucial assumption here, namely that distant particles cannot influence each other. They named this the *locality assumption* and even went as far as to suggest that quantum mechanics be modified to include this assumption, thus making it a complete theory again. That is to say, they put the Copenhagen interpretation into the ring with another theory (one that Brown calls *local realism*). Once again, as seems to happen so often in successful thought experiments, the second theory accords well with our folk intuitions about the world - so much so that the EPR Paradox became a powerful tool for the anti-Copenhagen crowd. Unfortunately for them, John Bell in 1964 managed to derive a mathematical result that could be used to pit local realism against the Copenhagen interpretation in a *real* experiment, and local realism made the wrong prediction. (See Brown, 1991 and Bell, 1987)

The main point in all of this is that we get further in understanding these thought experiments if we always assume that two theories must be doing battle. Brown's Platonic view of what Galileo and Einstein et al. were up to does not help us to decide whether their conclusions were right or wrong. That is to say, the assumption that they were exploring abstract universals does not help us decide whether the particular universals they discovered correspond to the universals that happen to operate as natural laws. Furthermore, the Platonic view does not give us any real insight into what the thought experimenters were actually up to in either case. Specifically, there is nothing in the Platonic story to explain what justified Galileo's reasoning or why the EPR Paradox failed in the end. Platonism just doesn't buy us anything here. On the other hand, viewing their work through the lens of Dueling Theories forces us to make clear which theories were being tested, leading in the case of Coupled Falling Bodies to the exposure of an unstated piece of folk theory, and in the case of the EPR Paradox, to a simple account of why the result, which seemed so persuasive, did not stand up in a real experiment.

Other Types of Thought Experiments

Although Brown focuses most of his attention on Platonic thought experiments, his taxonomy actually includes a number of other categories as well. Brown's full set of thought experiment categories is reproduced in Figure 2 below. Thought experiments are broken down into two main groups, *destructive* and *constructive*. Roughly speaking, constructive thought experiments build new theories, while destructive thought experiments invalidate old ones. Constructive thought experiments are further subdivided into *direct* (those that begin with common, unproblematic phenomena and end with a well articulated theory), *mediated* (those that start with a well articulated theory and help to reach a new conclusion), and *conjectural* (those that start with conjectured, rather than common and unproblematic phenomena and end with a well articulated theory.) Platonic thought experiments, such as Coupled Falling Bodies and the EPR Paradox, are those that are both destructive *and* directly constructive. That is, they start from common, unproblematic phenomena and end in both the destruction of an old theory and the construction of a new one in it's place.

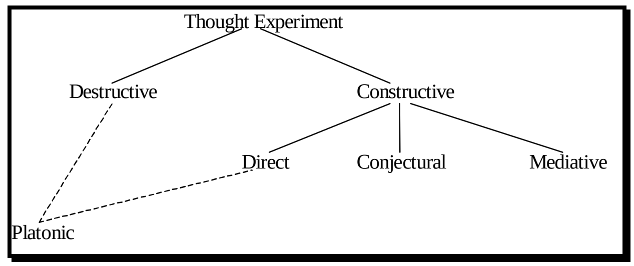

Figure 2: Brown's taxonomy of thought experiments. Adapted from (Brown, 1991)

Most of the examples Brown uses in any category are well accounted for by the Dueling Theories idea, but in some of Brown's examples, one has to dig a little deeper to get at the "other" theory involved. Perhaps the best example of this is Schrödinger's Cat, a destructive thought experiment aimed at the Copenhagen Interpretation of quantum mechanics (Schrödinger, 1935). In this experiment, a quantum event is amplified to have an effect on the macroscopic world - a cat in a box either dies or does not. The idea is that if the Copenhagen interpretation is correct, and reality is indeterminate, then the cat in the box is both alive and dead at the same time. The second theory coming into play here is just our deeply entrenched theory of the determinacy of the world around us. Calling us on this theory is a brilliant move, in which Schrödinger bets that scientists can accept quantum uncertainty only as long as they can cut it off from their folk intuitions about how the macroscopic world works - a sort of philosophical "not in my back yard" mentality.

The only problematic area for the Dueling Theories story is in Brown's "mediative constructive" category. Mediative experiments seem to play an articulatory role only. In these cases, a theory is already well established, and the thought experiment behaves like a diagram illustrating some particular, perhaps counter-intuitive, aspect of it. Most of the examples here are difficult to analyze for competing theories, but I suspect that is because they are not the same sort of thing as the thought experiments in the other categories. Consider one of Brown's examples - Maxwell's demon. One of the consequences of the theories Maxwell was developing was that heat could (with infinitely small probability) flow from a cold body to a hot body, rather than the other way around. This is, of course, wildly counter-intuitive. So Maxwell told a cute little story about a chamber of hot gas molecules and another of cold gas molecules connected by a gate. A little demon controls the gate and only lets fast (hot) molecules enter the hot chamber and slow (cold) molecules enter the cold chamber. All this story really does for us is provide a framework for understanding a counter-intuitive theory. Maxwell was not trying to prove anything new, and as a result neither Brown's conception of thought experiments nor my own has much to say about his thought experiment. I suspect we would do much better to reclassify mediative thought experiments as "thought experimental illustrations" and leave it at that.

Thought Experiments in Cognitive Science

The most famous thought experiments in the cognitive sciences are destructive in nature. This is true of both Putnam's Twin Earth experiment, in which the target is all non-reference-based theories of semantic content, and of Jackson's Mary the Colorblind Scientist experiment, in which the target is the scientific describability of phenomenal experience. It is also true of Searle's (1980) Chinese Room, arguably the most famous thought experiment in cognitive science (practically talked to death in the last 20 years.) For reasons of space, I will apply the Dueling Theories analysis in detail only to Searle's experiment - I will not develop the analysis of Putnam or Jackson any further. I want to show: a) that the Dueling Theories analysis can be applied, and b) that it would lead Searle's opponents directly to a particular line of attack.

The theory that Searle is attempting to demolish with the Chinese Room is the "Strong AI" theory that a computer programmed to behave indistinguishably from a human who understands Chinese would *really understand* Chinese. The thought experiment works by trading on our strong folk intuition (i.e. theory) that a non-Chinese-speaking human taught to behave like [a computer programmed to behave indistinguishably from a human who understands Chinese] would *not really understand* Chinese. Laying out the two theories explicitly makes it easy to see that in order for them to be in conflict, the computer and the human must be equivalent in their ability to execute computer programs. But that would clearly not be the case in the real world. Computers are much better and faster at executing programs than humans. Therefore, Searle must have a fictional super-human in mind, which leads us to ask why we feel entitled to any a priori judgements about what such humans would or would not "understand" about the task they were performing. The Dueling Theories analysis thus leads directly to a productive line of attack on Searle - one that has, in fact, been pursued in the literature (Hofstadter and Dennett, 1982).

The Empirical Basis of Thought Experiments

Brook's (1999) paper emphasizes a slightly different way of looking at thought experiments - one which seems at least partially at odds with Brown's conception, and which differs in emphasis from mine. He asks:

> …are thought experiments also empirical, at least in part? Yes; they are merely a particular way of manipulating material stored in memory, material originally gained from experience. (p. 104)
>
> Since the materials and at least most of the relationships of the imagined situation are derived from experience, thought experiments are thus a *kind* of empirical investigation. (p. 106, Brook's italics)

Two related questions lurk in the above. First, what is what the nature of the "material stored in memory", and second, what is the sense of "empirical" that Brook is employing?

Concerning the first question, it seems clear that the material referred to must not simply be memories of events that the thought experimenter has witnessed. If that were the case, then most observers would have no opinion about what might happen even in the somewhat ordinary situation imagined in the Coupled Falling Bodies experiment, to say nothing of the strange situations of the EPR Paradox, Twin Earth, Schrödinger's Cat, the Chinese Room, and Mary the Colorblind Scientist. The material stored in memory must have some kind of predictive power in order for it to be applied to the novel situation of a thought experiment. Therefore, it must consist of generalizations over past experience - that is, theories, folk or scientific, about the world.

Concerning the second question, the word "empirical" is a loaded one, and calling an activity empirical can have a number of different meanings. In the strongest sense, an empirical activity might consist of observing events in the world and carefully recording them for future use. But that is clearly not the sense being used here. In a weaker sense, an empirical activity might be any activity that takes into account memories of actually experienced events. But for reasons stated above, this cannot the correct sense either - we could not perform thought experiments based solely on remembered events. But there is another sense in which an activity is empirical if it involves generalizations (theories) that were derived, at least in part, from actual empirical experiences. This must be closest to the sense that Brook has in mind, and it also fits reasonably well with the Dueling Theories idea.

Brook goes on to situate thought experiments within an abductive, Popperian view of scientific progress, pointing out that thought experiments have historically played a role on both sides of the generate and test paradigm. It is easy to see how hypothesis generation can be aided by thought experiments. As Brook notes, "hypothesis generation is pretty much a pure act of the imagination." As such, it's not hard to see that thought experimentation can help here. Lots of counterfactual "what ifs" are bound to be involved, each of which is likely to be a thought experiment. On the Dueling Theories account, a thought experiment can also help to crystallize a folk theory that previously went unnoticed. This is not exactly the same thing as generating a hypothesis, but for the purposes of scientific investigation, it is the same thing - in order to test a theory, it must first be articulated and acknowledged.

Where I take some issue with Brook is concerning the role of thought experiments on the testing side of the generate and test paradigm. On the one hand, Brook makes the uncontroversial claim that thought experiments *have* historically played a role in testing hypotheses - the paradigm case being Galileo's rejection of Aristotelian physics based on the Coupled Falling Bodies experiment. Where Brook and I may part company is on the question of whether thought experiments *ought to* play such a strong role in testing hypotheses. Recall that on the Dueling Theories account, Galileo's thought experiment made use of an unarticulated folk theory of ordinary objects to make it's point. Hence, his conclusion should not have been that Aristotle's theory was ruled out as a possibility but that the theory ran afoul of some very ordinary intuitions about the world. This is a very important result in and of itself. It's just that the appropriate reaction was not for Galileo to smile triumphantly from his armchair, but for him to get up and figure out how to test the two theories with a *real* experiment.

The goal of Brook's treatment (and my own) is to see how much thought experiments can do for us in the cognitive sciences. Brook's emphasis on the empirical nature of thought experiments leads to the idea that they can in some cases stand in for real hypothesis testing.[3] What I hope to have shown with the Dueling Theories account is that the usefulness of thought experiments actually lies in their meta-level ability to test theories of the world against each other, and not in their ability to test the world directly. There are two further reasons why I would avoid using the term "empirical" to describe thought experiments - one technical and one psychological. The technical reason is that folk theories and intuitions may not need to be based on anything empirical at all. In fact, there are some thought experiments, such as Putnam's Twin Earth, Searle's Chinese Room, and even the EPR Paradox for which it is difficult to nail down exactly what the relevant empirical evidence supporting our intuitions actually is. For that reason, it's safer to view thought experiments as a theory evaluation activity and leave the question of the empirical nature of the theories under test open to be evaluated on a case by case basis. The psychological reason is that we might mistakenly encourage the idea that thought experiments are empirical in a stronger sense than what we really mean. And if we do that, we've done ourselves a disservice.

Conclusion

If I am right, then thought experiments are best viewed as a picturesque arena in which two competing theories of the world do battle. Sometimes the Dueling Theories are explicitly stated ahead of time, but often one of the two lurks in the background assumptions. In fact, what gives thought experiments such as Coupled Falling Bodies and the EPR Paradox their rhetorical power is precisely their accord with some of our most deeply entrenched folk theories of the world of middle-sized dry goods. Thought experiments are more useful if they clearly contain some kind of empirical material. This empirical material is not direct observation, but instead takes the form of predictions encoded by theories formed on the bases of empirical data. But this is not the same as saying that thought experiments can be empirical in the way that real experiments are. Despite their potential for polemical power, thought experiments can never actually tell us anything about the world. No thought experiment can ever be as good as the corresponding real experiment.

[3] From personal communication. Brook also discusses the use of thought experiments for the elimination of possibilities. This idea also relies on the characterization of thought experimentation as an empirical activity.

Acknowledgements

I would like to acknowledge the help and support of Andy Brook. Thanks also to Ron Boring, Leo Ferres, and Finn Makela for many useful discussions. This work was supported in part by Canada's National Sciences and Engineering Research Council (NSERC).

References

Bell, J. S. (1987). *Speakable and Unspeakable in Quantum Mechanics*. Cambridge: Cambridge University Press.

Brook, A. (1999). Does philosophy offer cognitive science distinctive methods? *Proceedings of the 21st Annual Conference of the Cognitive Science Society* (pp. 102-108). New York: LEA.

Brown, J. R. (1991). *Thought Experiments in the Natural Sciences*. London: Routledge.

Einstein, A., Podolsky B. and Rosen N. (1935). Can quantum mechanical description of reality be considered complete? *Physical Review*.

Hofstadter, Douglas R. and Dennett, Daniel C. (1982) Reflections. In Douglas R. Hofstadter and Daniel C. Dennett (Eds.) (1982). *The Mind's I: Fantasies and Reflections on Self and Soul*. Toronto: Bantam Books. 373-382.

Jackson, F. (1991). What Mary didn't know. *Journal of Philosophy*. 83-5: 291-295. Reprinted in D. Rosenthal (Ed.) *The Nature of Mind*. New York: Oxford University Press.

Putnam, H. (1975). The meaning of meaning. *Mind, Language and Reality: Philosophical Essays*. Cambridge: Cambridge University Press.

Schrödinger, E. (1935). The present situation in quantum mechanics. Translated and reprinted in J. Wheeler and W. Zurek, *Quantum Theory and Measurement*. Princeton, NJ: Princeton University Press.

Searle, John R. (1980). Minds, brains, and programs. *BBS*. Volume 3. Reprinted in Hofstadter and Dennett. *Op. Cit.* 353-373.