Galileo and the Indispensability of Scientific Thought Experiment

Tamar Szabó Gendler

ABSTRACT

By carefully examining one of the most famous thought experiments in the history of science—that by which Galileo is said to have refuted the Aristotelian theory that heavier bodies fall faster than lighter ones—I attempt to show that thought experiments play a distinctive role in scientific inquiry. Reasoning about particular entities within the context of an imaginary scenario can lead to rationally justified conclusions that—given the same initial information—would not be rationally justifiable on the basis of a straightforward argument.

1 Argumentative reconstruction

Philosophers who are opposed to all things spooky tend to think that thought experiments in science are (at least in principle) eliminable, and that whatever demonstrative force they have is the result of their being sound arguments dressed up in heuristically appealing clothing. On such a view, a scientific thought experiment's justificatory force comes from the fact that it can be reconstructed as an argument with explicit premises that make no reference to imaginary particulars.

My goal in this article is to challenge this view by carefully examining one of the most famous thought experiments in the history of science: that by which Galileo is said to have refuted the Aristotelian theory that heavier bodies fall faster than lighter ones. I will try to show that the thought-experimental format of Galileo's presentation plays an indispensable role in the persuasiveness of his case against the Aristotelian, and that a similar degree of persuasiveness could not be obtained on the grounds of explicit argument alone.

1.1 The Elimination Thesis

The view that thought experiments lead to justified conclusions because they are arguments finds clear articulation and powerful defence in a recent pair of papers by John Norton.[1] In those papers, Norton puts forth a hypothesis about thought experiments which he calls the *Elimination Thesis*. Paraphrasing slightly, his thesis is this:

The Elimination Thesis: Any conclusion reached by a (successful) scientific thought experiment will also be demonstrable by a non-thought-experimental argument.[2]

As initially formulated, the thesis is ambiguous; it is compatible with both a weaker reading, which I will call the *Dispensability Thesis*, and a stronger reading, which I will call the *Derivativity Thesis*.[3] In order to formulate these versions, however, I need first to clarify what a number of terms in the thesis mean, and to describe briefly the sorts of arguments which might be offered in its favour.

1.2 Clarification of terminology

A number of terms in the thesis require further elaboration. Let me begin with 'thought experiment' and 'non-thought-experimental argument'. To draw a conclusion on the basis of a *thought experiment* is to make a judgement about what would happen if the particular state of affairs described in some imaginary scenario were actually to obtain.[4] One might then use that judgement in

---

[1] Norton [1991] and [1996].

[2] Cf. Norton [1991], p. 131 and Norton [1996], p. 336. I should make it clear at the outset that although I am using Norton's articulation and defence of this thesis as a convenient jumping-off point, the specifics of his position are not my target. For a detailed comparison of my views with Norton's, cf. Section **3.4** below.

[3] Cf. Norton [1996], pp. 354–8.

[4] For other recent characterizations of (scientific) thought experiments see, for instance: Bealer [1998]; Brown [1986, 1991a, b, 1993a, b, 1995]; Carrier [1993]; Gooding [1990, 1992, 1993, 1994]; Hacking [1993]; Humphreys [1993]; Irvine [1991]; Janis [1991]; Kujundzic [1993]; Laymon [1991]; Lipton [1993]; Massey [1991]; Nersessian [1984, 1992, 1993]; Norton [1991, 1993, 1996]; Prudovsky [1989]; Shepard [1994]; Sorensen [1992a, b]; Wilkes [1988]; as well as other articles in Horowitz and Massey [1991]. The *locus classicus* is, of course, Mach [1883/ 1960, 1897/1976]; for other early discussions cf. also Hesse [1966]; Koyré [1960]; Kuhn [1964]; Popper [1959/1992]. For current discussion, cf. Souder [undated].

developing a more general theory, just as one might use the result of an actual experiment. By contrast, to draw a conclusion on the basis of a *non-thought-experimental argument* is to be led by a process of inductive or deductive reasoning from a set of explicit premises which make no reference to particular hypothetical or counterfactual states of affairs to a correspondingly general conclusion. Again, one might use that conclusion as the basis for endorsing one or another general theory about the phenomena in question. So thought experiments differ from non-thought-experimental arguments in two crucial respects: first, they are not presented as arguments, but rather as invitations to contemplate a way that the world might (have) be(en); and second, they make essential reference to particular hypothetical and counterfactual states of affairs.[5]

What the Elimination Thesis says is that any good scientific thought experiment can be transformed into a non-thought-experimental argument without loss of demonstrative force. Given the characterization just offered, what an elimination will involve is first a process of argumentative reconstruction in which the narrative presentation is replaced by a series of explicit premises sufficient to establish the desired result, and then a process in which those premises that make reference to hypotheticals, counterfactuals, and particulars are replaced by premises in which no such reference is made. If the Elimination Thesis is correct, such a process will preserve completely the thought experiment's demonstrative force.

What is meant by 'demonstrative force'? I will suggest two problematic readings and then one that I will endorse. If the claim that 'any conclusion reached by a good thought experiment will also be demonstrable by a non-thought-experimental argument' means no more than that in the reconstruction of a mature science, the conclusions that were (as a matter of fact) reached by thought experiments can be derived from more fundamental principles by means of inference schemes licensed within the science, then the thesis is trivially true. Even if the development of Newtonian mechanics relied on a series of crucial thought experiments, its textbook presentation might well establish particular conclusions on the basis of more conventional forms of argument.

On the other hand, there is a reading of 'demonstrative force' according to which the thesis is trivially false. If the claim is taken to mean that—as a matter of psychological fact—any conclusion which was reached by a good thought experiment might also have been demonstrated to the person who reached the conclusion by means of a non-thought-experimental argument, then the Elimination Thesis is certainly false. None would doubt the important heuristic and illustrative role played by thought experiments in

---

[5] Cf. Norton [1991], p. 129 and [1996], p. 336.

scientific exploration, and the crucial tasks they play in instruction and informal demonstration.

The proper reading of 'demonstrative force' makes the Elimination Thesis epistemologically interesting. On this reading, demonstrative force concerns the role that thought experiments play in living bodies of knowledge: after the moment of discovery and before the end of inquiry. It concerns whether a particular conclusion based on a particular process of reasoning (thought experiment) is thereby *justified*—whether if such a process leads to true beliefs, those beliefs should count as knowledge. So the issue raised by the Elimination Thesis is this: can reasoning about (reasonably) specific entities within the context of an imaginary scenario lead to rationally justified conclusions that—given the same initial information—would not be rationally justifiable on the basis of a straightforward argument? In the next section, I consider two reasons that one might think such eliminations are possible.

1.3 The negative argument and the positive argument

The Elimination Thesis (that is, the thesis that thought experiments are dispensable) can be defended with two arguments, one negative, one positive. So, for instance, John Norton argues (negatively) as follows: thought experiments must be arguments because there is nothing else for them to be.

> Thought experiments in physics provide or purport to provide us informa-
> tion about the physical world. Since they are *thought* experiments rather
> than *physical* experiments, this information does not come from the
> reporting of new physical data. Thus there is only one non-controversial
> source from which the information can come: it is elicited from information
> we already have by an identifiable argument (Norton [1991], p. 129).

Norton considers this position almost trivial: 'the alternative,' he writes, 'is to suppose that thought experiments provide some new and even mysterious route to knowledge of the physical world' (Norton [1991], p. 129).[6] So the negative argument contends that if we have obtained new information about the empirical world without having obtained new *empirical* information about the empirical world, the only way we *could* have done so is by means of an argument.

The positive reason that one might think that thought experiments are just arguments in disguise is that 'the *analysis* and *appraisal* of a thought experiment will involve reconstructing it explicitly as an argument', so that 'a good thought experiment is a good argument, a bad thought experiment is a bad argument' (Norton [1991], p. 131; [1996], p. 335). So the positive argument amounts to saying that if a thought experiment can be *reconstructed* as an

---

[6] For an endorsement of this alternative, cf. Brown [1991a, b, 1993a, b, 1995].

argument, then what it was all along *was* an argument.[7] Even if the reason I come to know something by contemplating a thought-experimental scenario doesn't *seem* to be because there is an argument into which the thought experiment can be reconstructed, it is. The reason my belief is *justified* is because, in the end, the thought experiment *was* a disguised argument all along.

1.4 The Dispensability Thesis and the Derivativity Thesis

We are now in a position to recognize that the Elimination Thesis as originally formulated and defended actually involves two distinct claims. These might be stated as follows:

The Dispensability Thesis: Any good scientific thought experiment can be replaced, without loss of demonstrative force, by a non-thought-experimental argument.

The Derivativity Thesis: The justificatory force of any good scientific thought experiment can only be explained by the fact that it can be replaced, without loss of demonstrative force, by a non-thought-experimental argument.

Loosely put, the Dispensability Thesis says that we can always get from here to there without appeal to a thought experiment. If a thought experiment legitimately transports us from one state of belief to another, a non-thought-experimental argument could too. Thought experiments may be convenient and efficient ways of reaching conclusions about the physical world, but they have only the advantage that a car has over walking; they get us where we want to go much more quickly, but they don't get us anywhere we couldn't reach by more pedestrian means.

The Derivativity Thesis says that not only can any good scientific thought experiment be replaced, without loss of demonstrative force, by a non-thought-experimental argument, but that to the extent that a good scientific thought experiment has demonstrative force, it is *because,* deep down, the thought experiment *is* an argument. We may be misled by the surface features of the case to think that something non-argumentational is doing justificatory work, but we are wrong. The *reason* the Dispensability Thesis is true is that all that was *ever* justificatorily at play was something argumentative. What looked like a car turned out to be propelled by foot power all along (like a child's go-car, or one of the vehicles on *The Flintstones*). So the Dispensability Thesis says we *can* get by without what we commonly call thought experiments; the Derivativity Thesis tells us that we already *do*.

---

[7] 'The workings and achievements of any thought experiment can be revealed and captured fully in an explicit argument which employs the same resources' (Norton [1996], p. 339; cf. also ibid., pp. 357–8).

In the next section, I challenge the Dispensability Thesis by showing that it does not hold true of a widely acclaimed thought experiment of Galileo's. I choose this case for two reasons. First, since this particular example is generally treated as the paradigm of an effective thought experiment, diagnosing the source of its success is itself a worthwhile endeavour.[8] Second, challenging the Dispensability Thesis in this way allows me to shed light on the Derivativity Thesis as well. Obviously, if the Dispensability Thesis is false, the Derivativity Thesis is too; the more interesting question is whether some alternative explanation can be offered of the thought experiment's success. I try to say something positive about this question in Section 3.

2. Galileo's thought experiment and its reconstruction

2.1 Galileo's thought experiment

Perhaps the most famous thought experiment in the history of Western science is the thought experiment with which Galileo is credited with having refuted the Aristotelian view that the speed with which a body falls is directly proportional to its weight.[9] The thought experiment appears in his last and most mature work, the *Discourse Concerning Two New Sciences*, in the context of a more general discussion of the possibility and nature of motion in a void.[10] Galileo's goal in the section as a whole is to establish that 'if one were to remove entirely the resistance of the medium, all materials would

---

[8] For some of the many discussions of this and related thought experiments of Galileo, cf. Brown [1986, 1991a, b, 1993a, b, 1995]; Cargile [1987]; Clement [1983]; Koyré [1939/1979, 1960]; Kuhn [1964]; Norton [1996]; Prudovsky [1989]; Sorensen [1992a]. In my discussion below, I follow the somewhat unfortunate practice of considering this thought experiment outside of both its historical and textual contexts. As a partial remedy to this misleading presentation, I refer the reader to some of the many general discussions of Galileo's work and its context; one might fruitfully begin with: Butts and Pitt [1978]; Claggett [1959]; Clavelin [1974]; Cooper [1935]; Damerow *et al.* [1992]; Dijksterhuis [1961]; Drake [1978, 1989, 1990]. See also sources listed in the next three footnotes.

[9] Challenges to the Aristotelian thesis—both empirical and conceptual—had appeared in a number of mid- and late sixteenth-century works. (For relevant passages, see: (Cardan) Cooper [1934], pp. 7–77; Damerow *et al.* [1992], p. 365; (Tartaglia) Drake and Drabkin [1969], pp. 63–143, esp. 120ff.; Damerow *et al.* [1992], p. 378; (Benedetti) Drake and Drabkin [1969], pp. 147–237, esp. 206; Drake and Drabkin [1969], pp. 31–41; Dijksterhuis [1961], pp. 269–71; Drake [1989], pp. 27–30; (Stevin) Cooper [1935], pp. 77–80; Dijksterhuis [1961], pp. 324–9.)

[10] Galileo himself had produced a less conclusive version of the famous thought experiment as early as 1590 in an unfinished and unpublished dialogue on motion; cf. Cooper [1935], pp. 80–90; Drake and Drabkin [1969], pp. 331–77; Galilei [1590/1960], pp. 26–38, esp. 29 (National Edition, p. 265). For an interesting discussion of whether Galileo ever actually performed such an experiment, cf. Cooper [1935]; Drake [1978, 1989, 1990]; Drake and Drabkin [1969]; Koyré [1960]; Segre [1989]. Since my primary purpose in this paper is not historical, I will focus only on Galileo's 1638 presentation of the refutation, bracketing the interesting question (a question not without philosophical interest) of why it was that such a simple and obvious mistake remained part of the West's scientific world view for nearly 2000 years.

descend with equal speed' (Galilei [1638/1989], p. 116);[11] the thought experiment in question leads to the weaker conclusion that 'both great and small bodies, *of the same* [*material*], are moved with like speeds' (Galilei [1638/1989], p. 109, italics added, bracketed word replaced).

The view that Galileo is challenging is that 'moveables differing in heaviness are moved in the same medium with unequal speeds, which maintain to one another the same ratio as their weights [*gravità*]' (Galilei [1638/1989], p. 106). That is, he is challenging the view that heavier bodies fall faster than lighter ones, and that they do so in direct proportion to their heaviness. On the version Galileo takes himself to be opposing, the proportionality is linear; 'a moveable ten times as heavy as another, is moved ten times as fast' as the other (Galilei [1638/1989], p. 106).[12]

The famous thought experiment, rephrased slightly, is the following. Imagine that a heavy and a light body are strapped together and dropped from a significant height.[13] What would the Aristotelian expect to be the natural speed of their combination? On the one hand, the lighter body should slow down the heavier one while the heavier body speeds up the lighter one, so their combination should fall with a speed that lies between the natural speeds of its components. (That is, if the heavy body falls at a rate of 8, and the light body at a rate of 4, then their combination should fall at a rate between the two (cf. Galilei [1638/1989], p. 107).) On the other hand, since the weight of the two bodies combined is greater than the weight of the heavy body alone, their combination should fall with a natural speed greater than that of the heavy body. (That is, if the heavy body falls at a rate of 8 and the light body with a rate of 4, their combination should fall at a rate greater than 8.) But then the combined body is predicted to fall both more quickly, and more slowly, than the heavy body alone (cf. Galilei [1638/1989], pp. 107–8). The way out of this paradox is to assume that the natural speed with which a body falls is independent of its weight: 'both great and small bodies ... are moved with like speeds' (Galilei [1638/1989], p. 109).

2.2 Reconstruction of the Galileo case

Transformed into an argument that conforms to the strictures of the Elimination Thesis, Galileo's reasoning can be reconstructed as follows.[14]

---

[11] Here and elsewhere, I have made use of Drake's 1974/1989 translation of the *Discorsi*. For easy cross-referencing with the more widely available (though less reliable) Crew and De Salvio translation, page references are to the National Edition (except where noted).

[12] Cf. Aristotle, *Physics*, 215a24–216a21; *On the Heavens (De Caelo)*, 301b.

[13] Note that in the remarks that follow, all references to bodies should be understood as referring to bodies of the same material. For the purposes of my discussion, this constraint is irrelevant.

[14] For Norton's reconstruction of this case, cf. Norton [1996], pp. 340–5; I discuss his version in Section **3.4** below.

The first claim of the Aristotelian is that:

(1) Natural speed is mediative.

That is, natural speed is a property such that if a body A has natural speed s_1, and a body B has natural speed s_2, the natural speed of the combined body A–B will fall between s_1 and s_2.

The second premise of the reconstruction is that:

(2) Weight is additive.

That is, weight is a property such that if body A has weight w_1, and body B has weight w_2, the weight of the combined body A–B will be equal to the sum of w_1 and w_2.

From these two premises (plus the assumption that not all weights and natural speeds are either zero or infinite), it follows that:

(3) Natural speed is not directly proportional to weight.

For the first is a mediative property, whereas the second is an additive property, and a mediative property cannot be directly proportional to one that is additive. Furthermore, the only way to maintain (1), (2), and (3) simultaneously is to assume that all natural speeds are the same. Then weight might be additive and natural speed (in a vacuous sense) mediative, with no contradiction thereby implied. Thus natural speed is shown to be independent of weight.

2.3 Four ways out for the Aristotelian

If the Dispensability Thesis is true, then Galileo's thought experiment should be replaceable by some non-thought-experimental argument without loss of demonstrative force. My goal in the next two sections (**2.3** and **2.4**) is to show that the reconstruction presented in Section **2.2** is not such an argument.[15]

I begin my case by pointing out that there are a number of 'ways out' for the defender of the view that natural speed is directly correlated with weight—a view which, for the sake of convenience, I will call the Aristotelian view. These ways out involve denying premises (1) and (2) by proposing a series of alternative hypotheses about the physical properties of strapped-bodies, that is, bodies of the sort described by the thought experiment. The point of talking about these ways out is to show that there are ways to maintain the negation of (3) by adopting alternatives to (1) and (2), and adopting these may well be less disruptive to the Aristotelian picture than adopting (3). What I will suggest below is that these ways out, though logically available, run counter to certain tacit knowledge about the physical world. It is for this reason that, when the

---

[15] Below I suggest reasons for thinking that this will be true for *any* argumentative reconstruction that conforms to the strictures of the elimination thesis.

case is presented as a thought experiment, they do not even occur to us. To block them as moves in a straight argument, however, requires metaphysical commitments that seem not to be at work in the thought experiment itself. What these commitments are, and what role I think they *actually* play in Galileo's reasoning, is a point I will turn to after presenting the four ways out.

The first ways out would be for the Aristotelian to deny that the properties in question are *determinate* for strapped-together bodies in one of the following two ways.

(4) Natural speed is not physically determinate for strapped-bodies.[16]

(5) Weight is not physically determinate for strapped-bodies.

That is, she might reject (1) or (2) on the grounds that they presuppose that natural speed and weight are properties that apply universally, even to bodies that are in some way monstrous.[17] Since strapped-bodies are odd entities, she might say, they need not be governed by the sorts of laws that govern ordinary objects. In particular, they need not have determinate natural speeds or weights.

The third way out for the Aristotelian would be to avoid the conflict between (1) and (2) by saying that there *is* a fact of the matter about whether a strapped-body is one body or two, and that its physical properties in falling will depend on the answer to this question. She might say:

(6) Natural speed and weight are mediative for strapped-bodies that are *united.* Natural speed and weight are additive for strapped-bodies that are *unified.*

That is, sometimes when two bodies are strapped together, they are merely *united* and remain, as a matter of fact, two objects; sometimes, when they are strapped together, they are *unified* and form, as a matter of fact, a single object. In the first case, both weight and speed will be mediative; so that the combined body will have a weight intermediate between those of the two original bodies, and fall with a natural speed that lies between the two original speeds. In the second case, both properties will be additive; so that the weight of the unified body will be equal to the sum of the weights of its component parts, and its natural speed correspondingly equal to the natural speeds of the two combined. Since the mediativity of the properties holds only with respect to united pairs of objects, and the additivity only with respect to unified single objects, there is no

---

[16] For a version of this 'way out', cf. Koyré [1960], p. 51.

[17] Galileo pre-emptively deals with this by getting a concession from Simplicio straight away that 'for every heavy falling body there us a speed determined by nature such that this cannot be increased or diminished except by using force or opposing some impediment to it' (Galilei [1638/1989], p. 107); cf. also Drake's fn. 40 in Galilei [1638/1989] at (modern pagination) p. 66.

way that the Aristotelian can be forced to a contradiction. What she *is* forced to accept, however, are radical discontinuities in nature. A body, united, might be falling steadily at a rate of, say, six, and suddenly, should its parts happen to become unified, begin falling at a rate of, say, twelve.

But the Aristotelian can avoid the problem of discontinuity. A fourth way would be for her to say that given two bodies that fall together there is a fact of the matter about their degree of connectedness, and that this determines their physical properties when falling. The claim would be:

(7) Natural speed and weight for strapped-bodies are determined by a *degree of connectedness* (C) such that the speed/weight of B_1-strapped-to-B_2 where B_1 has w_1 and B_2 has w_2 will be: $(C)(w_1 + w_2) + (1 - C)((w_1 + w_2)/2.$[18]

We let C measure the degree of connectedness between the two bodies; that is, we let it be a number between zero and one that corresponds to the degree to which two bodies that fall together are unified: if the bodies are completely unified, C will take a value of one; if the bodies are completely disunified (that is, united), C will take a value of zero. For intermediate cases, the value will be between these two, and the speed and weight of the combined body will lie between the mean and the sum of the two initial values. So if the two bodies are completely united, the additive law will apply completely; if the bodies are merely unified, the mediative law will apply throughout; and for intermediate cases, some proportional average will be found between them. Thus the assumption that natural speed is a function of weight can be maintained, and it can be maintained without violation of continuity. How? Under the assumption that degree of connectedness is a relevant physical property. That this way out too seems not to be a live option brings us to the point where I will suggest my alternative explanation of what is going on.

2.4 What the reconstruction misses

I want to begin by thinking about ways in which the four ways out might be blocked. And I want to start with the most obvious: appeal to two broad, defeasible, tacit assumptions, each of which captures an important feature of our representation of experienced reality. One is that, for any body that one might encounter, there is a determinate fact concerning its weight and natural speed. That is:

(8) Natural speed and weight are physically determined.

---

[18] To keep the equation minimally complicated, I have made a number of trivial simplifying assumptions. I have assumed that the units for measuring weight and natural speed correspond so that the number representing an object's weight is the same as the number representing its natural speed; and I have assumed that the natural speed of two merely unified bodies is the mean of their individual natural speeds.

The other is that there is *no* determinate fact whether strapped-bodies are one object or two. That is:

(9) Entification is not physically determined.

What (8) says is that a particular question about natural properties has a determinate answer. Any body, no matter how oddly shaped, will have a particular weight and a particular natural speed that are fixed by the world. What (9) says is that a particular question of entification has an indeterminate answer. Whether we consider a strapped-together body to be a single object, or two objects held together by a strap, or indefinitely many objects held together by internal forces, is merely a question of the aspect under which we choose to view that object. The answer to the question 'how many objects?' does not follow from any *physical* property we might discover; it is a question about our words, not a question about the world.

These two premises are sufficient to eliminate the 'ways out' enumerated above. If (8) holds, then (4) and (5) are not available as lines of escape; if (9) holds, then neither (6) nor (7) can be appealed to as a means of avoiding the Galilean conclusion. And the *way* in which they eliminate (4)–(7) is very different from the way that a simple reassertion of (1) and (2) would. They show not only *that* there is something wrong with (4)–(7) as descriptions of the way the world is, but *why* there is something wrong with them. They show *what* about our tacit understanding of physical reality, and of our instincts concerning plausible candidates for physically relevant and irrelevant properties, is missed by someone who appeals to (4) or (5) or (6) or even (7).

What this reveals is that the initial reconstruction of the Galilean thought experiment (as presented in Section **2.2**) fails to capture what is really doing the work in the case. As hypotheses about the ways strapped-bodies might behave in fall, (4), (5), (6), and especially (7) are in principle available as alternatives to (1) and (2); just as natural speed might be mediative and weight additive, it might be that both natural speed and weight depend on the degree to which the two bodies are connected. So if (1)–(3) were truly capturing what is going on in the Galilean thought experiment, there would be ways out for the Aristotelian that would allow her at least to shift the burden of proof back to the Galilean.

That these ways out do not seem available when the thought experiment is presented in its unreconstructed form shows that this eliminative reconstruction has failed to capture its original demonstrative force. (What has been lost is the way in which, by evoking tacit knowledge about the how falling bodies actually behave, the thought experiment pre-emptively precludes such ways out.) Accordingly, I tried to come up with a reconstruction sufficiently strong to rule out (4)–(7) in a similarly categorical and decisive manner. This involved appeal to two rather comprehensive and metaphysical-sounding

principles, namely (8) and (9). (8) and (9) give background support to (1) and (2) and thereby help to establish (3).

But just as (1) and (2) are too weak to capture the way in which alternative hypotheses concerning fall are ruled out by the thought experiment, (8) and (9) are too strong. They represent approximate articulations of defeasible assumptions about the physical world. But as they stand, they articulate principles that have less certainty than the conclusion they are taken to support. Prior to contemplation of the case Galileo describes, the Aristotelian may be committed to *something* like (8) and *something* like (9), but he is certainly not committed to them unmodified. To reconstruct the case as an explicit argument with some version of (8) and (9) among the premises would require enumerating outright their defeasibility conditions. But this is something he does not know how to do. Contemplation of the case Galileo describes *brings him to see* that these principles are not defeated in *this* case. And it is this recognition that serves as the basis for the case's power. No austere argumentative reconstruction will be able to do this, because part of the thought experiment's function is to bring the Aristotelian to accept certain *premises*. In the next section, I will discuss what makes belief in these premises *new*, and what makes it *justified*.

3 Denying the Dispensability and Derivativity Theses

3.1 Rejecting reconstruction: what the thought experiment does

If the Dispensability Thesis were correct, then the conclusion established by Galileo's thought experiment—that 'both great and small bodies ... are moved with like speeds' (Galilei [1638/1989], p. 109)—*should* be demonstrable by means of a non-thought-experimental argument. 'Demonstrable' here means: rationally justifiable on the basis of the same background conditions. So let us spell out what the background conditions in question are.

For the Aristotelian, daily experience seems to confirm the theory that heavier bodies fall faster than lighter ones. Just as the Galilean sees cases where lighter bodies fall more slowly than heavy ones as exceptional, so the Aristotelian sees as crying out for explanation those cases where the rate of fall is (nearly) equal. That is, when gold beaten into a very thin leaf reaches the ground more slowly than a solid lump of the same material, the Galilean must posit some factor that explains the divergence of this result from the generally expected outcome.[19] Similarly, when the Aristotelian sees two stones of very different weights fall to the ground with like speeds, the circumstance requires diagnosis and explanation.

---

[19] Cf. Galilei [1638/1989], p. 109.

So far all I have pointed out is the possibility of maintaining theoretical commitments in the face of apparent counter-evidence. This can be done by appealing to additional principles that explain away anomalous data by showing that the phenomena at issue are subject to the fundamental principle in question, but that the world's complexity has prevented them from manifesting this. So the Galilean might appeal to air resistance, the Aristotelian to the fact that the bodies have not been dropped from a height sufficiently great.[20] Such explaining-away of recalcitrant exceptions is not a desperate move by a failing paradigm; it is a fundamental element of doing normal science in a non-ideal world.[21]

The point of this discussion is to give a better sense of the background conditions under which the argumentative reconstruction must be demonstratively forceful if the Dispensability Thesis is to be shown to hold in this case. The thesis tells us that some non-thought-experimental reconstruction of the case Galileo presents should be able to do the same thing that the thought-experimental version does: lead the Aristotelian *from the same background assumptions to the same rationally justified conclusions*. The standard reconstruction presented in Section **2.2** fails in this regard, as does the strengthened version presented in Section **2.4**. And, I contend, *any* argument satisfying the Elimination Thesis is likely fail in the same way.

Why? Because prior to the thought experiment, the Aristotelian is explicitly committed to the *negation* of (3), and this background commitment serves as a filter through which apparently contrary evidence will inevitably be reinterpreted. Any argument which satisfies the Elimination Thesis (that is, any argument with explicit premises which make no reference to particular hypothetical or counterfactual states of affairs) can be reframed by the Aristotelian as a *reductio*. (1) or (2) or (8) or (9) will simply be *denied*, once their implications are made evident. What is remarkable about the thought-experimental presentation is that it is able to undermine this framework-shaping assumption.

---

[20] In Galileo's dialogue, when Salviati points out that a rock of two pounds and a rock of twenty pounds will strike the ground nearly simultaneously when dropped from a height of one or two hundred feet, Simplicio retorts that this may just be a result of not having given the objects enough falling-time for the differences to become apparent (see Galilei [1638/1989], pp. 109–10). Simplicio says: 'Perhaps from very great heights, of thousands of bracchia [discrepancies] would follow which [are] not seen at these lesser heights.' That is, Simplicio suggests that perhaps the reason we do not observe the sorts of differences which Aristotle's theory predicts is that we are making observations under non-idealized circumstances, and that were we to eliminate distorting elements—like the fact that the fall is only a few hundred feet—the true phenomena would reveal themselves. Galileo provides Salviati with a rather sharp retort (Galilei [1638/1989], p. 110). See also (pagination here refers to the modern text) Galilei [1638/1989], fn. 44 at p. 69.

[21] Of course, it is a commonplace in the history of science that at a certain point the burden of positing literal or metaphoric epicycles becomes too great, and the theory—especially when there is a simpler alternative available—collapses under the weight of its own internal complexity. But it is a similarly well-established commonplace in philosophy that theories are underdetermined by evidence, and that extra-scientific considerations do some of the work in determining theory choice.

So I conclude, tentatively, that the Dispensability Thesis (and *a fortiori* the Derivativity Thesis) is false. Suppose, however, that somehow it were possible to come up with an argumentative reconstruction that almost exactly captures the strength and limits I have attributed to Galileo's thought experiment.[22] Would we then be justified in accepting the deeper methodological claim put forth in the Derivativity Thesis: that the *justificatory force* of whatever beliefs I hold via thought experiment is a function of the thought experiment's argumentational essence? In the next section (**3.2**), I offer reasons for thinking that the answer to this question is 'no'.

3.2 Rejecting the positive argument: what makes these beliefs *new*?

The Derivativity Thesis says that the justificatory force of any good scientific thought experiment can only be explained by the fact that it can be replaced, without loss of demonstrative force, by a non-thought-experimental argument. So rejection of the thesis can take two forms: denying that the justificatory force of a particular scientific thought experiment *can* be explained this way, and denying that it can *only* be explained this way. The simplest way of doing the former, of course, would be to show that the Dispensability Thesis is not true of some thought experiment; obviously, if there is no non-thought-experimental argument with which the thought experiment can be replaced without loss of demonstrative force, then the justificatory force of the thought experiment cannot be explained by the possibility of such replacement. But for the sake of argument, I am supposing that the Dispensability Thesis is (at least approximately) true. This leaves two avenues for denying the Derivativity Thesis: denying that the argumentative reconstruction explains the justificatory force of some thought experiment *at all*, and denying that it explains such justificatory force *entirely*.

As I discussed above (Section **1.3**), two sorts of defence are offered for the Derivativity Thesis. The negative argument contends that thought experiments are arguments because there is nothing else for them to *be*; the positive argument contends that thought experiments are arguments because their 'analysis and appraisal' involves explicit argumentative reconstruction. In this section I will address the positive thesis, suggesting that it fails to get at what is most interesting about thought-experimental reasoning; in the final section, I offer some thoughts about what sorts of alternative justifications are available such that the negative thesis too is untenable.

It is a mistake, I contend, to think that the *reason* conclusions to thought

---

[22] Suppose, for instance, we were to enumerate the defeasibility conditions of (8) and (9), and we included these modified versions as premises alongside (1) and (2).

experiments are justified is because thought experiments have argumentational analogues (if indeed they do). Rather, I want to suggest that the Aristotelian comes to have novel justified true beliefs about the empirical world not because he has (whether he knows it or not) followed the path of a recognized argument form, but rather because he has performed an act of introspection that beings to light heretofore inarticulated and (because he lacked a theoretical framework in which to make sense of them) heretofore implausible tacit beliefs. There are two things that I need to show myself able to explain. The first is how it is that knowledge has been *gained*. In what way is it that the Aristotelian has come to believe something *new*? The second is how it is that *knowledge* has been gained. In what way is it that the Aristotelian has come to believe something *justified*? I will answer the first in the remainder of this section, and the second in Section **3.3**.

In addressing the issue of novelty, I will begin with brief remarks about what I think is *not* at issue, and then say something about where I think the important questions rest. One might say that the beliefs are not new since, in some sense, the Aristotelian had access to them before. After all, he has acquired no new knowledge of the external world; all he has done is reshuffle tacit beliefs he already held, coming to see their implications. But this view of what makes knowledge new is surely too stringent; it would, among other things, rule out all mathematical reasoning as a potential source of new knowledge. On the other hand, it seems too weak to say that a belief is new if it merely results from putting together two explicitly held beliefs that have not, for the individual in question, been previously connected. If I believe that snow is white and I believe that crows are black, but I have never thought about the two at the same time, it seems wrong to say that the belief that snow-is-white-and-crows-are-black should count as a *new* belief for me.[23] In any case, without spelling out precisely what it is that makes some beliefs new and others mere implications, there is a simple reason to think that the Aristotelian's belief that the speed at which a body falls is independent of its weight should count as a new belief for him: Until recently he was explicitly committed to the truth of its negation. This alone suggests that—whatever the implicatory relation between his prior commitments and this view about natural speed—the belief should count as new.

But there is a deeper and more interesting way that the belief is new, and that

---

[23] Except under very odd circumstances. Suppose I believe that crows are black because I live in a village where there are crows, and I have seen many of them. I also believe that snow is white, because I have read about it in books. In my village, it is taboo to think about black things and white things at the same time, because this is thought to allow the evil spirit access to the soul. One day, I leave my village for the north, and I observe a crow circling above a field of snow. I find the image aesthetically striking, and I ask myself why this is so. In analysing my response to the visual experience I am having, I realize with a start: 'Snow is white and crows are black.' In such a case, it seems plausible to suggest that this is a new belief.

is the following. The thought experiment that Galileo presents leads the Aristotelian to a reconfiguration of his conceptual commitments of a kind that lets him see familiar phenomena in a novel way. What the Galilean does is provide the Aristotelian with conceptual space for a new notion of the *kind of thing* natural speed might be: an independently ascertainable constant rather than a function of something more primitive (that is, rather than as a function of weight). It is in this way, by allowing the Aristotelian to make sense of a previously incomprehensible concept, that the thought experiment has led him to a belief that is properly taken as *new*.

What this suggests is that the Derivativity Thesis is *missing the point* of what makes the Galileo case work as it does. The recognition that natural speed is independent of weight comes not from tracing the implications of antecedent commitments to (1) and (2), which, after all, lead to the denial of a position to which the Aristotelian is explicitly committed (and thence to retreats such as the four ways out). The recognition comes from the sudden realization, on the part of the Aristotelian, of the conceptual possibility of a certain sort of physical property. Prior to contemplation of the case, there was no room on the Aristotelian picture for the thought that natural speed might be constant, not varying—that it might be dependent not on some specific features of the body in question, but only on the fact that it is a body at all.[24] After contemplation of the case, there seems to be no conceptual space for the view that it might be variable.

Tempting as it might be to digress, this is not the place for a general discussion of incommensurability across theory-change. What is important for our purposes is only this: one of the things that enables this rather striking shift in the representation of physical reality is that the Aristotelian recognizes that there are experientially possible objects—strapped-together bodies—for which the defeasibility conditions of (8) and (9) are not met (that is, objects of which (8) and (9) hold true), and that these are objects for which his old notion of natural speed simply *does not make sense*. If entification is arbitrary and natural speed and weight are fixed by the world, then a feature-dependent notion of natural speed is just plain incoherent. So one way of thinking about how the thought experiment works is this: it brings the Aristotelian to recognize the inadequacy of his conceptual framework for dealing with phenomena which—through the contemplation of this imaginary case—he comes to recognize as always having been part of his world.

---

[24] To get a sense of how odd this transition is, try thinking about *weight* as something dependent not on the specifics of the body in question, but as something constant for all bodies. (That is, to ascertain a body's weight, we would not need to know anything more about it that the simple fact that it is a body.) Clearly this would be a major conceptual readjustment; one might even be inclined to say that we aren't talking about *weight* any more, since whatever sort of thing weight is, it is surely something that depends upon specific features of the bodies to which it applies. The analogy is not perfect, since part of what happens as a result of thinking about the Galileo case is that it becomes apparent that there is no physical application for the Aristotelian idea of natural speed; like phlogiston, it disappears into the ether of abandoned concepts.

What this suggests is that 'the *analysis* and *appraisal* of a thought experiment' need not 'involve reconstructing it explicitly as an argument', in such a way that 'a good thought experiment is a good argument, a bad thought experiment is a bad argument' (Norton [1991], p. 129; [1996], p. 335). After all, the argument from (1) and (2) to (3) is no better or worse than the argument from (not-3) to (not-1) or (not-2). Like an experiment, *part* of what makes a thought experiment good or bad is the validity of the procedure by which the same result can be repeatedly obtained. But another thing that distinguishes good thought experiments from bad is their ability to direct the reader's attention to inadequacies in her conceptual scheme that she herself recognizes immediately, as soon as they are pointed out to her. It is *this,* I want to suggest, that grounds her new beliefs. Of course, I have said nothing so far about what might make these beliefs *justified*. It is to this issue that I turn in the Section **3.3**.

3.3 Rejecting the negative argument: what makes these beliefs *knowledge?*

Thought experiments work in a variety of ways. By describing appropriately selected imaginary scenarios, they provide contexts within which sense can be made of previously incomprehensible conceptual distinctions.[25] This happens when two features that are constantly conjoined in our representations of all actual cases are imaginatively separated in the thought-experimental scenario in a way that shows them to have been isolatable all along. And by describing specific situations, thought experiments, like analogical reasoning in general, can justify conclusions about particular cases without explicit or implicit appeal to more general absolute principles.[26] Many of the higher-level principles by which we negotiate the world are *defeasible*, and the determination of their applicability to particular situations must be made on a case-by-case basis. By bringing the reader to focus on particulars, thought experiments can help the reader distinguish warranted from unwarranted applications of the principle in question.[27]

So far, however, this has little to do with *justification*. After all, as Norton would ask, if the thought experiment is not an argument, why should we put faith in its conclusion? What I want to explore in this closing section is one possible answer: that thought experiments rely on a certain sort of *constructive*

---

[25] By this I mean they do something like what the answer to a riddle does in making suddenly intelligible what previously appeared to be a nonsensical description. Cf. Cavell [1979], pp. 156–7.

[26] For an articulation and defence of such a view of analogical reasoning, cf. Sunstein [1993] and [1996].

[27] This often happens when the particulars are sufficiently well-sketched to evoke practical as well as theoretical responses. Cf. also the literature on mental modelling, references to which can be found in Nersessian [1993], as well as at Thomas [undated].

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*.

An experiment-in-thought is an *actual* experiment; the person conducting the experiment asks herself: 'What would I say/judge/expect were I to encounter circumstances XYZ?' and then *finds out* the (apparent) answer. This technique is common in linguistics, where the methodology is used to ascertain the grammaticality of sentences, the meanings of phrases, the taxonomic categories of words, and so on.[28] And it is, on one view at least, a central element of moral reasoning: we think about particular imaginary cases, observe the judgements that they evoke in us, and use these judgements as fixed points in developing our moral theories.[29]

How does this connect with the Galileo case? What kind of experiment-in-thought plays a role there? Answer: by thinking about the case in question, we discover what sorts of motions and objects we think are possible in the world. Do we think objects can be strapped together? Yes, we do. Do we think objects fall with radical discontinuities in speed? No, we think they do not. Do we think entification is something that is fixed by the world? No, we do not. Do we think weight and natural speed are fixed by the world? Yes, we do. We *come to recognize* that we have these beliefs by contemplating the imaginary case in question; thinking about the case is what brings us to the realization that we believe what we do. *And*—and this is where the justificatory work comes in— the fact that we have these beliefs gives us *prima facie* warrant to think that they are true.

But why? Why should we think that our pre-theoretical beliefs about the structure of the physical world are reliable? In the *Science of Mechanics*, a few pages after coining the expression *Gedankenexperiment*, Mach writes:

> Everything which we observe imprints itself *uncomprehended* and *unanalyzed* in our percepts and ideas, which then, in their turn, mimic the process of nature in their most general and most striking features. In these accumulated experiences we posses a treasure-store which is ever close at hand, and of which only the smallest portion is embodied in clear articulate thought. The circumstance that it is far easier to resort to these experiences than it is to nature herself, and that they are, notwithstanding this, free, in the sense indicated, from all subjectivity, invests them with high value (Mach [1883/1960], p. 36).

---

[28] Cf. Thomason [1991], p. 247: 'When linguists want to test hypotheses about the structure of a particular language, their methodology crucially involves thought experiments in a ... literal sense: real experiments carried out in thinking.'

[29] Cf., to choose an example nearly at random, Thomson [1986], p. 257: 'it is ... our moral views about examples, stories, and cases which constitute ... data for moral theorizing.' I do not intend to be taking a stand here on questions of moral epistemology.

So one possible explanation is the one Mach gives. We have stores of unarticulated knowledge of the world which is not organized under any theoretical framework. Argument will not give us access to that knowledge, because the knowledge is not propositionally available. Framed properly, however, a thought experiment can tap into it, and—much like an ordinary experiment—allow us to make use of information about the world which was, in some sense, there all along, if only we had known how to systematize it into patterns of which we are able to make sense.[30]

This, of course, is the beginning not the end of an answer to the question. But it is sufficient for the modest aim of this article. What I have been trying to show is that *something* besides argument might give justificatory force to thought-experimental reasoning. The alternative I have proposed has been this: by focusing on imaginary scenarios and making reference to particulars, thought experiments can provide a fulcrum for the reorganization of conceptual commitments; this explains the way in which they can provide us with novel *information* without empirical input. And by bringing the reader to perform experiments in thought, thought experiments can lead us to reject shaky (and ultimately false) theoretical commitments in light of newly systematized but previously inarticulable *knowledge* about the way the world is.

The justificatory force of thought experiments is thus parasitic on the extent to which the messy twisted web of background beliefs that underpin our navigation of the world are rightly considered knowledge. To establish this, on coherentist or evolutionary or empiricist grounds, would be an enormous undertaking, and one which I will not even begin to endeavour here.[31] But I hope I have given you some sense, at least, of why I am not convinced that even if the Dispensability Thesis is true, the Derivativity Thesis must be true as well. For even if it *could* be replaced by an equally effective argument, the *justificatory force* of a thought experiment might still be based on its capacity to make available in a theoretical way those tacit practical commitments which enable us to negotiate the physical world.

3.4 Constructivism and the contrast with Norton and Brown

Before concluding, let me try to articulate more clearly how the position I have been advocating differs from those of James Robert Brown and John Norton, both of whom offer detailed but conflicting discussions of precisely this thought experiment.[32] Brown's discussions of the case can be found

---

[30] For further discussion of these themes, cf. Kuhn [1964].
[31] For one such (evolutionarily based) attempt, cf. Shepard [1994]; a similar explanation is offered in Sorensen [1992a].
[32] I thank a non-anonymous referee for *BJPS*—John Norton, in fact—for suggesting that I include such a section.

throughout his works,[33] but I will focus on his most explicit version, which appears in his book (Brown [1991a]), and which is framed partly as a critique of Norton. Norton's own discussion of the case, which is framed partly as a critique of Brown, can be found in his 1996 article (Norton [1996]).

Brown contends that the Galileo thought experiment is what he calls a *platonic thought experiment*, that is 'a single thought experiment which destroys an old or existing theory and simultaneously generates a new one' (Brown [1991a], p. 77, cf. pp. 43–5). According to Brown, thought experiments such as Galileo's give us '*a priori* knowledge of nature' (Brown [1991a], p. 77); through consideration of such thought experiments, Brown maintains, we gain direct quasi-perceptual knowledge of relations between independently existing abstract entities (Brown [1991a], p. 76). In the case of Galileo's thought experiment, the platonic law of nature to which we gain access is that 'all bodies fall at the same rate'.[34]

Bracketing questions concerning the ontological status of laws of nature, Brown's basic contention is that Galileo's thought experiment brings the heretofore mistaken Aristotelian to see the truth of a certain platonic law of nature, and that it does so neither (a) by introducing the Aristotelian to novel empirical data, nor (b) by showing the law to be a logical truth, nor (c) by making minimal modifications to the Aristotelian's earlier theory, nor (d) by deriving that law from antecedent commitments which the Aristotelian is in a position to maintain[35]. Of these, (a) and (b) are apparently uncontroversial,[36] and Brown argues for (c)—whose truth I also accept for reasons enumerated in Section **2.3** above—by contending that the post-thought-experimental justification for belief in the Galilean theory is higher than pre-thought-experimental justification for belief in its Aristotelian counterpart. The crucial claim, then, is (d): that the Aristotelian cannot be brought to the Galilean conclusion by argument alone.

It is precisely this, of course, which Norton denies. So before returning to Brown's own positive account of the thought experiment, let us consider Norton's proposed argumentative reconstruction of the case.[37] Norton suggests that Galileo's reasoning is as follows (Norton 1996, pp. 341–2). (The comments in brackets that follow each of Norton's steps are mine.)

---

[33] Cf. Brown [1991a, b, 1993b, 1995]. In [1991a] he speaks of this as 'the best (i.e. my favourite) [thought experiment]' (p. 1).

[34] For reasons that Galileo himself addresses, this law is true only of bodies falling in a vacuum; cf. Galilei [1638/1989], p. 109, which is discussed above.

[35] Cf. (1), (2), (3) in Brown [1991a], pp. 77–9.

[36] Concerning (a), it is clear that the thought experiment involves no new sensory input, nor is the Aristotelian provided with novel data concerning the behavior of objects in the world. Concerning (b), it is not logically impossible (*simpliciter*) that the speed at which bodies fall should depend on some non-universal feature(s).

[37] Norton writes: 'there can be little question that what Galileo gives us here is simply an argument' (Norton [1996], p. 341).

(N1) Assumption [for *reductio* proof]: the speed of fall of bodies in a given medium is proportionate to their weights.

[This is equivalent to the negation of (3) in my reconstruction in Section **2.1**; the brackets around 'for *reductio* proof' are my own .]

(N2) From (N1): if a large stone falls with 8 degrees of speed, a smaller stone half its weight will fall with 4 degrees of speed.

[This is an application of the (N1)-law to a particular case. On my understanding of the Elimination Thesis (cf. Section **1.2** above), invocation of this particular makes Norton's reconstruction—strictly speaking—a thought-experimental argument. But since his reconstruction relies essentially only on the incompatibility of (N1), (N3), and (N5), I will neglect this *ad hominem* point. The reader should be careful, however, not to focus on (N2), (N4), and (N6) in thinking about the forcefulness of Norton's reconstruction.]

(N3) Assumption: if a slower falling stone is connected to a faster falling stone, the slower will retard the faster and the faster speed the slower.

[This is equivalent to (1)—that natural speed is mediative—in my reconstruction in Section **2.1**.]

(N4) From N3: if the two stones of (N2) are connected, their composite will fall slower than 8 degrees of speed.

[This is an application of the (N3)-law to the example presented in (N2).]

(N5) Assumption: the composite of the two weights has greater weight than the larger.

[This is equivalent to (2)—that weight is additive—in my reconstruction in Section **2.1**.]

(N6) From (N1) and (N5): The composite will fall faster than 8 degrees.

[This is an application of the (N1)-law to the example presented in (N2) as governed by the (N5)-law.]

(N7) Conclusions (N4) and (N6) contradict.

[As the reconstruction in Section **2.1** demonstrates, a parallel result can be obtained from (N1), (N3), and (N5) alone, without reference to the particular applications that generate (N4) and (N6).]

(N8) Therefore, we must reject assumption (N1).

[Since (N7) is derived from (N1), (N3) and (N5), then one of these three must be rejected. Norton takes (N3)—that natural speed is mediative—and (N5)— that weight is additive—as given, so the only option is to reject (N1). As my

discussion in Section **2.3** shows, however, there are ways of denying (N3) and (N5) which are—at least *prima facie*—not completely outrageous.]

Norton then goes on to discuss a final step, which he attributes to Brown:

(N9) Therefore all stones fall alike.

[Norton attributes this conclusion to Brown in part on the assumption that Brown accepts that (N8) follows straightforwardly from (N1)–(N7). For reasons that I discuss below, I think this is a misreading of Brown's view. But in any case, as both Norton and Brown recognize, (N9) does not follow from (N8) without additional premises. All that (N8) tells us is that the speed of fall of bodies in a given medium is not proportionate to their weights (that is, the negation of (N1)); it has no bearing on the issue of whether it might depend on some other factor.[38] Galileo himself is trying to show that 'if one were to remove entirely the resistance of the medium, all materials would descend with equal speed' (Galilei [1638/1989], p. 116), but since precisely what is at issue in the section as a whole is the possibility and nature of motion in a void, this idealization is not permitted as one of the assumptions in the reconstruction.]

Clearly, the move from (N8) to (N9) depends upon an auxiliary assumption:

(N8a) Assumption: the speed of fall of bodies depends only on their weights.

If (N8a) is accepted, then (N9) follows from (N3) and (N5),[39] and if (N8a) is not accepted, then (N9) does not. But, Norton notes, Galileo's characters are not in a position to accept (N8a), nor, if we are considering the fall of bodies *in media*, are we. For the speed at which a body falls *in media* is dependent in part on its *shape*: gold leaf falls more slowly than a gold nugget of the same weight (cf. Galilei [1638/1989], p. 109), just as a closed parachute falls more quickly than an open one. So if we wish to derive (N9) from (N8), we will need to appeal to (N8a).

In short, Norton concludes: 'our degree of belief in the final outcome of the thought experiment turns out to depend on our degree of belief in the assumption' (N8a). How so? Well, 'the transition from (N8) to (N9) is, at worst, a fallacious inference to a falsehood [if (N8a) is denied]; or, at best, valid only insofar as it is invoked in special cases in which assumption [(N8a)] holds, such as the fall of very heavy compact objects in very rare media. This final step looks more like a clumsy fudge or a stumble than a leap into the Platonic world of laws' (Norton 1996, p. 345). And this outcome, Norton contends, supports the previously-cited thesis that 'a good thought experiment is a good argument, a

---

[38] Cf. Brown: it is not a logical contradiction to suppose that 'bodies might fall at different speeds depending on their colours or on their chemical composition' (Brown [1991a], p. 78).

[39] Norton presents this argument as a second *reductio*; cf. Norton [1996], p. 343.

bad thought experiment is a bad argument' (Norton [1991], p. 131; [1996], p. 335). For the strength of the thought-experimental conclusion correlates directly with the strength of the corresponding argument.

Now, I think the discussion of (N8a) turns out to be a bit of a red herring in the dispute between Brown and Norton because, as far as I can tell, Brown's concession that 'the thought experiment is an argument insofar as it destroys the Aristotelian theory' (Norton [1996], p. 342; cf. Brown [1991a], p. 78) is not—or at least should not be—a concession by Brown that (N8) follows logically from (N1)–(N7). Like any *reductio*, all that the argument shows is that at least one of the premises needs to be rejected; it does not show us *which* premise. (The point of my discussion in Section **2.3** was to show precisely this.) So Brown need not concede—as Norton takes him to—that (N8) follows from (N1) (brackets omitted)–(N7). But, of course, Brown is committed to the view that (N8) is true; indeed, Brown is committed to the view that (N9) is true. So how does Brown think that we know (N8) and (N9), if not by argument from (N1)–(N7) and (N8a)?

Brown's answer, in so far as I understand it, is that thinking about the case forces us (or the Aristotelian) to use our platonic perceptual capacities to see that (N9) is a platonic law of nature (that is, a connection between independently existing abstract entities). But since it seems to me that a perfectly good explanation of what is going on can be offered without appeal to such laws, ontological parsimony suggests that Norton may be partly wrong without Brown being fully right. So let me say what I think is right and wrong in both views as a way of putting my own in context. To give my view a label that contrasts it with Norton's self-proclaimed empiricism and Brown's self-proclaimed platonism, let me call my own position *constructivism* (Kantian associations intended).

I agree with (my reading of) Brown that (N1)–(N7) does not get us (N8); the *reductio* tells us that *something* is wrong with the Aristotelian picture, but it does not tell us *what* is wrong. And I agree with Norton that something is being left out of Brown's story by his neglect of (N8a); indeed, I don't think that the thought experiment in question shows anything more than that natural speed is independent of weight. But what my argument in Section **2.4** aims to show is that the sorts of considerations that allow the conclusion of (N8) from (N1)–(N7) are not available to the Aristotelian as premises before she has followed Galileo's instructions for guided contemplation. It is by focusing on some imaginary particular strapped-body that she comes to realize her non-defeated defeasible commitments to my own (8) (the physical determination of natural speed and weight) and (9) (the lack of physical determination of entification). And therein lies my answer to both Norton and Brown.

Now, if Norton wants to call this sort of guided contemplation *argument*, and Brown wants to call it *platonic intuition*, then we're all in agreement. I

prefer to describe it as a reconfiguration of internal conceptual space. By introducing novel categories by which we make sense of the world, this reconfiguration allows us to recognize the significance of certain previously unsystematized beliefs. For the reasons that I discussed in Section **3.2**, these beliefs are new; and for the reasons I discussed in Section **3.3**, they are justified. Thus the thought experiment brings us to new knowledge of the world, and it does so by means of non-argumentative, non-platonic, guided contemplation of a particular scenario.

4 Conclusion

In this article, I have offered reasons for thinking that a certain view about thought experiments in science is false. The view is that any scientific thought experiment can be reconstructed as a non-thought-experimental argument without loss of demonstrative force.[40] In Section **1**, I explained the philosophical motivations for adopting such a view, and distinguished two versions of the position. The first—the Dispensability Thesis—concerns the replaceability of thought experiments; the second—the Derivativity Thesis—concerns their justificatory force. In the remainder of the article, I offered reasons for thinking that both of these theses are false. Through a detailed discussion of a thought experiment of Galileo's, I tried to show that the standard argumentative reconstruction of the case fails to capture its justificatory power, and I suggested reasons to think that any other argumentative reconstruction would be likely to fail in similar ways. I then argued that even if one were to provide an argumentative reconstruction that did almost perfectly capture the thought experiment's demonstrative force, this would not show that the *reason* the thought experiment is successful is because, deep down, it is nothing more than an argument in disguise. I suggested that, to the contrary, the success of the thought experiment may be a result of the way in which it invites the reader's constructive participation, depicts particulars in ways that make manifest practical knowledge, and describes an imaginary scenario wherein relevant features can be separated from those that are inessential to the question at issue.

Acknowledgements

---

[40] In Section **3.4** I also briefly offered reasons for rejecting another alternative—that thought experiments offer us special access to platonic laws of nature.

Department of Philosophy
Syracuse University
Syracuse, NY 13244-1170
USA

References

Bealer, G. [1998]: 'Intuition and the Autonomy of Philosophy', in M. DePaul and W. Ramsey (*eds*), *Rethinking Intuition*, Savage, MD, Rowman and Littlefield Publishers (in press).

Brown, J. R. [1986]: 'Thought Experiments Since the Scientific Revolution', *International Studies in the Philosophy of Science*, **1**, pp. 1–14.

Brown, J. R. [1991a]: *The Laboratory of the Mind: Thought Experiments in the Natural Sciences*, New York and London, Routledge.

Brown, J. R. [1991b]: 'Thought Experiments: A Platonic Account', in T. Horowitz and G. Massey [1991], pp. 119–28.

Brown, J. R. [1993a]: 'Author's Response (to Norton [1993])', *Metascience* (new series), **3**, pp. 38–40.

Brown, J. R. [1993b]: 'Why Empiricism Won't Work', in D. Hull *et al.* [1993], pp. 271–9.

Brown, J. R. [1995]: 'Critical Notice: Roy Sorensen's *Thought Experiments*', *Canadian Journal of Philosophy*, **25**, pp. 135–42.

Butts, R. E. and Pitt, J. C. (*eds*) [1978]: *New Perspectives on Galileo*, Dordrecht and Boston, Reidel.

Cargile, J. [1987]: 'Definitions and Counterexamples', *Philosophy*, **62**, pp. 179–93.

Carrier, M. [1993]: 'Review of Horowitz and Massey (*eds*), *Thought Experiments in Science and Philosophy*', *Erkenntnis*, **39**, pp. 413–19.

Cavell, S. [1979]: *The Claim of Reason: Wittgenstein, Skepticism, Morality and Tragedy*, New York and Oxford, Clarendon Press.

Claggett, M. [1959]: *The Science of Mechanics in the Middle Ages*, Madison, University of Wisconsin Press.

Clavelin, M. [1974]: *The Natural Philosophy of Galileo: Essays on the Origins and Formation of Classical Mechanics*, Cambridge, MA, MIT Press, trans. A. J. Pomerans.

Clement, J. [1983]: 'A Conceptual Model Discussed by Galileu and Used Intuitively by Physics Students', in D. Gentner and A. L. Stevens (*eds*), *Mental Models*, Hillsdale, NJ, Lawrence Erlbaum Associates, pp. 325–40.

Cooper, L. [1935]: *Aristotle, Galileo, and the Tower of Pisa*, Ithaca, NY, Cornell University Press.

Damerow, P. *et al.* [1992]: *Exploring the Limits of Pre-classical Mechanics: A Study of Conceptual Development in Early Modern Science: Free Fall and Compounded*

Motion in the Work of Descartes, Galileo and Beeckman, New York, Springer-Verlag.

Dijksterhuis, E. J. [1961, repr. 1986]: *The Mechanization of the World Picture: Pythagoras to Newton*, Princeton, NJ, Princeton University Press, trans. C. Dikshoorn.

Drake, S. [1978]: *Galileo at Work: His Scientific Biography*, Chicago. IL, University of Chicago Press.

Drake, S. [1989]: *History of Free Fall: Aristotle to Galileo*, Toronto, Wall and Thompson.

Drake, S. [1990]: *Galileo: Pioneer Scientist*, Toronto, University of Toronto Press.

Drake S. and Drabkin, I. E. [1969]: *Mechanics in Sixteenth Century Italy: Selections from Tartaglia, Benedetti, Guido Ubaldo, Galileo*, Madison, WI, University of Wisconsin Press.

Galilei, G. [1590/1960]: *On Motion*, Madison, WI, University of Wisconsin Press, trans. I. E. Drabkin.

Galilei, G. [1638/1914, repr. 1954]: *Dialogues Concerning Two New Sciences*, New York, Dover, trans. H. Crew and A. de Salvio.

Galilei, G. [1638/1974, rev. 1989]: *Two New Sciences, Including Centers of Gravity and Force of Gravity and Force of Percussion*, Toronto, Wall and Thompson, trans. S. Drake.

Giere, R. (*ed.*) [1992]: *Minnesota Studies in the Philosophy of Science: Cognitive Models of Science*, XV, Minneapolis, MN, University of Minnesota Press.

Gooding, D. C. [1990]: *Experiment and the Making of Meaning*, Dordrecht, Kluwer.

Gooding, D. C. [1992]: 'The Procedural Turn; or Why do Thought Experiments Work?', in R. Giere [1992], pp. 45–76.

Gooding, D. C. [1993]: 'What is *Experimental* about Thought Experiments?' in D. Hull *et al.* [1993], pp. 280–90.

Gooding, D. C. [1994]: 'Imaginary Science' (review of Horowitz and Massey [1991], Sorensen [1992a], Brown [1991a]), *British Journal for the Philosophy of Science*, 45, pp. 1029–45.

Hacking, I. [1993]: 'Do Thought Experiments Have a Life of Their Own?' in D. Hull *et al.* [1993], pp. 302–8.

Hesse, M. [1966]: *Models and Analogies in Science*, Notre Dame, IN, University of Notre Dame Press.

Horowitz, T. and Massey, G. (*eds*) [1991]: *Thought Experiments in Science and Philosophy*, Savage, MD, Rowman and Littlefield Publishers.

Hull, D., Forbes, M. and Ohrulik, K. (*eds*) [1993]: *Proceedings of the Philosophy of Science Association, 2*, East Lansing, MI, Philosophy of Science Association.

Humphreys, P. [1993]: 'Seven Theses on Thought Experiments', in J. Earman, A. Janis, and G. Massey (*eds*). *Philosophical Problems of the Internal and External Worlds: Essays on the Philosophy of Adolf Grünbaum*, Pittsburgh, PA, University of Pittsburgh Press, pp. 205–27.

Irvine, A. D. [1991]: 'On the Nature of Thought Experiments in Scientific Reasoning', in T. Horowitz and G. Massey [1991], pp. 149–65.

Janis, A. I. [1991]: 'Can Thought Experiments Fail?' in T. Horowitz and G. Massey [1991], pp. 113–18.

Koyré, A. [1939/1979]: *Galileo Studies*, Atlantic Highlands, Humanities Press, trans. J. Mepham.

Koyré, A. [1960]: 'Galileo's Treatise *De Motu Gravium*: the Use and Misuse of an Imaginary Experiment', in *Metaphysics and Measurement*, London, Chapman and Hall [1968], pp. 44–88.

Kuhn, T. [1964]: 'A Function for Thought Experiments', reprinted with new pagination in *The Essential Tension*, Chicago, IL, University of Chicago Press [1977], pp. 240–65.

Kujundzic, N. [1993], 'How Does the Laboratory of the Mind Work? Critical notice of James Robert Brown, *The Laboratory of the Mind*', *Dialogue, Canadian Review of Philosophy/Revue Canadienne de Philosophie*, **32**, pp. 573–7.

Laymon, R. [1991]: 'Thought Experiments of Stevin, Mach and Gouy: Thought Experiments as Ideal Limits and as Semantic Domains', in T. Horowitz and G. Massey [1991], pp. 167–92.

Lipton, P. [1993] 'Review of Horowitz and Massey (*eds*), *Thought Experiments in Science and Philosophy*', *Ratio* (new series), **6**, pp. 82–6.

Mach, E. [1883/1960]: *The Science of Mechanics*, sixth edition, La Salle, IL, Open Court Publishers trans. T. McCormack.

Mach, E. [1897/1976]: 'On Thought Experiments', in *Knowledge and Error*, sixth edition, Dordrecht, Reidel, trans. T. McCormack and P. Foulkes, pp. 134–47.

Massey, G. [1991]: '*Gedankenexperimente*: Where Science and Philosophy Meet', unpublished manuscript.

Nersessian, N. [1984]: *Faraday to Einstein: Constructing Meaning in Scientific Theories*, Dordrecht and Boston, Kluwer.

Nersessian, N. [1992]: 'How Do Scientists Think? Capturing the Dynamics of Conceptual Change in Science', in Ronald Giere [1992], pp. 3–44.

Nersessian, N. [1993]: 'In the Theoretician's Laboratory: Thought Experiment as Mental Modeling', in D. Hull *et al.* [1993], pp. 291–301.

Norton, J. [1991]: 'Thought Experiments in Einstein's Work', in T. Horowitz and G. Massey [1991], pp. 129–48.

Norton, J. [1993]: 'Seeing the Laws of Nature: Review of James Robert Brown, *The Laboratory of the Mind*', *Metascience* (new series), **3**, pp. 33–8.

Norton, J. [1996]: 'Are Thought Experiments Just What You Thought?' *Canadian Journal of Philosophy*, **26**, pp. 333–66.

Popper, K. R. [1959/1992]: *The Logic of Scientific Discovery*, New York, Routledge, trans. K. R. Popper.

Prudovsky, G. [1989]: 'The Confirmation of the Superposition Principle: On the Role of a Constructive Thought Experiment in Galileo's *Discorsi*', *Studies in History and Philosophy of Science*, **20**, pp. 453–68.

Segre, M. [1989]: 'Galileo, Viviani, and the Tower of Pisa', *Studies in History and Philosophy of Science*, **20**, pp. 435–51.

Shepard, R. [1994]: 'Mind and World: From Newton, Einstein, and Darwin to Principles of Mind', unpublished William James Lectures, Harvard University.

Sorensen, R. [1992a]: *Thought Experiments*, New York and Oxford, Oxford University Press.

Sorensen, R. [1992b]: 'Thought Experiments and the Epistemology of Laws', *Canadian Journal of Philosophy*, **22**, pp. 15–44.

Souder, L. [undated]: 'A Home Page for Thought Experiments', http://astro.ocis.temple.edu/~souder/thought/index.html.

Sunstein, C. [1993]: 'On Analogical Reasoning', *Harvard Law Review*, **106**, pp. 741–91.

Sunstein, C. [1996]: 'Political Conflict and Legal Agreement', *Tanner Lectures on Human Values*, *18*, Salt Lake City, UT, University of Utah Press.

Thomas, N. J. T. [undated]. 'Imagination, Mental Imagery, Consciousness, and Cognition: Scientific, Philosophical and Historical Approaches', http://web.calstatela.edu/faculty/nthomas/index.htm.

Thomason, S. [1991]: 'Thought Experiments in Linguistics', in T. Horowitz and G. Massey [1991], pp. 247–57.

Thomson, J. J. [1986]: *Rights, Restitution and Risk*, Cambridge, MA, Harvard University Press.

Wilkes, K. V. [1988]: *Real People: Personal Identity without Thought Experiments*, Oxford, Clarendon Press.