Figure 3   Falling Stone with Rotating Earth
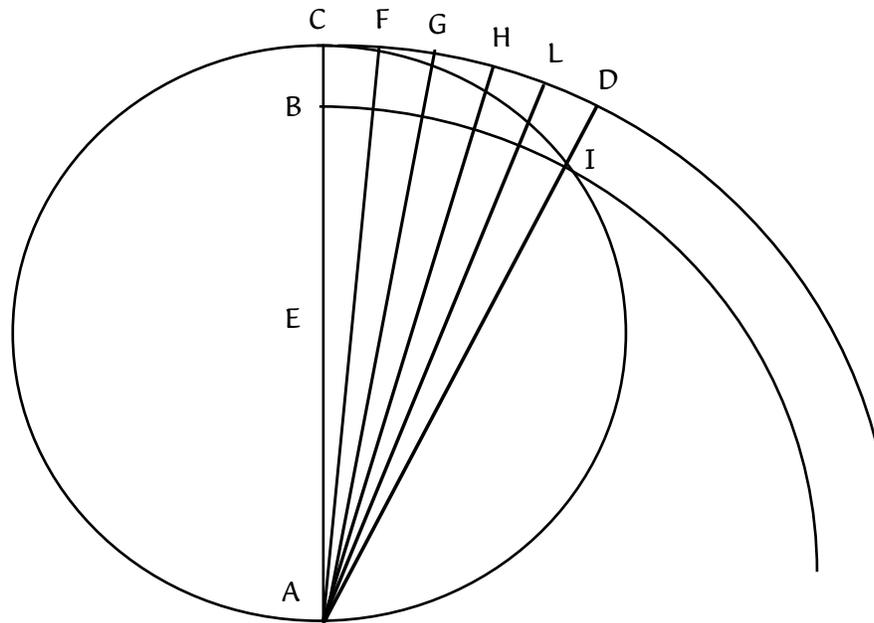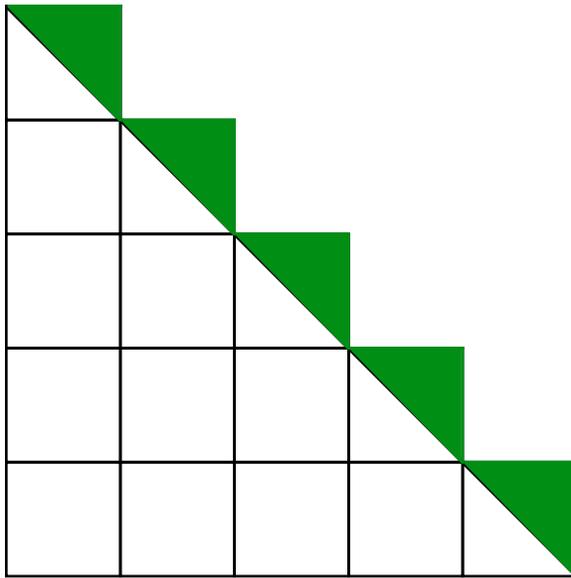
Figure 2   The sum of the first n numbers (n = 5)

Figures

Figure 1   Saunderson's calculating devices

Gendler, T. S. (1998) "Galileo and the Indispensability of Scientific Thought Experiment,"
British Journal for the Philosophy of Science, **49**, 397-424.

Koyré, A. (1978) *Galileo Studies*, transl. J. Mepham, Atlantic Highlands: Humanities
Press.

Kuhn, T. (1977) "A Function for Thought Experiments," pp. 240-265 in *The Essential
Tension*, Chicago: University of Chicago Press.

Leibniz, G. (1976) *Philosophical Papers and Letters*, 2nd edn., ed. Loemker, L.,
Dordrecht/Boston: D. Reidel.

Norton, J. (1993) "Seeing the Laws of Nature: Review of James Robert Brown (1991)",
Metascience (new series), **3**, 33-38.

Norton, J. (1996) "Are Thought Experiments Just What You Thought?" *Canadian Journal
of Philosophy*, **26**, 333-366.

Bibliography

Adam, C. and Tannery, P. (1908) *Oeuvres de Descartes,* vol X, Paris: L. Cerf.

Brown, J. R. (1986) "Thought Experiments Since the Scientific Revolution", *International Studies in the Philosophy of Science*, **1**, 1-14.

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

Brown, J. R. (1993) "Author's Response (to Norton 1993)", *Metascience* (new series), **3**, 38-40.

Brown, J. R. (1997) "Proofs and Pictures", *British Journal for the Philosophy of Science*, **48,** 161-180.

Brown, J. R. (forthcoming) *Philosophy of Mathematics*, London: Routledge.

Diderot, D. (1749) *Lettre sur les aveugles, à l'usage de ceux qui voient*, London; 73-146 in (Diderot 1964).

Diderot, D. (1964) *Œuvres philosophiques*, ed. P. Vernière, Paris: Garnier.

Feyerabend, P. (1975) *Against Method*, London: New Left Books.

Fisher, A. (1988) *The Logic of Real Arguments*, Cambridge: Cambridge University Press.

Galilei, G. (1638) *Discorsi e Dimostrazioni Matematiche, intorno à due nuoue scienze*, Leyden: Elzevir. All references given with the pagination from volume VIII of the Edizione Nazionale of Galileo's *Opere* edited by A. Favaro, Florence 1890-1910.

Galilei, G. (1967) *Dialogue Concerning the Two Chief World System, Ptolemaic and Copernican*, 2nd edn., transl. with revised notes by Stillman Drake, Berkeley: University of Caifornia Press.

Galilei, G. (1974) *Two New Sciences*, transl. and ed. of Galilei (1638) by Stillman Drake, Madison: University of Wisconsin Press.

[7] "Galileo's account of free fall did two distinct things: it destroyed Aristotle's view that heavier objects fall faster; and second, it established a new account that all objects fall at the same speed" (Brown 1991, 43).

[8] "Contemplation of the case Galileo describes *brings him to see* that these principles are not defeated in *this* case... No austere reconstruction will be able to do this, because part of the thought experiment's function is to bring the Aristotelian to accept certain *premises*" (Gendler 1989, 408).

[9] Gendler believes "that the thought experiment in question shows [nothing] more than that natural speed is independent of weight", but that the sorts of consideration that allow the conclusion of (A8) from the reductio argument (A1)-(A7) "are not available to the Aritsotelian as premises before she has followed Galileo's instructions for guided contemplation" (Gendler 1998, 419).

[10] Norton's assumption 5 is "the composite of the two weights has greater weight than the larger" (Norton 1996, 342); Gendler's corresponding second assumption is the more austere "weight is additive"(Gendler 1998, 404).

[11] As Norton points out (1996, 344), Galileo allows that gold's being beaten into a leaf or a rock's being crushed into powder will make it behave differently in falling through a medium (Galilei 1638, 109). Brown refers to the reanalysis of Eötvos' experiment performed by Fischbach in 1986, which appeared to show the existence of a "fifth force" acting against gravity, and which depended on atomic composition (Brown 1991, 78, 167).

On thought experiments as *a priori* science   ©R. T. W. Arthur

Notes

[1]"Thought experiments are arguments which (i) posit hypothetical or counterfactual states of affairs, and (ii) invoke particulars irrelevant to the generality of the conclusion" (Norton 1996, 336).

[2] (Diderot 1964, 120-121). See Newton's Letters to Bentley, and Clarke's Boyle lectures. One of Leibniz's favourite images is that of the city looked at from the perspective of various approaches as compared to the view from above, which he claims is analogous to the points of view of created substances as compared to that of God. See his Letter to Thomasius (1669), Discourse on Metaphysics (1686) §14, Monadology (1714) §57, in (Leibniz 1976, 97, 312, 648, resp.).

[3] Diderot is rightly fascinated by these calculating devices invented by Saunderson, an illustration of which is given in figure 1.

[4] Descartes' solution dates from late 1618, see (Adam and Tannery 1908, 76), and Koyré's discussion in Part II of his classic work (Koyré 1978, 79-86).

[5] Fisher leaves out the "natural speed" premise, my (A2); this turns out to be crucial in my analysis. Fisher has (A3) follow from (A1), whereas I take it to follow from (A2), a premise accepted by Galileo.

[6] Following Galileo, I shall treat "being of the same material" as equivalent to "having the same specific gravity" for the purposes of this argument.

Galileo's case, how the confusion in the original experiment can be shown up using formal logic, once the thought experiment has done its clarificatory work. In this sense, in the context of reformulation of knowledge, I think Norton's reconstruction thesis is valid: the reformulation of thought experiments as arguments is a vital part of the scientific process. Nevertheless, exactly why it was that the Aristotelian argument begged the question could still only be seen once it was understood that a rival natural interpretation could be used to describe the same phenomena.

In conclusion, then, I do not think thought experiments are simply reducible to arguments without epistemic loss. I agree with Tamar Gendler that the formulated argument may still not have the requisite persuasive force, precisely because what is at issue is what can be presupposed, and this alters the acceptability-status of the premises. Thought experiments go beyond arguments in providing an easily visualizable—or rather, I should say in deference to Saunderson, graspable— imaginative reconstruction of the phenomenon at issue. This playing through of the phenomenon or situation not only helps to bring out its essence—as in the proof-sketch of the sum of the first n integers $1 + 2 + 3 + ... + n = (n^2 + 1)/2$. But, at least in the case of those thought experiments leading to conceptual innovation, it helps us to articulate a new "natural interpretation" that is self-consistent. It is in this sense that thought experiments lead to new knowledge *a priori*.

On thought experiments as *a priori* science   ©R. T. W. Arthur

I want to end with a link back to the discussion of the first section, by making a comparison between dream interpretation and the interpretation of thought experiments. For there is, I believe, no univocal interpretation of a given dream. What appears to be going on, at least as we emerge into the waking state, is that ideas following one another in a certain sequence also evoke other tenuously connected ideas laterally, with a resulting confusion as to which ideas are supposed to have a real sequential connection. The interpretive vehicle that's normally brought to bear in making sense of this in conscious thought is, as it were, working through a cloud. Still there's an analogy with thought experiments in that "working through a cloud", I suggest, is what we're always doing to a greater or lesser extent. Perhaps I'm a particularly woolly thinker, and this is only true of me, but I'm always struggling for greater clarity, and thought experiments seem to me to be an example, just as formal arguments themselves are, of our attempts to achieve a clarity of the concepts that we're using.

On this way of conceiving things, a thought experiment is an attempt to gain clarity, to articulate the concepts we have formed of the world. Now I don't think there are pre-existent concepts in a really precise sense, just as I do not think there are arguments until we elaborate them and make them explicit. We have some sort of inkling of them, some intuition of them, but that isn't because those ideas are really there and we're not quite understanding them, it's because we haven't quite succeeded in formulating them. Likwise, we do not generally reason with formal logic; we are more likely to formulate our reasoning using explicit premises and rules of inference when we're struggling to gain clarity, or perhaps when we've achieved it and are trying to persuade others. We use it as a culmination of a process of clarification in our concept formulation and attempt to persuade.

The examples of the Aristotelian thought experiment demonstrating the impossibility of the Earth's motion, contradicted by the Galilean experiment showing its consistency with the observed facts, shows not only this clarification, but also, in

contradiction, namely that weight is nonoperative for bodies falling at their natural speeds in the same medium. But from this it follows that —so far as other factors like shape and the effects of the so-called fifth force may be ignored[11]— a body moving in the same medium at its natural speed will be neither accelerated nor retarded by the addition of another body of the same specific gravity.

Here new knowledge has been achieved, an old theory has been undermined and a new one simultaneously generated to replace it, but without the addition of any new empirical data. Thus it meets all the requirements that Brown proposes for a "Platonic thought experiment", but without his problematic notion of Platonic perception. Moreover, it is *a priori* in precisely the sense Brown articulates: "it is not based on new empirical evidence nor is it merely logically derived from old data" (Brown 1991, 77). I hasten to add, as does Brown, that this does not mean the resulting theory is infallible, or has no empirical content. That knowledge reached by such *a prori* methods still has empirical content, and the example of Ptolemy's thought experiment being undermined by Galileo's subsequent one should serve to show that the process of unearthing tacit assumptions or natural interpretations can never be assumed to be complete.

I should also re-emphasize the limited scope of this interpretation: it is intended to apply only to those thought experiments which refute an existing theory by *reductio ad absurdum*, and simultaneously suggest a constructive interpretation free of contradiction. But it would therefore seem to have application to the majority of cases that Brown calls Platonic thought experiments: Einstein and Infeld's idealized experiment with an elevator to argue for the equivalence of accelerated motion and motion in a gravitational field (Brown 1991, 17-20), Leibniz's thought experiment to establish that the quantity "motive force" that is conserved is mv^2 (Brown 1991, 43-45), and so forth.

V *Conclusion*

natural speed. This is the deeper reason why natural speed cannot be proportional to weight. "From this we conclude that both great and small bodies, of the same specific gravity, are moved with like speeds" (Galileo 1638, 109).

Here my reading of the thought experiment differs significantly from those of Norton and Gendler. For they both have the additivity of weight as an independent premise in the original *reductio*, a premise all the interlocutors must agree to in order for the contradiction to follow, and for the proportionality of natural speed to weight to be refuted.[10] On my reading, by contrast, Galileo only says that the composite of the two stones of the same specific gravity is larger than the bigger stone; one stone's being heavier than another when both have attained their natural speed only follows if weight is operative in these circumstances. Thus we may now construct the argument:

(B1) Every falling body acquires a natural speed, i.e. a "definite speed fixed by nature, a speed which cannot be increased or diminished except by the use of force or some impediment that retards it" (= A2).

(B2) But weight is not operative when bodies have attained their natural speed.

(B3) Two unequal bodies of the same material falling in the same medium cannot affect each other's natural motion through their weights. Therefore

(B4) "great and small bodies of the same specific gravities move with like speeds" (= A8).

To reiterate: the original *reductio* successfully refutes the proportionality of natural speed to weight, since the contradiction depends only on this premise, the natural speed hypothesis (premise A2), and the fact that the combined body is larger. But Galileo then puts the thought experiment to work as an "external measure of comparison" (Feyerabend 1975, 76) to discover the natural interpretation that leads to the contradiction (here the tacit assumption that when a falling body has attained its natural speed in a given medium, it weight is still operative). At the same time, this suggests a new natural interpretation that will, when it replaces the old one, remove the

under Natural Interpretation	Falling stone *proves*	Motion of the earth *predicts*
All motion is operative	Earth at rest	Oblique motion of falling stone
Only relative motion is operative	No *relative* motion between starting point and earth	No relative motion between starting point and stone

I find this analysis of Feyerabend's persuasive, but also of general applicability to the kind of thought experiments that Brown terms "Platonic". Let me try to support that by giving an analysis of the falling stones thought experiment of the previous section along the same lines.

There Galileo found a contradiction when seemingly incontrovertible Aristotelian premises were applied to the case of a body composed of two unequal stones of the same material falling under their own gravity. Simplicio is perplexed by this: "I find myself in a tangle, because it still appears to me that the smaller stone added to the larger adds weight to it; and by adding weight, I don't see why it should not add speed to it, or at least not diminish this [speed] in it" (Galileo 1638, 108). But this, Galileo has Salviati respond, is to "commit another error", "because it is not true that the smaller stone adds [*accresca*] weight to the larger" (*ibidem*). That is, he argues that this interpretation of the additivity of weight as a natural property of bodies in all circumstances does not hold for falling bodies. "We feel weight on our shoulders when we try to oppose the motion that the burdening weight would make; but if we descended with the same speed with which such a heavy body would naturally fall, how would you have it press and weigh on us?" (Galileo 1638, 108). If a body in a medium has already attained its natural speed, then there's a balance between its weight and the resistance. It will no longer feel its weight, nor the weight of a body added to it. Weight is not operative for a body moving at its

earth and *then inquire* what changes will remove the contradiction" (Feyerabend 1975, 77). This occurs not immediately, but after considerable further discussion:

> SALV.: With respect to the earth, the tower, and ourselves, all of which keep moving with the diurnal motion along with the stone, the diurnal motion is as if it did not exist; it remains insensible, imperceptible, and without any effect whatever. All that remains observable is the motion which we lack, and that is the grazing drop to the base of the tower. You are not the first to feel great repugnance toward recognizing this non-operative quality of motion among the things which share it in common. (Galilei 1967, 171)

Feyerabend remarks: "The interpretation which Galileo uses restores the senses to their position as instruments of exploration, but only with respect to the reality of relative motion. Motion 'among things which share it in common' 'is non-operative', that is, it remains insensible, imperceptible and without any effect whatever'" (Feyerabend 1975, 78). Galileo has used the Copernican doctrine to discover a tacit natural interpretation lying behind the Aristotelian position, namely that all motion is operative. Proceeding "counterinductively" (Feyerabend 1975, 77) he has discovered what needs to be changed in order to make the Copernican doctrine compatible with the facts that apparently refuted it. "Galileo's first step, in his joint examination of the Copernican doctrine and of a familiar but hidden natural interpretation, consists therefore in *replacing the latter by a different interpretation*" (Feyerabend 1975, 78-79), namely by the interpretation that only relative motion is operative. This yields the following schema (adapted from Feyerabend's in his 1975, 87):

Figure 3   Falling Stone with Rotating Earth
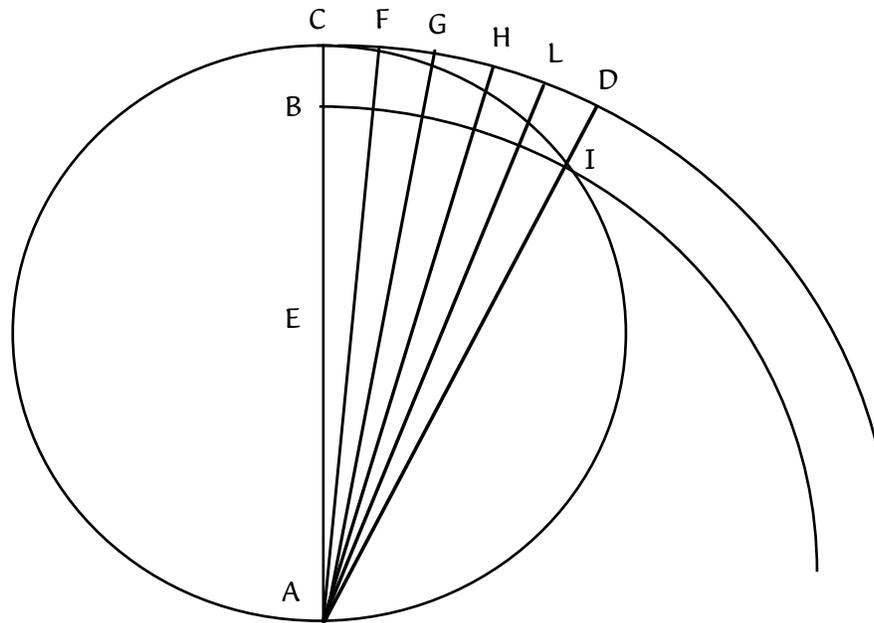
On Feyerabend's reading of this, Galileo offers a new *natural interpretation* of the

phenomenon in question, the fall of the stone down the side of the tower. Or rather, he

uses the Copernican view —apparently directly contradicted by the empirical facts— as

a kind of "external measure of comparison" (Feyerabend 1975, 76) to discover a new

natural interpretation. "Turning the argument around, we *first assert* the motion of the

Here the middle term is the straight and perpendicular fall of the stone:

> SALV.: But wasn't it concluded a little while ago that we could not have any
> knowledge of this fall being straight and perpendicular unless it were first known that
> the earth stood still? Therefore in your syllogism the certainty of the middle term is
> drawn from the uncertainty of the conclusion. Thus you see how, and how badly, it is
> a paralogism (Galilei 1967, 140).

That is, if the stone fell with a compound of two motions, one straight down, the other a
rotation, it would not fall in a straight line. As Galileo goes on to suggest, if the fall were
at a uniform speed, and the earth's rotation were also uniform, the compounded motion
would be an Archimedean spiral (Galilei 1967, 164). Since however, it accelerates
downwards, its compounded motion will be something closer to the arc CI of the circle
depicted in figure 3 (Galilei 1967, 165):

On thought experiments as *a priori* science   ©R. T. W. Arthur

persuade his token Aristotelian to consider a reinterpretation of the phenomenon under

discussion. He does this by asking him how things would look on the Copernican view:

> SALV.: But if it happened that the earth rotated, and consequently carried along the
>
> tower, and if the falling stone were seen to graze the side of the tower just the same,
>
> what would its motion have to be?
>
> SIMP.: In that case one would have to say "its motions", for there would be one with
>
> which it went from top to bottom, and another one needed for the path of the tower.
>
> SALV.: The motion would then be a compound of two motions; the one with which it
>
> measures the tower, and the other with which it follows it. From this compounding it
>
> would follow that the rock would no longer describe that simple straight perpendicular
>
> line, but a slanting one, and perhaps not straight.
>
> SIMP.: I don't know about its not being straight, but I understand well enough that it
>
> would have to be slanting, and different from the straight perpendicular line it would
>
> describe with the earth motionless...
>
> SALV.: Then here clear and evident is the paralogism of Aristotle and of Ptolemy,
>
> discovered by you yourself. They take as known that which is intended to be proved
>
> (Galilei 1967, 139-140).

That is, Aristotle has committed a *petitio principii*: the conclusion, namely the

motionlessness of the earth, is already presupposed in the first premise. Galileo spells

out the syllogism explicitly, and then deconstructs it:

> P1: If the earth rotates from west to east, then a body falling freely from a high tower
>
> will land a long way to the west of the base of the tower, and thus will not fall along a
>
> straight and perpendicular line directed toward the center of the earth.
>
> P2: But a stone falling from a high tower lands at the foot of the tower exactly under
>
> where it was dropped, i.e. falls in a straight and perpendicular line etc.
>
> Conclusion: The earth does not rotate from west to east.

> rock was let fall, being carried by the whirling of the earth, would travel many
>
> hundred of yards to the east in the time the rock would consume in its fall, and
>
> the rock ought to strike the earth that distance away from the base of the tower
>
> (Galilei 1967, 126).

This is adduced by Galileo as "the strongest reason" among several thought

experiments offered by Aristotle and Ptolemy as effective arguments against the earth's

motion: a projectile thrown to a great height vertically upwards would land miles to the

west of its launching place, a cannonball shot to the north or south would curve to the

west, and ones shot to the east or west a distance of, say, five miles while the earth

rotated three miles underneath, would land two miles east or eight miles west,

respectively (Galilei 1967, 126-127); a bird on leaving its perch would see it whisked

away a mile to the east in a few seconds, and clouds and winds would also rush

westwards from the earth at a comparable rate (Galilei 1967, 132); and so forth. The

principle of all these thought experiments is the same as that underlying the tower

argument, and can thus be accommodated to it, as Galileo has Salviati note in the

Dialogue. He then continues:

> SALVIATI: In order to begin to untie these knots, I ask Simplicio by what means he
>
> would prove that freely falling bodies go along straight and perpendicular lines
>
> directed toward the center, should anyone refuse to grant this to Aristotle and
>
> Ptolemy.
>
> SIMPLICIO: By means of the senses, which assure us that the tower is straight and
>
> perpendicular, and which show us that a falling stone goes along grazing it, without
>
> deviating a hairsbreadth to one side or the other, and strikes at the foot of the tower
>
> exactly under the place from which it was dropped (Galilei 1967, 139).

As Feyerabend comments, "the correctness of the observation is not in question"

(Feyerabend 1975, 71). No new empirical data are invoked. What Galileo wants to do is

consistency of Galileo's redescription in terms of his own principles.[8] She offers an austere reconstruction of the thought experiment that is in accordance with Norton's eliminative thesis (all references to paticulars are eliminated), as an argument with only two premises, and which delivers the desired conclusion (A7) soundly.[9] But she points out, once this is done (however it is done), and the Aristotelian can see how the conclusion follows, he or she is in a position to find ways out by denying one or more of the assumptions. "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" (Gendler 1998, 407). What has been lost in the reconstruction, Gendler argues, is "the way in which, by evoking tacit knowledge about how the falling bodies actually behave, the thought experiment pre-emptively precludes such ways out" (Gendler 1998, 407).

Now I do not want to reiterate all Gendler's arguments here. Rather what I wish now to suggest is that her analysis is highly reminiscent of Paul Feyerabend's construal of Galileo's method of argumentation. Her conclusion that thought experiments persuade in a way that cannot be reduced to explicitly reconstructed arguments can be reached, I argue, by means of Feyerabend's notion of "natural interpretation", which he uses to particularly good effect in analysing a different thought experiment concerning falling bodies given by Galileo in his *Dialogo* (Feyerabend 1975, 73ff.).

IV *Natural Interpretations*

As Feyerabend mentions, Galileo's target in this section of the *Dialogo* is an argument, effectively a thought experiment, that not only convinced Tycho Brahe, but was used by Galileo himself in his *Trattato della sfera* (Feyerabend 1975, 70). It is that

> heavy bodies, ... falling down from on high, go by a straight and vertical line to
> the surface of the earth. This is considered an irrefutable argument for the earth
> being motionless. For if it made the diurnal rotation, a tower from whose top a

Norton, on the other hand, construes the move from (A7) to (A8) (in his numbering, from 8 to 9) as a second argument, and thus not a direct perception of a universal law of nature. Moreover, on his view this second argument depends critically on the assumption (numbered 8a by him) that "the speed of fall of bodies depends *only* on their weights". This, he points out, is an assumption *we* make all too naturally when reading Galileo's text, since we modern readers "are usually unable to resist the temptation of dismissing the medium through which bodies are falling as a confounding distraction that should be idealized away" (Norton 1996, 344). Once the assumption is made (A8) follows trivially. The trouble is, however, that the Aristotelian (here represented by Simplicio) is extremely unlikely to accept it. Worse, Salviati cannot legitimately appeal to it either, since the context of the argument at this stage is that of bodies of unequal weights falling through the same medium; and in a given medium speed depends also on the resistance of the medium. Thus, Norton concludes,

> On the argument view the transition from 8 to 9 [i.e. from my (A7) to (A8)] is, at worst, as fallacious inference to a falsehood; or, at best, valid only insofar as it is invoked in special cases in which assumption 8a holds, such as the fall of very heavy, compact objects in very rare media. This final step now looks more like a clumsy fudge or a stumble than a leap into the world of Platonic laws (Norton 1996, 345).

This is a disappointing result. If we admire Galileo's thought experiment because we surreptitiously assume the truth of its conclusion, whilst an Aristotelian would simply fail to be persuaded because he could not accept one or more of the premises, then it is simply a bad thought experiment. Both Galileo's contemporaries and we moderns have been taken in by Galileo's powerful rhetoric.

Tamar Gendler refuses to accept this. She argues that one of the things that the thought experiment does is to represent the phenomenon of fall in such a way that the Aristotelian is reluctantly persuaded of the joint untenability of his premises, and the self-

(A4) If a body whose natural speed is 8 degrees of speed is combined with a body whose natural speed is 4, their composite will fall with a speed less than 8 degrees of speed.

Also

(A5) the two stones joined together make a composite stone larger than the one which moved with eight degrees of speed;

But now from (A1) and (A5) it follows that

(A6) this composite stone will fall with a speed greater than 8 degrees of speed.

Obviously (A4) and (A6) together entail a contradiction. Now assuming the truth of the independent premises (A2) and (A5), it follows that (A1) is false:

(A7) = ~(A2) A heavier body does not fall through the same medium more rapidly than a lighter one made of the same material.

That is, the speed of fall of bodies does not depend on their weights. From this Galileo concludes, after a little more dialogue, that

(A8) "great and small bodies of the same specific gravities move with like speeds".

Now Brown grants that the thought experiment, insofar as it can be regarded as a successful refutation of Aristotle's assumption —that is, in yielding conclusion (A7)— is a *reductio* argument. But it does more: it establishes a new theory: "the question of which [body] falls faster is obviously resolved by having all objects fall at the same speed." (Brown 1991, 77). It is this feature, the construction of a new theory, that makes the experiment "Platonic".[7]

argument for the instability of the earth because it is not deductively valid (Fisher 1988, 122-3), or Galileo's argument against the Aristotelians because some of his premises, whilst acceptable to his opponents, are now known to be false (Fisher 1988, 96). For this reason, following Fisher's example, I shall stick as closely as possible to Galileo's own words. My reconstruction differs from his,[5] as well as from those of Norton and Gendler in certain respects whose importance will be made clear later. (The fact that all our reconstructions differ already supports Fisher's point about the non-triviality of reconstruction, and casts some doubt on the idea that it is straightforwardly equivalent to one of these reconstructions.)

Let us begin by identifying the supposition that the *reductio* is intended to refute, the Aristotelian premise that the natural speed depends on weight, i.e. that

> (A1) a heavier body falls through the same medium more rapidly than a lighter
> one made of the same material.[6]

Galileo then gains Simplicio's assent to a crucial premise,

> (A2) Every falling body acquires a natural speed, i.e. a "definite speed fixed by
> nature, a speed which cannot be increased or diminished except by the use of
> force or some impediment that retards it".

But from (A2) it follows that

> (A3) if we had two bodies whose natural speeds were unequal, it is evident that
> on uniting the two, the more rapid one would be partly retarded by the slower,
> and the slower one would be somewhat speeded up by the faster.

Therefore, particularizing,

definite speed fixed by nature, a speed which cannot be increased or diminished except by the use of force or some impediment that retards it. [Simplicio acquiesces.]

Then if we had two bodies whose natural speeds were unequal, it is evident that on uniting the two, the more rapid one would be partly retarded by the slower, and the slower one would be somewhat speeded up by the faster. [Again Simplicio agrees.]

But if this is so, and if it is also true that a large stone moves with eight degrees of speed, say, while the slower moves with four, then when they are joined together, their composite will move with a speed less than eight degrees. But the two stones joined together make a larger stone than the one which moved with eight degrees of speed; therefore this greater stone moves less rapidly than the lighter; which is contrary to your supposition. Thus you see how, from the supposition that a heavier body moves more rapidly than a lighter one, I infer that the heavier body moves less rapidly (Galilei 1638, 107-108; Galilei 1974, 66-67).

It is natural to construe this as an argument—in fact, quite a beautiful example of a *reductio ad absurdum*. Indeed, an excellent little textbook by Alec Fisher that I have used in my logic classes begins with precisely this argument (Fisher 1988, 1-2) and later analyzes it in detail (91-98). One of Fisher's main points in his book is that it is no trivial matter to reconstruct a natural argument (that is, one given in its original context and wording) like Galileo's. In order to facilitate this he suggests applying what he calls "the assertibility question" —"What argument or evidence would justify me in asserting the conclusion?"—in order to determine whether the various inferential steps are valid, or the premises acceptable —acceptable, that is, according to "appropriate standards of evidence or appropriate standards of what is possible" (Fisher 1988, 27). This requires some sensitivity: we do not want to be rejecting a perfectly good argument because we have adopted inappropriate standards of proof or evidence —rejecting, say, Darwin's

the original thought experiment cannot be reconstructed as an explicit argument without some epistemic loss. This has been argued in detail by Tamar Szabó Gendler (Gendler 1998), by reference to one of Galileo's most celebrated thought experiments, one that has served as a touchstone for much of this debate, and which I shall now consider.

III *Galileo's Falling Bodies*

This thought experiment occurs in Galileo's *Discorsi (Two New Sciences)*, in the context of a consideration of Aristotle's proof of the impossibility of motion in a void. One of the assumptions on which Aristotle's proof depends is that bodies of differing weights [*gravità*], moving in the same medium, have natural speeds directly proportional to their weights, so that, "for example, a moving body ten times as heavy as another moves ten times as fast" (Galilei 1638, 106); the other is that the speeds of the same moving body in different mediums are in inverse ratio to the crassitudes or densities of the mediums" (*ibid.;* Drake 1974, 65). Galileo, in opposition to this, argues first that "great and small bodies of the same specific gravities move with like speeds", and then that "if one were to remove entirely the resistance of the medium, all materials would descend with equal speed" (Galilei 1638, 109, 116). The thought experiment occurs in arguing the first point. Galileo has Sagredo claim that he has done an experiment in which a hundred-pound cannonball and a half-ounce musket ball are dropped through a height of two hundred bracchia , and has determined that the cannonball "does not anticipate the musket ball's arrival on the ground by even half a span" (Galilei 1638, 106-107). The dialogue continues:

> SALVIATI: But even without experiment, it is possible to prove clearly that a heavier moving body does not move more rapidly than another less heavy one, provided both bodies are of the same material, and in short such as those mentioned by Aristotle. But tell me, Simplicio, whether you accept that each falling body acquires a

argues, "always eliminable without compromising the conclusion", precisely by this reconstruction of the thought experiment as an explicit argument.

Thus, concerning Brown's related figure-demonstration above, Norton argues that the commentary he supplies is an argument, and this is "just what the reconstruction thesis calls for", for without such an argument the figure is bereft of persuasive power (Norton 1996, 353). But, one might counter, what about the non-propositional elements of the demonstration? After all, one could turn Norton's argument on its head: without the figure, his commentary, far from being an argument, is not even intelligible. Norton might want to claim that the figure is in principle replaceabe by a description of it in words, but I think this is irrelevant to the success of the figure as a demonstration-sketch, and thus also to the point in question. Indeed, Brown himself has plenty of very persuasive arguments about the non-eliminability of figures and diagrams in mathematics, which might also apply to non-propositional features of thought experiments (Brown 1997, Brown (forthcoming)).

I will not argue this point further here. Suffice to say that even if Brown's notion of Platonic perception is irrelevant epistemically, this does not entail that all non-propositional features of a thought experiment are irrelevant epistemically. In the remainder of this paper I shall disregard the elimination thesis, and will interpret "argument" in a wide enough sense to include Norton's "particulars irrelevant to the generality of the conclusion".

But what about the reconstruction thesis in this case? I am quite prepared to accept that a thought experiment can be reconstructed as an explicit argument, and even that such reconstructions are an essential part of the process of science. But the question is whether a given reconstruction is epistemically equivalent to the original thought experiment. Here I think Norton underestimates the difficulty of making explicit the relevant background assumptions. In many cases this is far from trivial, especially in the case of the innovative ("Platonic") experiments I am considering here. In such cases

that the figure would be any less efficacious for him in yielding the theorem. So Brown's view passes the "Saunderson test", and my dream criticism fails.

Still, this does not appear to oblige us to accept Brown's idea of a Platonic revelation of the theorem through a mere study of the figure. For we need to be in possession of quite detailed background knowledge in order to be able to interpret the figure at all. (I myself, primed with a quite different body of background knowledge from my historical studies, might have taken it for a proof of the t^2 law for falling bodies, as it is almost identical with the figure that Descartes drew for Beeckman in solving that problem in 1618.[4]) It is only by virtue of this tacit background that the figure acquires any demonstrative force.

Norton makes a related point. He believes that what happens when we "see" the theorem is that we run through a kind of commentary on the figure, filling in at least tacitly the steps that take us to the conclusion. This can be done with greater or lesser speed depending on one's mathematical ability, but such a step-wise reconstruction leaves no room for the holistic act of intuition that Brown conceives Platonic perception to be. Moreover, Norton claims, when this reconstruction is made explicit, what we have is an informal argument, whose soundness we can determine. This supports his general thesis that "any thought experiment can be reconstructed explicitly as an argument", or more precisely:

> *Reconstruction Thesis*: All thought experiments can be reconstructed as arguments based on tacit or explicit assumptions. Belief in the outcome-conclusion of the thought experiment is justified only insofar as the reconstructed argument can justify the conclusion (Norton 1996, 339).

This thesis is closely related to a thesis Norton had proposed earlier, the "elimination thesis". Certain "particulars irrelevant to the generality of the conclusion" of a thought experiment, such as falling musket balls, light beams and twisting buckets are, Norton

351). An example is the proof-sketch Brown offers of the theorem that the sum of the first n integers, $\Sigma_{1\,\text{to}\,n}\, n = (n^2 + 1)/2$ (see, e.g. Brown 1977, 170). It consists solely in the following figure:

Figure 2 The sum of the first n (five) numbers
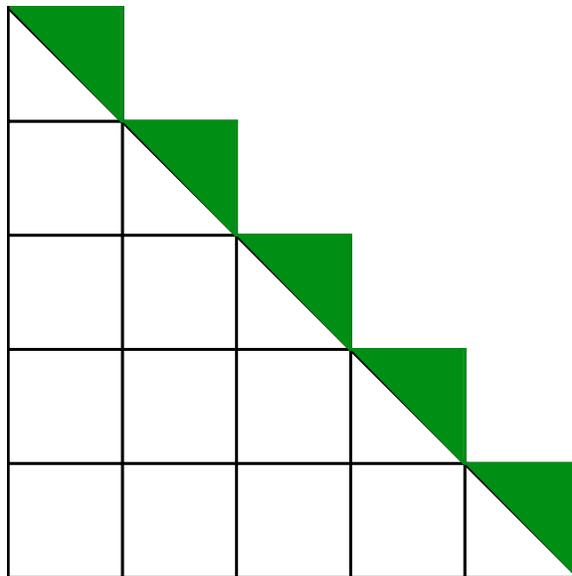
"Brown's idea," explains Norton, "is that we immediately see the truth of the theorem in this figure without supplement of any text and that we do so without elaborate mathematical inferences. The suggestion is that this moment of mathematical revelation coincides with the grasping of mathematical law" (Norton 1996, 352).

Now this talk of "seeing the truth of the theorem" (like so much of our philosophical discourse) involves a sighted bias. Does this bias vitiate Brown's idea of perceiving laws, as my dream appeared to suggest? Actually, no. To appreciate this it is only necessary to revert to the real Saunderson and his ingenious mechanical devices. The task of constructing the above figure on one of his boards (see Figure 1) would have been trivial for a mathematician of his accomplishment, and there is no reason to think

by the human mind. This is an objective view of *a priori* knowledge — it posits a non-sensory perception of independently existing objects" (Brown 1991, 86-87). Of course, one can be mistaken about what one sees; hence the *a priori* knowledge gained through this kind of perception is fallible, according to Brown, thus explaining the fallibility of thought experiments.

That said, my dream seems interpretable as an attempt to break the Brown line, if I may call it that, from thought experiments to *a priori* physics by denying the persuasiveness of the Platonic vision to a blindman. The dream-me seems to be suggesting that the idea of "seeing the laws of nature" would be no more persuasive to a sightless person than would the argument from design of Newton, Clarke and Leibniz to Diderot's Saunderson.

II *Thought Experiment as Picturesque Argument*

Thus one might argue that the non-propositional features of a thought experiment are inessential elements. The neat illustrations that make them so visually appealing, such as Einstein racing his light beam, or Newton's twisting bucket, serve as aids to comprehension, but are not epistemically indispensable. If a thought experiment is to persuade someone —anyone, including a sightless Saunderson— then it must be reformulatable as an argument. This is the gist of Norton's criticism of Brown's position. "Platonic perception of laws is essentially irrelevant epistemically whenever it transcends argument since its products are only reliable insofar as they can be checked by explicit argumentation"
(Norton 1996, 335).

In making his case, Norton notes that Brown, in his attempts to explain Platonic perception, freely appeals to proof-sketches in mathematics that incorporate visual elements to deliver the theorem in question. "These demonstrations are the analog in mathematics of Platonic thought experiments in the natural sciences" (Norton 1996,

aesthetic, which Saunderson as a blind man could not share; and which, generalizing, can be seen to be contingent on the possessing of certain organs of sense. Man might have evolved, or might yet evolve, Diderot suggests, with a different set of sensory organs, yielding an entirely different outlook in which visual harmony would simply not appear (Diderot 1964, 120-122). On the one hand this has the effect of undercutting the force of the argument from design, but Diderot's deeper point is that a blind person is not distracted by the visual wonders or mirages that the sighted person is. The blind person has his feet on the ground, as it were; Saunderson's grasp of reality comes, not from contemplating abstract objects with a geometer's intuition, but from holding objects in his hand, from manipulating his algebraic-cum-geometric calculating devices.[3]

Figure 1   Saunderson's calculating devices

With this relevant background knowledge in place, I propose to interpret my dream as a cryptic encapsulation of a criticism of the position Brown takes in his book, which I had just read at that time. Brown is an unabashed Platonist in mathematics, and he sees thought experiments as a natural means for extending the Platonist argument from mathematics to physics. If mathematical truth depends on relations between real mathematical objects, relations that have been codified in mathematical facts and are transparent to our intuition, then perhaps the argument can be extended to physics using thought experiments as a Trojan Horse. One can imagine that the fundamental truths of physics, just like those of mathematics, are true of the world but can nonetheless be intuited *a priori* by the physicist in a direct way. Thought experiments, on this view, are one of the most powerful means at our disposal for discovering that *a priori* knowledge.

This is just what Brown argues. In his chapter "Seeing the Laws of Nature" he proposes that physical laws are universals, and that "universals (properties and relations) have an existence of their own, and like mathematical objects can be grasped

On thought experiments as *a priori* science   ©R. T. W. Arthur

Brown's work has been devoted to showing that the sense of touch had nothing to do with reality, holding up and pointing to a page of Brown's *c.v.* as evidence! Considering this alleged core thesis, I comment "Saunderson would hardly agree". To this the mysterious figure responds with the unfathomable remark: "You told me that your own imbecile is always walking around saying that you are him"! At this point, I wake up.

Now I'm not one of those who believes that every aspect of a dream has some deep and abiding meaning. I awoke under the impression that the last comment had some profound relevance for personal identity, only to be swiftly disillusioned in the cold light of day. Nevertheless, I can throw light on other features of the dream fragment that may be completely obscure to others, but about which I have privileged knowledge. I know, for instance, the identity of Saunderson: he was a distinguished 18th century algebraist (1682-1739), the fourth occupant of the Lucasian Chair at Cambridge after Barrow, Newton and Whiston, and who also happened to be totally blind. But this reference is to a fictionalized version of the same man: Denis Diderot, in his *Lettre sur les aveugles* (Letter on the Blind) (1749), made brilliant use of a (forged) first-hand report on Saunderson's dying moments to make some canny points in favor of evolutionary materialism. The gist of this was that a person like Saunderson lived solely in a world of touch, he operated in a purely tactile space; he did not have to correlate this with a visual space, as we do, because he had none, having been blind almost from birth. As a result of this reliance on different sense organs he had a different view of the world, and in particular the *order* of the world, than might somebody whose cues for such things as order and harmony were predominantly visual (Diderot 1964, 122-124). Diderot uses this to make an interesting point about the *argument from design* for God's existence touted by Isaac Newton and Samuel Clarke, a point which applies equally to Leibniz's conception of God as creating the world with a maximal pre-established harmony, the main paradigm of which is a visual perspective.[2] Diderot intimates that this argument from design derives its force from our appreciation of visual orderliness, the visual

<u>On Thought Experiments as A Priori Science</u>

In what follows I shall be considering the merits of two questions: whether thought experiments should be regarded as a species of argument, and whether certain thought experiments may be construed as yielding knowledge of the physical world *a priori*. The first is the view of John Norton, who urges it in several articles in vehement opposition to the second view, proposed by Jim Brown in his book *The Laboratory of the Mind* and elsewhere. Norton is alarmed by Brown's argument that thought experiments of a certain kind not only refute some existing theory, but enable one to "see" the relevant laws of nature by "a kind of perception". Brown calls a thought experiment of this kind "Platonic", and argues that it is "is *a priori* in that it is not based on new empirical evidence nor is it merely logically derived from the old data". In an effort to block this path to Platonism, Norton insists that "thought experiments are merely picturesque arguments and in no way remarkable epistemologically" (1996, 334).[1] Although I am no more inclined to accept Brown's Platonism than Norton is, I argue that in his zeal to block it, Norton fails to realize that thought experiments of this kind certainly *are* remarkable epistemologically. Not only are they not reducible to formulated arguments without loss of vital pedagogical elements, they are paradigms of how scientific advances may be made *a priori*.

I. *A Brown Study*

There's a double pun in my title for this section. It is, of course, a study of Brown's position, but there's a less obvious pun on its content and form. For I intend to begin with an apparently irrelevant and absent-minded digression about a dream I had a few days before Christmas 1995, for which I beg indulgence.

The dream is little more than a vivid fragment, which I recorded as soon as I awoke. In it a mysterious character in a Fedora hat is trying to persuade me that all Jim