

# Thought Experiments: Determining Their Meaning

Igal Galili

Published online: 7 November 2007  
© Springer Science+Business Media B.V. 2007

**Abstract** This paper considers thought experiment as a special scientific tool that mediates between theory and experiment by mental simulation. To clarify the meaning of thought experiment, as required in teaching science, we followed the relevant episodes throughout the history of science paying attention to the epistemological status of the performed activity. A definition of *thought experiment* is suggested and its meaning is analyzed using two-dimensional conceptual variation. This method allows one to represent thought experiment in comparison with the congenerous conceptual constructs also defined. A similar approach is used to classify the uses of thought experiments, mainly for the purpose of science curriculum.

## 1 Introduction and Background

Thought Experiments (TEs) have attracted the attention of many researchers in modern day science and philosophy. Although we can trace their use back to the dawn of natural philosophy, TE is commonly associated with the physicist Ernst Mach (1883/1989) and his notion of thought (“mental”) experiment (*Gedankenexperiment*).<sup>1</sup> He made it an integral part of his revision of mechanics, the most established domain of physics knowledge at that time. TEs helped Mach in providing mechanics with an empirical (instead of metaphysical) foundation.

In fact, a variety of arguments, fragments of debates, and scientific presentations are currently identified as TEs. What initially was understood by Mach as TE was not more

---

<sup>1</sup> An anonymous reviewer argued that the term was already used by Ørsted in 1811 (Ørsted, H. C.: 1811, *Naturvidenskabelige Skrifter*, 3, 151–190), as discussed in Witt-Hansen, J.: 1976, H. C. Ørsted, ‘Immanuel Kant, and the Thought Experiment’, *Danish Yearbook of Philosophy* 13, 48–65, and in D. Cohnitz of the University of Düsseldorf in his forthcoming article. This fact, however, cannot diminish the leading role of Mach in introducing this construct into active use, particularly in education.

---

I. Galili (✉)  
Science Teaching Center, The Hebrew University of Jerusalem, Jerusalem 91904, Israel  
e-mail: igal@vms.huji.ac.il

than the idea of an experiment, which did not require actual performance because of the *obviousness* of the results expected. It was Einstein, who essentially upgraded TE as a concept, making it a standard research tool of fundamental importance, “a principal means through which scientists change their conceptual structures” (Nersessian 1993).

The concept of TE has been discussed in the philosophy of science and many of its features have been elaborated upon since Mach (e.g., Popper 1968; Kuhn 1977; Brown 1991, 2002, 2004; Sorensen 1992; Nersessian 1993; Norton 2004a, b, c). Popper (1968) stated three *uses* of “imaginary experiments” in physics: critical, heuristic, and apologetic, in the context of his critique of the foundations of quantum mechanics. Based on the history of physics, Brown, in 1991, classified TEs into three major classes: constructive, destructive and their combination—the Platonic type. Several definitions of TE were provided while addressing different aspects of this concept.

In parallel, TE was praised as possessing a high potential for the teaching–learning process (Helm and Gilbert 1985; Helm et al. 1985; Matthews 1994; Gilbert and Reiner 2000; Reiner and Burko 2003). The proponents of this use often cite Mach (1896/1976), who first elaborated the merits of TE for education. The entry of science education into the discourse on the nature of TE and its use in science is natural due to the role of conceptual clarification that the philosophy of science provides both to the science (the *object* in the semantic triangle: science–science education—philosophy of science) and science education (the vertex of *sign* in that same triangle, Tseitlin and Galili 2005).

The term *thought experiment* is seldom used by scientists. Thus, Niels Bohr never used this term while discussing the TEs of Einstein. He mentioned the “devices proposed by Einstein” and their “pseudo-realistic style” (Bohr 1949/1959: 226). In presenting their ideas, prominent physicists may not even specify whether they are addressing real or imaginary experiments (e.g., Feynman et al. 1965: 6–3–6–4; Penrose 1997: 254–255) and just write “consider an experimental set-up ...”. Born (1944), in his review of contemporary science, ignored TE by stating a clear dichotomy of physics practice—theory versus experiment. The same author included many of TEs, although never named them this way, in his renowned presentation of *Einstein’s Theory of Relativity* (Born 1924/1965).

Einstein opened his seminal 1905 paper on the electrodynamics of the moving bodies with a TE of a magnet and conductor in relative motion but did not use the term TE (Einstein 1905/1952a: 37–65). In his *Autobiographical Notes* the TE of chasing light (Einstein 1949/1979: 53) was termed “a paradox” and the TE that explained the status of inertial frames by describing the people “who know only a small part of the earth’s surface” was termed an *analogy* (*ibid.*: 29). In the *Evolution of Physics* (Einstein and Infeld 1938) we find an “idealized experiment” with regard to the famous Einstein’s elevator (*ibid.*: 226–235). Other popular science books that use TE abundantly avoid this term too, preferring “conceptual (or imaginary) experiment”, or just “paradox” (e.g., Park 1988).

However, the fact remains that although physicists refrain from defining TE, in practice, they use it very often. This fact suggests that they consider TE so obvious that it does not require definition.<sup>2</sup> Such a situation challenges philosophers, historians, and educators all

<sup>2</sup> Somewhat similar to scientists, historians of science often do not mention TEs as a special means of furthering scientific claims. For example, Moss (1993), in her comprehensive study of Galileo, referred to his famous TEs as “dialectical arguments”, “eloquent examples”, “analogies”, “experiments”, but never—“thought experiments”. Similarly, among the philosophers of science, Giere (1984) in his course of philosophy of science, surveys the major methods of scientific reasoning and never mentioned “thought experiment”. In his analysis of the modern science (1999), “thought experiment” does not appear. He speaks about *prototypes* and *things*, and *models*, as “providing guidelines for the relevant similarity judgments” (*ibid.*: 123).

of whom reflect on the methods of scientific practice. Miller (1986) addressed TE as a specific form of scientific research employing imagery in scientific thought. Sorensen (1992) and Cushing (1998) stated that TE exposes the essence of scientific ideas. Popper (1934/1968) wrote that TE serves as a major tool of scientific discourse, and Matthews (1994) argued that TE provides significant conceptual benefits to the learners of science. TEs serve as a leading device in non-formal presentations of science (“conceptual experiments”) because of their vivid and appealing style of depicting fundamental ideas of physics while avoiding both heavy formalism and real experimentation (e.g., Harrison 1981; Park 1988).

Several researchers elaborated on the merits of using TE in teaching physics (e.g., Stinner 1990, 2005; Gilbert and Reiner 2000; Reiner and Gilbert 2000; Lattery 2001), and those who follow this discourse might have a feeling that TE are ubiquitous through all history of physics. At the same time, the epistemological status of TE usually remains on the sidelines in *educational* discourse and is never mentioned in textbooks as if “it is known to everyone”.<sup>3</sup> Lack of discussion on the subject may lead to confusion on the part of students with what TE both “*can*” and “*cannot do*”, and misinterpretation of this fundamental element of scientific research. Thus, reading Koyré’s (1968: 75) analysis of Galileo’s use of TE one could be puzzled by the following remark:

I do not reproach him on this account; on the contrary, I should like to claim for him the glory and the merit of having known how to dispense with experiments ... (shown to be indispensable by the very fact of his having been able to dispense with them): yet the experiments were unrealizable in practice with the facilities at his disposal.<sup>4</sup>

Scientists often apply TE spontaneously, as a model, without epistemological explanation (e.g., Peierls 1980); and philosophers continue to debate its epistemological status (Brown 2002, 2004; Gendler 2004; Norton 2004). Our major concern here, however, is the use of TE in education. Physics textbooks normally lack a definition of TE despite their frequent employment. Various activities are addressed as TE in educational research discourse often making this concept indistinguishable from merely *thinking* about physics.

This situation encourages conducting a survey of TE relevant for educational purposes. In this paper, after placing TE into historical context, the epistemological status of TE will be discussed together with its features and possible definitions. Two-dimensional variation of features will help to clarify the specificity of TE in comparison with other constructs. A similar approach will be applied regarding the typology of using this concept. All these will facilitate considering the implications of TE for science education.

## 2 Historical Perspective

TE represents a mental activity involving imagination and theoretical thinking addressing real objects. In retrospect one may identify TE among the major investigative tools of science prior to the scientific revolution of the 17th century. The inductive–deductive

<sup>3</sup> In these words Newton introduced the concepts of time and space in his *Principia* (1999: 408). Posteriorly, after Einstein, we know that Newton was wrong in this subject.

<sup>4</sup> Koyré was considering the experiments with falling bodies. Vacuums became available to researchers later in the 17th century due to Otto von Guericke and Robert Boyle. Koyré noted the imperfect nature of the instruments available to Galileo, the absence of a reliable clock, etc.

method of scientific exploration often took the form of TE to interpret and explain natural phenomena (e.g., Losee 1993: 6–13).<sup>5</sup> For example, in proving the spherical shape of the Earth, Aristotle thought about what would happen if a lump of mass is added to the spherical earth. Referring to his previously introduced principle of seeking the center of the universe, he inferred that the earth would “continue to move until it occupies the center [of the universe] equally every way” (Aristotle 1952a: 388–389) and restore the spherical shape. A similar method was employed to determining the nature of the cosmic-edge (Harrison, 1981: 104–107).

In the following Hellenistic period of science, although the empirical activities came to the fore, the hypothetico-deductive method of scientific exploration was preserved (Mason 1962: 48–60; Russo 2004: 185–187). Ptolemy used a TE based on Aristotle’s *On the Heavens* (1952a: 388), to show the impossibility of Earth’s movement. Imagining the Earth rotating, he wrote in the *Almagest* (Ptolemy 1952: 12):

those things that were not at rest on the earth would seem to have a movement contrary to it and never would a cloud be seen to move towards the east nor anything else that flew or was thrown into the air. For the earth would always outstrip them in its eastward motion, so that all other bodies would seem to be left behind and to move towards the west.

This argumentation applied a certain theoretical view to an imaginary (for Ptolemy) situation of Earth’s rotation and led him to the conclusion: the Earth is at rest. Important scientific discoveries employed TE: Archimedes demonstrated the law of the lever (Feather 1959: 162–164), and Philoponus disproved the Aristotelian self-pushing movement of projectiles (antiperistasis) (Cohen and Drabkin 1966: 221–223; Pedersen and Pihl 1974: 124).

In medieval science, although Grosseteste and Roger Bacon, in the 13<sup>th</sup> century, limited the all-prevailing theoretical account by the necessity to observe real objects, “the prerogative of experimental science” (Losee 1993: 32, 58), the regular way of doing natural philosophy remained *secundum imaginationem* (“according to imagination”), even if the results of this treatment contradicted the actually observed (Murdoch and Sylla 1978; King 1991). TE served as a major tool of argumentations for Buridan, Oresme and Albert of Saxony. In their writings we find the famous TE of a spinning top (rejecting the external pushing force), the lance pointed at both ends (rejecting self-pushing propulsion), the boat continuing to move by inertia (rejecting the Aristotelian force–motion relationship) (Buridan 1357/1959: 533).

All these TEs supported the new theoretical concept—impetus. Among them one finds the important heuristic TE which introduced into physics pendulum motion. Albert of Saxony considered the falling of a heavy body through a hole perforating the Earth globe and, arguing by impetus, predicted oscillatory motion of the body—“titubando” (Clagett 1959: 566). In the same way, Oresme applied impetus to reason the non-retarded falling of bodies from the mast of the moving ship (Grant 1977: 67–68; Dugas 1986: 63). Buridan and Oresme speculated about the rotation of the Earth and based on the principle of kinematic relativity inferred the uncertainty of the on-ground observer to decide whether either Earth or heavens rotate (Buridan 1357/1959; Grant 1977: 64–68). However, TE in medieval times, although numerous, never proceeded by researchers to the stage in which its results were subsequently tested by a real experiment: theorizing was sufficient.

<sup>5</sup> In fact, this was a revolutionary swing from the empirical knowledge about nature, expressed in objective terms and accumulated in the old civilizations (Sumerian, Babylonian, and Egyptian), to the scientific knowledge as practiced in the Classical Greece.

Real experimentation of controlled variables came to the fore during the scientific revolution of the 17th century, in the empiricist manifesto of Francis Bacon and in practice by Boyle.<sup>6</sup> However, the transition into this new way of doing science was not abrupt, and it took about two centuries of gradual development during which Leonardo da Vinci, Maurolico, Grimaldi, Benedetti, Galileo, Kepler and Descartes remained with one foot in the medieval science.<sup>7</sup>

TE followed this development. In Galileo's writings thought and real experiments are deeply interwoven and often indistinguishable. Some of his experiments proved to be imaginary only (Cohen 1950/1953; Dijksterhuis 1986; Cushing 1998: 81–84). The blurred demarcation between true and thought experiments could be explained by the fact that *experiment* was understood by Galileo in a special way: as “a contrived occurrence determined in its entirety by a phenomenon and to no extent by accidents” (McAllister 1996). Also phenomena were understood differently: as “fundamental modes in which physical reality manifests itself” (ibid.). All these made both real and thought experiments practically signify the same, the “idealized experiments”, contrasting with real experiments in the modern sense.

In any case, Galileo, Descartes, Newton, Huygens, as well as many others, continued to use TEs in their research and theory presentations. The new feature was that TE could precede a real experiment which constituted a verification stage. In modern science the design of an experiment is normally preceded by many TEs, mediating between theory and actual experimentation.

Although *experiment* became a symbol of modern science, TE, often subtle and controversial, remains among the central tools of research, regularly used in scientific debates professional interactions and especially teaching. Therefore, presenting of the scientific methodology without TE is deficient.<sup>8</sup>

### 3 Epistemological Status

The epistemological status of TE is not obvious and has been intensively debated. Norton provided a provocative description:

A scientist—a Galileo, Newton, Darwin or Einstein—presents us with some vexing problem. We are perplexed. In a few words of simple prose, the scientist then conjures up an experiment, purely in thought. We follow, replicating its falling bodies or spinning buckets in our minds, and our uncertainty evaporates.

---

<sup>6</sup> Another opinion is that the break with the medieval science is due to Leonardo da Vinci (Draper 1890: 233), at least, in the continental Europe. This view is not too convincing since Leonardo held the medieval theoretical views (Randall 1957: 207–218). His strong empirical affiliation was not sufficient to produce a revolutionary change in science.

<sup>7</sup> Although the notion of “The scientific revolution” introduced by Koyré (1943) is presently a commonplace, there is a counter view among historians who state the evolution in science: “There was no such thing as the Scientific Revolution” (Shapin 1996: 1).

<sup>8</sup> Real and thought experiments are deeply interwoven. Shamos (1959: 46, 287, 318) in his compendium of *Great Experiments in Physics* included Newton's TEs addressing absolute and relative movements, Maxwell's TEs concerning electromotive force, and Einstein's TE of a magnet and conductor, all in the rubric of “experiments” without any note.

However, the quote of Koyré in the Introduction, as well as other statements praising TE as a device used to investigate *Nature*, may mislead: why do we need real experiments if we can investigate Nature by means of pure thought?

In fact, Einstein addressed this subject directly. In his *Autobiographical Notes* (Einstein 1949/1979: 11) he wrote:

... it appeared that it was possible to get certain knowledge of the objects of experience by means of pure thinking, this “wonder” rested upon an *error*. Nevertheless, for anyone who experiences it for the first time, it is marvelous enough that man is capable at all to reach such a degree of certainty and purity in pure thinking as the Greeks showed us for the first time to be possible in geometry. (italics added)

In this way, the adult Einstein answered the 12-year-old Einstein, fascinated by “the holy geometry” shown to him by his uncle. Einstein explained (*ibid.*):

... This primitive idea, which probably also lies at the bottom of the well known Kantian problematic concerning the possibility of “synthetic judgments *a priori*,” rests *obviously* upon the fact that the relation of geometrical concepts to objects of direct experience (rigid rod, finite interval, etc.) was unconsciously present. (second italics added)

To the adult Einstein it was even a “primitive idea” that proving a new theorem in geometry (within the same set of axioms) is not *synthetic* but rather *analytic* knowledge, that is, implicitly “existing” in the system of axioms and definitions of Euclidian geometry. In a sense, the new theorem indeed represents new knowledge about the real objects (“rods and intervals”), but in another epistemological sense, it does not. It is the latter sense that matches Kuhn’s (1977: 252) assertion: “a thought experiment can teach nothing that was not known before”—a statement that may puzzle those familiar with the crucial role TEs played in the development of classical and modern physics. This issue exactly lies at the heart of the polemics of the philosophers of science (Brown 1991, 2004; McAllister 1996; Arthur 1999; Gendler 2004; Norton 2004b) concerning the *a priori* nature of knowledge as provided by TE.<sup>9</sup> Within this variety Einstein’s position is elucidating: TE can produce new (analytic) knowledge based on certain theory and thus inform about reality, possibly beyond the available empirical knowledge, but still only *within* the known conceptual framework used. In other words, *TE teaches us not about the world (directly) but rather about the theoretical framework we have used in that particular TE.*

One can elaborate this point using an example from the history of science. Galileo in his *Discorsi* (1638/1914: 107) presented the famous TE of falling bodies, originally suggested by Benedetti (e.g., Dijksterhuis 1986).<sup>10</sup> Using an exquisite logic that impressed generations thereafter, Galileo demonstrated the inconsistency of the Aristotelian theory with its loosely defined concepts of “body”, “heavy”, and “light”. Galileo showed that the

<sup>9</sup> Thus, Brown ascribes to the Platonic type of TE the ability to construct new knowledge as in Galileo’s TE of falling bodies; in contrast, Norton refutes this idea: “pure thought cannot conjure up new knowledge” (Norton 2004b: 9).

<sup>10</sup> For different reasons, in many of his important statements Galileo did not acknowledge their authors, the brilliant minds of the medieval and renaissance periods. This behavior could be understood as breaking with the popular scholastic tradition heavily based on the citation of an authority. In the following years, seeking simplicity of a linear presentation, a whole collection of ideas were ascribed to Galileo (e.g., Cohen 1950/1993; Crombey 1959; Gliozzi 1962; Dugas 1986; Koyré 1968; Dijksterhuis 1986).

statement “heavy bodies fall faster” was erroneous within that specific theory, but *not* more than that.<sup>11</sup> Unlike what is often declared to students and stated in discussions (e.g., Brown 2004), Galileo did not claim this TE to be the *proof* that all bodies fall in the same way (acceleration).<sup>12</sup> Although Simplicio, a peripatetic philosopher, was ready to accept the compelling logical refutation of the Aristotelian law, Galileo did not stop and proceeded to empirical considerations: (ibid.: 110):

It is clear that Aristotle could not have made the *experiment*, yet he wishes to give us the impression of his having performed it when he speaks of such an effort as one which we see. (italics added)

Why did he? Why did the reference to experiment become important for Galileo? Perhaps he knew that he struggled with the great all-inclusive theory, and he, Galileo, had no other *theory* to replace it. *Experiment* was recruited to surpass a mere correction of logic. In a long discussion Galileo provided a detailed elaboration of empirical considerations concerning falling of bodies within the medium of decreasing density (ibid.: 116). It was not sufficient for Galileo himself,<sup>13</sup> but he concluded:

Having *observed* this I came to the conclusion that in a medium totally devoid of resistance all bodies would fall with the same speed. (italics added)

This reasoning led Galileo to the *empirical* law of falling bodies (Galileo’s law), which later, after Newton’s theory was introduced, appeared to be only *approximately* correct. Benedetti-Galileo’s TE could indeed guide Galileo’s thinking about falling: if Aristotle was wrong, perhaps one should seek a *theory* which predicts falling independent of mass?—a possibility. As Popper put it (1934/1968: 443):

... the use of the imaginary experiments in critical argumentation is, undoubtedly, legitimate: it amounts to an attempt to show that certain *possibilities* were overlooked by the author of a theory. (italics added)

The Aristotelian *theory* of falling was severely shaken, but no comprehensive *theory* replaced it before Newton. This is in contrast to Brown (1991: 77) who wrote:

<sup>11</sup> In another TE that Galileo introduced in his early work *De Motu Antiquiora* he in a similar way pointed at the *logical error* in the Aristotelian law of falling bodies, this time in the context of the bodies sinking in liquids of different densities. The same TE could be found in Buridan’s *Questions on Physics*, Book VII, Qu. 7 (ed. Paris 1509) (Moody 1994: 188–189).

<sup>12</sup> And indeed they do not. Within the Newtonian theory, the acceleration with which two masses approach each other directly depends on their masses. However, close to the Earth and for the masses much smaller than that of the Earth, in a *very good approximation*, they all fall with the same acceleration (Galileo’s law). Ignoring the approximate nature of Galileo’s law in the modern teaching of physics presents a *conceptually* deficient instruction.

<sup>13</sup> This was a remarkable recourse to experiment by Galileo who showed quite an opposite attitude to experiment in other cases, preferring convincing reasoning free of impeding factors which may mask the phenomenon. We read in *Dialogo* (Galilei 1632: 145): “Did you [Salviati] make an experiment? (Salviati represented Galileo)—No, I do not need it, as without any experience I can affirm that it is so, because it cannot be otherwise” and: “I without experiment am certain that the effect will follow as I tell you, because it is necessary that it should”. And the same Galileo: “...the fact that all human reasoning must be placed second to direct experience. Hence, they will philosophize better who give assent to propositions that depend upon manifest observations, than they who persist in opinions repugnant to the senses and supported only by probable reasons.” (Galilei 1613: 118). The case of Galileo may definitely confuse those trying to classify him as a rationalist or empiricist. The shortest summary was given by Koyré (1968: 43): “The new science is for him [Galileo] an experimental proof of Platonism”. This complexity causes a problem for “the poor teacher of elementary physics” who seeks a simple picture for his students (Cohen 1950/1993).



A platonic thought experiment is a single thought experiment which destroys an old and existing theory and simultaneously generates a new one: it is a priori in that it is not based on new empirical evidence nor is it merely logically derived from old data; and it is an advance in that the resulting theory is better than the predecessor theory.

The statement of the “simultaneous generation of a new theory” is too strong and inaccurate. The new theory of Newton was, however, more complex and did not approve Galileo’s law conceptually, beyond being a numerically excellent approximation.<sup>14</sup>

Stevin’s “wreath of spheres” hung over an edge (e.g., Dijksterhuis 1986: 326) illustrates the *heuristic* role of TE. The author drew on the obvious results, which resembled “natural experiments” of Aristotelian physics<sup>15</sup> and derived the ratio of weights at equilibrium on the inclined plane, previously known through a rather artificial derivation by Jordanus Nemerorarius (Dijksterhuis 1986: 248–251).<sup>16</sup> However, the theoretical meaning of this account could be understood only within Newton’s theory.

Imaginary performance provides TE with a very special role: to be a stage for theory performance, the “laboratory of the mind”. Free from the constraints of reality (such as heat, friction) TE creates “ideal experiments”, as Einstein used to call TE. The experimenter is allowed to “forget” about all technical limitations of equipment, costs, availability, etc. In a sense, TE presents a kind of mental *modeling* in theoretical physics (Peierls 1980) and is conceptually similar to computer simulation.

In the absence of a theory, TE is in an extremely difficult situation. When Greek natural philosophers debated the nature of the cosmic end, they placed an observer there, and speculated what would happen to the spear thrown across the cosmic edge. Several answers were produced in accordance with different ideas used: wall-like, marshy or cliff-like edge (e.g., Lucretius 1910: 59; Harrison 1981: 104–107).

In contrast, possessing a theory normally guides TE to a certain answer, even when a real experience is impossible. Such were the TEs of the body dropped into the hole through the Earth (the theory of impetus), considering the reality inside a cabin of the ship in uniform motion (the principle of relativity) (Galilei 1632: Second day); Galileo’s scaling of bones, showing the impossibility of giants (Galileo’s theory of statics) (Galilei 1638: 170), Newton’s “the lowest little moon”, showing universality of gravitation (the principle of parsimony and Newton’s theory of gravitation) (Newton 1687/1999: 805).

In its *heuristic* use, TE may anticipate future theoretical principles. In 1659, Huygens described the imaginary experience of a person rotating while attached to a large wheel. He *hypothesized* concerning the equivalence of centrifugal force and gravity (Dugas 1986: 194–197). Huygens’ description of reality in the rotational frame of reference anticipated the principle of equivalence of gravitational and inertial forces introduced much later by Einstein in the theory of general relativity (e.g., Reichenbach 1958: 223).

In 1930, in his debate with Bohr, Einstein suggested a brilliant TE of a clock in a suspended box (Bohr 1949/1959: 225–230). It was refuted by Bohr. Nonetheless this TE was extremely important. Bohr’s refutation drew on Einstein’s theory of relativity. It was

<sup>14</sup> See note 12.

<sup>15</sup> Natural experiments”, or analyses of actual phenomena in terms of a certain theory, as if arranged for testing certain theory, were among the major tools of Aristotelian physics. In our present day natural experiments continue to play a central role for “historical sciences” such as astrophysics, paleontology, geology, etc. (e.g., Diamond 1999: 424).

<sup>16</sup> By considering a chain of spheres suspended on a double edge, Stevin showed that the gliding force on the load placed on the inclined plane is in inverse proportion to the length of the inclined plane (e.g., Mach 1896/1976: 32–41).



thus shown that the *non-relativistic* quantum theory concerning the time–energy uncertainty requires better interpretation. This important result illustrated the power of TE to question the consistency of the theory used and its understanding.

“Schrödinger’s cat” TE, was similarly elucidating. By obtaining the absurd result by the quantum description applied to a cat, assumed in a mixed state: live–dead, it showed with great clarity the fundamental feature of the quantum mechanics, the existence of two types of objects: *macro* objects (cat), for which a mixed state is impossible, and *micro* objects (elementary particle) which allows such states (e.g., Cushing 1998: 311–313).

Perhaps the most representative is the TE suggested by EPR (Einstein et al. 1935), who claimed that quantum description was incomplete since it violated the principle of locality of physical events (e.g., Cushing 1998: 325).<sup>17</sup> Facing the empirical success of quantum mechanics, the authors were arguing for incompleteness of the theory, rather than its inaccuracy. For years this TE “floated in the air” without any chance to evaluate its conclusion. It was only after the *real experiment* by Aspect in 1981, and not before the *theoretical* account by Bell in 1964, that the quantum mechanical description was shown to be correct as it is in describing the microscopic objects subdued to non-locality (e.g., Penrose 1997: 64–66; Cushing 1994: 14–16). The TE of EPR was *correct*, although it produced the result which contradicted the theoretical principles outside the quantum mechanics. We were shown that EPR were *wrong* in their assumptions regarding the nature of the reality in the microworld.

These examples show how far the power of TE may go and support the conclusion that TE obeys certain theory (or theory-like framework), presenting its “view”, and cannot surpass it. If a paradox is found, its solution comes from a new theory.<sup>18</sup> It is in this sense that one can adopt the remark of Koyré that “Good physics is made *a priori*” (1968: 88). This feature was understood by Duhem and Mach who warned that TE may lead to confusion and erroneous inferences (Mach 1896/1976: 146). A *real* experiment is indispensable in analysis and conclusion regarding theory.

#### 4 Multiple Interpretations and Failure

The nature of mental simulation of a theory (views, ideas) may cause an interesting situation of contrasting results produced by the same TE. For instance, to demonstrate the absence of a void Aristotle reasoned imagining a body in motion (1952b):

... no one could say why a thing once set in motion [in void] should stop anywhere; or why should it stop here rather than here? So that a thing will either be at rest or must be moved ad infinitum, unless something more powerful gets in its way.

The conclusion of movement “ad infinitum” was considered to be an obvious absurd, that is, in contradiction with the Aristotle’s *theoretical framework*—a finite universe—in which any rectilinear motion could not be permanent. Today the very same TE, within the new theoretical framework, this same reasoning might serve an argument for the continuous inertial movement.

<sup>17</sup> EPR—Einstein, Podolsky, and Rosen (1935) pointed to the nature of quantum description. Einstein provided the non-formal description of this paradox featuring the quantum description of reality in his *Autobiographical notes* (1949: 83–87).

<sup>18</sup> Brown approached this view when wrote in another place (2004): “The thought experiment establishes a phenomenon; the explanation comes later”. The new theory provides such explanation.

Norton (2004c) defined thought-experiment/anti-thought-experiment pairs which demonstrate a statement and its opposite. One can say even more, that inferences from the same TE may be multiple, reflecting the multiplicity of theoretical approaches applied.<sup>19</sup>

This feature of TE may guide our meaning of TE being “wrong” which does *not* imply production of a physically incorrect result. The latter might testify for a deficiency of the theory applied—a valuable result. For example, Newton (1687/1999: 412–413) in his rotating bucket TE referred to the water surface shape.<sup>20</sup> Newton’s inference was the existence of absolute movement (that is, of the absolute space)—a correct result under the given theoretical assumptions. Mach (1883/1989) pointed to the fallacy in Newton’s assumptions, in Popper’s terms, “certain overlooked possibilities”.<sup>21</sup> The TE itself did not fail and to accuse it in conceptually wrong result is similar to accusing Michelson-Morley experiment of producing a result incompatible with classical mechanics.

In contrast, one can consider the TE suggested by Leibniz (1902/1968) concerning *falling bodies*. Leibniz intended to dismiss the Cartesian “quantity of motion”,  $mv$ , and replace it with  $mv^2$  (*vis viva*) as a more appropriate characteristic of motion. However, Descartes employed his “quantity of motion” to describe its conservation in collisions and never stated conservation of the quantity of motion in falling (he knew the results of Galileo). Therefore the critique of Leibniz was irrelevant. For this reason his TE failed to reach its goal.<sup>22</sup> As was understood later, both quantities energy and quantity of motion (momentum) adequately characterize motion.

Another kind of failure results the erroneous application of the theory (e.g., Norton 2004a).<sup>23</sup> In 1679, Newton suggested a TE which had to demonstrate Earth’s rotation (Schneer 1960: 105–108; Westfall 1977: 148–151). Newton speculated that a body dropped from a high tower would deviate to the east, contrary to Ptolemy’s claim in a similar TE. Newton proceeded and asked about the trajectory of the body if it could continue its fall inside the Earth without any resistance. His answer was a spiral trajectory converging to the center of the Earth. Hooke immediately found two errors in Newton’s result, publicly presented to the Royal Society in 1679 (Arnold 1990: 16–21). Firstly, Newton missed the additional displacement to the South (London is well to the North of the equator). The second error, however, was fundamental. The spiral trajectory was a misconception congenerous to the Aristotelian force-motion paradigm. Hooke stated an

<sup>19</sup> This claim matches McAllister’s (1996) view that the evidential significance of TEs depends on the particular historical environment where history is the history of scientific paradigms. As Brown (2004) put it: “A particular thought experiment might be rightly used one way in one historical situation and wrongly used in another”. Our reservation is only regarding *wrongly*, in favor of *differently*.

<sup>20</sup> Newton’s bucket experiment refers to the appearance of the water surface in a suspended bucket. The dependence of the water surface’s profile on the relative motion of the water and the bucket was considered in four possible relative situations: (1) water is at rest but the bucket is rotating, (2) both are rotating, (3) water is rotating but the bucket is at rest, (4) both are at rest.

<sup>21</sup> In fact, Bishop Berkeley, not acknowledged by Mach, provided the same critique much earlier: “clearly we cannot know the absolute motion of any body” (Berkeley 1721/1965: 270). Newton ignored the possible influence of the Earth and distant celestial bodies on the water in the bucket.

<sup>22</sup> Even if failed in its original intention this TE may be beneficial in science class. In the attractive context of scientific debate Leibniz introduced another important quantity—mechanical energy: *vis viva*—a prototype of kinetic energy and *vis mortua*—a prototype of elastic potential energy (Dugas 1986: 221). The true expression for kinetic energy requires  $\frac{1}{2}$  and the quantity  $m \cdot h$ , missed the factor  $g$ . These deficiencies, however, reflected the limits of the applied qualitative analysis.

<sup>23</sup> Norton (2004) exemplified erroneous application of special relativity in the TE of “rod and slot.”

elliptic trajectory (Kepler's first law), as required by the influence of the gravitational force with the inverse square distance dependence  $1/r^2$  (the claim made before the *Principia*).<sup>24</sup> This was a public humiliation from which Newton never completely recovered, but at the same time, it was this TE that instigated Newton's writing of the *Principia* and providing a full mathematical treatment of the motion under the gravitational force featured by the  $1/r^2$  dependence.

The considered examples demonstrate the meaning of failure concerning TE. It may happen for deficient logic or incorrect application of the theory. The failure of TE does not mean results contradicting reality or another theory not used in the considered TE, similarly to the regular experiment that does not become wrong if it produces the results contradicting the known theory.

## 5 TE Definition

We will consider definition of TE although the difficulty of defining this concept is known (Brown 2004):

It's difficult to say precisely what thought experiments are. Luckily, it's also unimportant.

Indeed, is concept definition important? In science the lesson was taught by Einstein whose revision of mechanics started from the definition of simultaneity, in contrast to Newton who avoided discussion of the *a priori* given concepts of time and space. In the new trend of the philosophy of science—operationalism—concept definition is of central importance and stipulates meaningful conceptual knowledge of physics (Bridgman 1936: 10–11).

In education, however, there are arguments. Those who are against providing definition argue for gradual construction of concepts with use. They often quote Bloom's taxonomy (Bloom et al. 1956) according to which knowledge of definitions is inferior, being of "low" ranking in cognition. This strategy, for instance, is often adopted in teaching energy, causing serious confusion amongst students and teachers (Galili and Lehavi 2006).<sup>25</sup> Definitions are especially important for novice who, unlike practitioners, learn science "from outside". This is unlike experts who construct concept understanding in a long process of its using, that is, "from inside". In their routine work, they may often afford to neglect definitions.

As defined by Mach (1883/1989), a TE represents an experiment, which need not be performed because of its obvious result. Considering the contributions of Stevinus and Archimedes, Mach remarked that there was "no need to perform" some experiments due to their great compelling power based on the "instinctive knowledge" (*ibid.*: 32–39). Later on, however, arguing for the use of TEs in education (1896/1976), Mach expanded his definition to including qualitative reasoning concerning any physical situation in which one makes predictions and solves problems.<sup>26</sup>

<sup>24</sup> Full of irony, this case possesses a great pedagogical potential, showing to the learner that the greatest mind are not immune against misconceptions of the kind students may do on our days.

<sup>25</sup> Another representative example of the importance of concept definition concerns the concept of weight. It has been shown that its definition is essential for phenomena explanation and may influence phenomena understanding (e.g. Galili and Kaplan 1997; Galili 2001; Galili and Lehavi 2006).

<sup>26</sup> An anonymous reviewer pointed out that Mach was at that time promoting his notion of *Physikalische Denkaufgaben* (literally, physical thought problems) in the teaching journal *Zeitschrift für den Physikalischen und Chemischen Unterricht* which he had co-edited with B. Schwal since 1887.

Closely to this definition Lakatos (1976, 1978) defined TE in mathematics as an *intuitive* (lacking solid formal reasoning) consideration showing, however, the truth of certain claims (Lakatos 1978: 65). He pointed to the old tradition of such approach in mathematics, originated in Greek *deykmine*, meaning showing, instead of demonstrating.

Sorensen (1992: 205) defined TE as:

an experiment that purports to achieve its aim without the benefit of execution.

Using *purports to achieve* instead of *achieves* seemingly indicates complexity. An experiment was previously defined by him as (*ibid.*: 186):

a procedure for answering or raising questions about the relationship between variables by varying one (or more) of them and tracking any response by the other or others.

This definition of “experiment”, however, seems to go beyond science. Scientific experiment presumes in addition certain theoretical considerations behind the question and procedure. For example, measuring the water level of the Nile providing it as a function of time, as it is performed for many centuries, answers some questions, but could doubtfully be considered as a scientific experiment.

Reiner and Gilbert (2000) wrote:

thought experiment is design of thought that is intended to test and/or convince others of the validity of a claim.

This definition is too inclusive, and Gendler (2004) refined it with regard to science:

to perform a *scientific* thought experiment is to reason about an imaginary scenario with the aim of confirming or disconfirming some hypothesis or theory *about the physical world* (italics added).

Here are several other characteristics which in a sense define TE too:

It is a special type of mental window through which the mind can grasp universal understandings. (Brown 1991).

Thought experiments are devices of the imagination used to investigate nature (Brown 2002).

Thought experiments are just picturesque argumentation of a hypothetical or counterfactual nature (Norton 2004).

A form of simulative mental based reasoning performed by a scientist (Nersessian 1993).

All these definitions reveal specific and complementary interpretations of TE. Learning from all and following the performed above epistemological analysis one may suggest the following definition:

*Thought experiment* is a set of hypothetico-deductive considerations regarding phenomena in the world of real objects, drawing on a certain theory (\*) (principle or view) that is used as a reference of validity.

The latter definition rules out the speculations about reality not within a scientific theory: supernatural forces, magical powers, and other fantasies. Not to be confused, Maxwell’s demon remains legitimate. He acts in a rational way, attempting to violate the

entropy increase (e.g., d'Arbo 1927/1950: 203; Feynman et al. 1965: 46–5), and it is upon the physicist to demonstrate by theoretical tools the impossibility of such demon and preserve the validity of the law. Furthermore, the latter definition also excludes a mere formal analysis (investigation of equation, dependence on parameters, etc.), manipulating with theoretical entities without addressing the real objects as in real experiment.

After Piaget (1970: 9) who defined schema as possessing a status similar to a theory in the learner's knowledge, one may in accord interpret TEs as performed by students drawing on their schemata. In accordance, TE obtains meaning also with regard to children's thinking and account of nature.

## 6 Establishing the Meaning by Variation of Features

To promote understanding of the concept one need to define it in a certain space of meaning and compare it with the conceptual environment. In fact, this is an old tradition to characterize objects by varying some property against certain scale. Classical Greek science characterized phenomena or processes by opposite features: coldness–heat, dryness–wetness, heaviness–lightness, dense–rare, rough–smooth, hard–soft, etc. (Dijksterhuis 1986: 22–23). For example, the elements were conceived on the scale of heaviness and lightness (levity) which determined the variation (Fig. 1).

Two pairs of the *first qualities* were postulated as fundamental: coldness–heat and dryness–wetness. Within this conceptualization, any real object can be located on the plane of the correspondent axes (Fig. 2).

To apply a similar approach to classification of concepts, one needs to identify their representative dimensions. Naturally, each of the two components of TE—Thought and Experiment—could be interpreted as a variation of a certain activity. In light of the previous discussion, the component of Thought should be related to Theory and so the first axis may denote the variation in the degree of affiliation to *Theory*. The opposite pole of the axis of mental activity could be defined as *Intuition*. One may place on this axis such activities as theoretical analysis, algorithm application, guessing.

The second dimension required for TE is related to *Experimentation*, understood as an organized activity of manipulation with *real objects*. Its opposite pole may denote manipulation with *formalism* (symbolic codification of theory). One may place on this axis such activities as laboratory experiment, modeling, and computer simulation.<sup>27</sup>

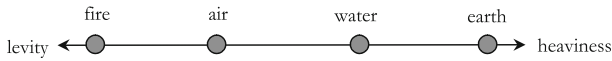
The two axes create a plane (Fig. 3) and thus match Mill's perception who wrote (1892: 99):

It cannot be said that to state the meaning of the word is the purpose of definition. The purpose is not to expound a name but to help to expound a classification.

Indeed, the four quadrants define several activities. Aside from TE we have obtained three constructs: IE—Intuitive Experimentation, IM—Intuitive Manipulation, and TM—Theoretical Manipulation. This conceptual quartet presents the space of variation, which by comparison may clarify the meaning of TE.

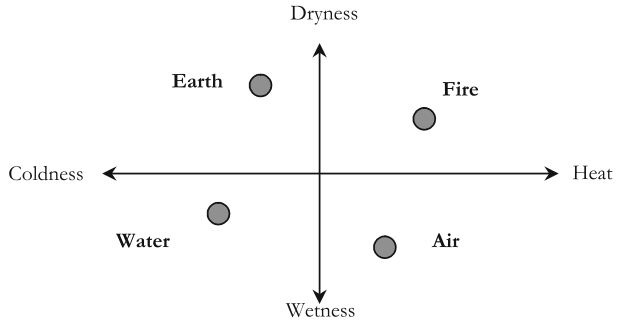
In particular, TM differs from TE in losing the experimental aspect, that is, the reference to manipulation with the *real* objects. Theoretical analysis of a problem is TM, for

<sup>27</sup> The choice of axis is not unique and may cause different classifications (e.g. experiment versus observation).

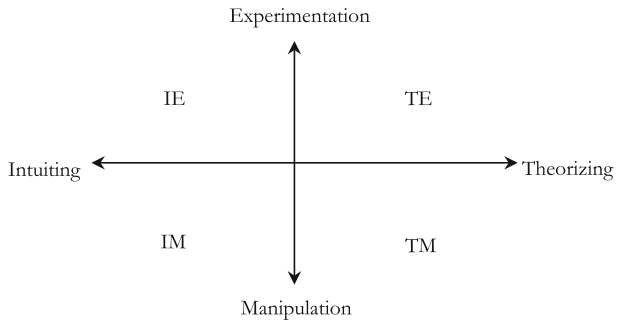


**Fig. 1** Greek conception of heaviness of elements can be presented on the axis of heaviness–lightness (levity) variation

**Fig. 2** Within the Greek science any object or substance could be characterized by its location on the plane of primary qualities. The closeness to the axes reflects the resulting quality of a compound object



**Fig. 3** The conceptual space created by the axes of theorizing–intuiting and experimentation–manipulation defines a cluster of relative activities



example, considering the way of derivation of the electromagnetic waves from Maxwell's equations. Real objects may appear in TM but indirectly.

Furthermore, by definition, IE differs from TE in losing the element of theorizing while preserving experimental character. Sports coaches talk about “muscle memory” when referring to such phenomena. Reiner and Gilbert (2000) used the notion of “body memory” to address the same, that is, the imaginary “manipulation” with certain objects (e.g., imagining throwing a ball into a basket, “distant surgery”, etc). These activities include various manifestations of non-propositional knowledge and accompany fantasies, dreams, etc.; they draw on physical experience. IE may appear in the head of researcher prior or in parallel to theoretical account.

Finally, IM differs from TE in that it lacks both theoretical and experimental reference. This activity may include calculations, solving problems according to a memorized algorithm or browsing one's memory seeking such an algorithm. IM may include analysis of data where one classifies it and/or seeks correlations between various parameters and does not use for that any theory.

Facing this space of variation one can better appreciate the specific nature of TE, which intricately combines theory with experiment. Moreover, if the difference between TE and IM is obvious since they are opposites in both senses, the distinguishing between TM and TE (as well as IE and TE) is often perplexing. A theoretical analysis of a phenomenon

(TM) and a reference to intuition in considering a situation (IE) are often confused with TE. Intuitively, physicists resist equating the TEs of Einstein in his debate with Bohr, with the effort of estimation made by a basketball player before throwing the ball. The suggested classification reflects this intuition.

The suggested classification helps to value the introduced definition (\*). Indeed, TE was defined by two *definiens*: (1) a hypothetico-deductive reasoning based on theory and (2) mental manipulation with physical objects. The former presents the *genus* of TE and the latter—its *differentia* (Copy 1972: 118, 134). By dropping the genus one may move from TE to IE (intuitive experiment), and dropping the differentia will transfer the construct from TE to TM (theory manipulation), and by dropping both of them one arrives at IM (intuitive manipulation). In isolation, without comparison between these constructs, novice may term each of them as a TE.

## 7 Illustration of Classification

To illustrate these introduced constructs in an educational context, let us consider a common problem from school physics: a projectile thrown at an angle to the horizon. Student need to find the trajectory and the maximal distance covered by the projectile. Four approaches could be imagined in accordance with the four constructs introduced.

1. TM. This approach requires applying Newtonian dynamics as learned in the class. A hypothetical student could think:

OK. This is a two-dimensional situation. The only force exerting on the thrown ball is the vertical gravitational force,  $mg$ . So, I introduce vertical and horizontal axes, decompose the motion, and use the Newton's Second Law for both axes. Then I solve the equations and get the equations of motion for both axes. Combining them I obtain the trajectory. From the trajectory I will get an expression for the distance covered by the ball as this corresponds to the intersection point with the horizontal axis. I will analyze this expression for its maximum and get the conditions for the greatest distance possible by the thrown ball.

This mental design represents TM. Some students could imagine the equations of motion and even solve them without writing. This kind of performance would fit the common requirement of a physics class.

2. TE. In this approach the student's thinking takes another, more qualitative path:

OK, I need to throw a ball as far as possible (the greatest distance). What does it mean? I need to give the ball the greatest speed possible; that is obvious. But the initial speed is already given... Well, the ball needs the maximal horizontal speed to cover the greatest distance. At the same time, it needs the greatest time to fly before it falls. Suppose I throw it at a steep angle; it will rise higher but will not get that far because of a small horizontal speed, and if I throw it on a shallow path it still will not get far because the flight will be short ... How can I achieve both goals? Well, I can throw the ball at half the angle between vertical and horizontal. This way I will equally satisfy both requirements of the flight time and the horizontal speed. This means, the angle should be  $45^\circ$ .



Now, what will its path look like? There are two motions of the ball. Gravity permanently pushes it down, so it is moving with constant acceleration all the way. At the same time, it also moves horizontally by inertia (no force there). OK, it should be a kind of a trajectory curved towards the ground. And the faster the ball goes, the less its path is curved by the gravity, so at the top where the ball stops going up, it will be most curved ... And ... the ball moves up exactly as it falls down (the same equation...), then the trajectory will be sort of symmetrical... Well... it should be like a parabola...

This thinking presents a qualitative solution drawing on theory, as Mach would probably suggest. The approach is heuristic; it uses theory, but not as a given procedure; rather in a hypothetico-deductive analysis, addressing a real object.

3. IE. In this case, thinking about the given situation, the student does not seek any theoretical knowledge. Trying to tackle the problem, the student refers rather intuitively to his experience accumulated in similar situations. This knowledge takes a non-propositional form, “body knowledge” (Reiner and Gilbert 2000), the sense data of trial and error experience accumulated in throwing objects. This knowledge suggests that the angle of throwing the ball to reach the greatest distance is close to the middle of the right angle. However, an experienced coach in athletics may refine this result, suggesting a slightly different angle which in fact takes into account the air resistance and is better to reach more distance in throwing the iron ball, for example. Another advice is provided by a football coach. He takes into account the air circulation around the spinning ball and suggests still another angle. A baseball coach may suggest yet another angle. All these results are known from practice without any theory. All these coaches made IEs.
4. IM. Here we should imagine the hypothetical student, who has no theory to apply, but might have sufficient relevant and reliable data. The data become the subject of parametric analysis, a manipulation seeking to make sense of the data:

OK, let’s try to make sense of it. Let’s vary the speed.... here are the results; there is a clear dependence. Let’s vary the angle ... Now, what else can influence it... The pull of gravity must be important. Let’s see, what was observed concerning the bodies thrown on the Moon.... OK, it also seems to influence in the inverse way... It is reasonable... Air: the resistance of the air should shorten the distance..” And so on.

Eventually, the student may arrive at a certain account of the projectile motion and realize the kind of trajectory produced and the condition for the greatest distance covered. As in IE the result will be not necessary a parabola. Look at the drawings of cannon shells by Leonardo (Clavi 1956: 281). His ingenious perception afforded a curve resembling parabola but smashed a little in its second half due to air resistance, a slightly asymmetrical trajectory. To solve the real situation theoretically could be a difficult problem; too many factors might affect the result in reality. In fact, however, common situations in physical laboratories are often quite similar: investigation of a “black box” of data without a clue to the theory which could guide.<sup>28</sup> These are IMs.

<sup>28</sup> In the course of IM the researcher may render a theoretical model, replacing a theory. This would change the activity to TM.

Of course, these four examples serve only as tentative and schematic illustrations, trying to provide a general idea of distinguishing between the different approaches that are often presented as TE. In reality, a mixture of any of the four types of activities can of course occur. This, however, does not remove the benefit of conceptual classification, which contributes to better understanding of the nature of possible cognitive strategies in making physics and identifying them in educational research.

## 8 Typology of Uses

In addressing the *structure* of TEs Brown (1991: 33–48) suggested their *destructive* and *constructive types*. In contrast, this paper suggests that one should rather address the *uses* of TEs. Variation approach provides a different path to render a reasonable classification. There could be a variety of goals in using TE: research, education, popularization of science. One representative dimension could be the extent to which the TE draws on a theory, as opposed to drawing on intuition, appearance, experience, common sense, or a naive view. Another dimension could be the goal of the use, which may vary from *falsifying* to *confirming*. These dimensions create the required plane of variation (Fig. 4).

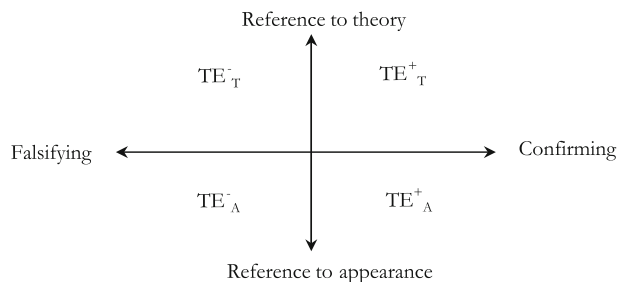
The four quadrants in this plane produce possible cases. The uses located in the lower half of this plane ( $TE_{\bar{A}}$ ,  $TE_{\bar{A}}^{+}$ —reference to appearance) are more common in *introductory* contexts. For example,  $TE_{\bar{A}}^{+}$  uses often include dropping a body from the mast of a moving ship, Newton's bucket and Stevin's chain, all already discussed in this paper.

The use of  $TE_{\bar{A}}$  also draws on appearance but in this use the TE also challenges and criticizes certain theoretical views. For example, this way Buridan-Oresme questioned the stationary Earth by imagining a rotating observer; Benedetti-Galileo and Buridan-Galileo falsified Aristotelian concepts of falling bodies; Leibniz sought falsifying the Cartesian quantity-of-motion while considering falling; Buridan, arguing by an arrow, falsified the Aristotelian self-pushing mechanism of a projectile.

Both  $TE_{\bar{A}}$  and  $TE_{\bar{A}}^{+}$  draw on the appearance and so match the original definition of Mach who required obviousness of the results. Appearance is often persuasive when it comes to intuition, experience and common sense. Although pedagogical considerations may suggest such a use, it, however, remains on a shaky ground in physics. The teacher may explain the origin of the limited validity of appearance and point to the required theoretical fortification which may come later in the course.

The upper-half of this plane includes the uses of TE with reference to theory:  $TE_{\bar{T}}$  and  $TE_{\bar{T}}^{+}$ . Their performance and results may be not obvious.  $TE_{\bar{T}}^{+}$  matches Popper's *apologetic* and *heuristic* uses and  $TE_{\bar{T}}$  corresponds to Popper's *critical* use, challenging the theory and possibly falsifying it.

**Fig. 4** Typology of uses of TE



$TE_T^+$  may reveal certain non-trivial and hidden possibilities embedded in the given theory. For example, such use is often given by a teacher to the “twins’ paradox” in the theory of special relativity: a person making a round trip in a spaceship moving with high speed remains younger than her twin stayed at home (e.g. Taylor and Wheeler 1997: 125–126; Park 1988: 297–300). The same use is given to Einstein’s TE demonstrating mass–energy relationship (Einstein 1905/1952b: 69–71; Born 1924/1965: 283–286), Schrödinger’s cat demonstrating the dichotomy of micro and macro objects (Cushing 1994: 38–39), Einstein’s elevator introducing the equivalence principle (Einstein and Infeld 1938: 226–235), Maxwell’s demon questioning the entropy growth principle (Feynman et al. 1965: 46–5). All these TEs by their anti-intuitive results promote understanding of the sophisticated aspects of particular physical theories. All of them are consistent with the correspondent theory and therefore illustrate *apologetic* and *heuristic* uses of TE.

In contrast,  $TE_T^-$  points to a contradiction, inconsistency, or incompatibility within certain theory. Examples of this use, common at physics lectures, are: Einstein’s concerning a magnet and conductor in relative motion (revealing the problem of non-relativistic electromagnetism), Einstein’s chasing the beam of light (revealing the problem in non relativistic optics), EPR (revealing the problem of quantum mechanics with the principle of locality), Einstein’s clock in the box (challenging non-relativistic quantum mechanics), entanglement TEs (which challenge the principle of causality in quantum mechanics). All these uses are commonly employed by lecturers intending to articulate conceptual problems of the particular theory. Such using of TE is often sophisticated and helpful in revealing the intricate nature of physical theories, especially its unresolved and not well understood aspects.

It is important to realize that the same TE could be used in different purposes of disciplinary and/or social nature. What once used to serve scientists as a tool in their debates can now serve as an educational tool. The same TE could be considered as critical with regard to one theory while at the same time it might confirm another. For example, the famous TE by Poisson, predicting in 1818 a bright spot in the middle of the shadow area behind a circular screen (e.g., Hecht 1996: 1041), was originally suggested in a critical heuristic use ( $TE_T^-$ ) to discredit the wave theory of light. Immediately after, it was performed as a real experiment by Arago. Since then it is used as  $TE_T^+$  for a *confirmation* of the wave theory of light, as well as  $TE_T^-$  for *falsification* of Newton’s particle theory of light.

This perspective matches Popper’s approach (1934/1968) and suggests that we may not consider the unique classification of a TE ascribing it to a particular *category* (confirming or criticizing type). One may rather talk about certain *use* of this TE in accordance with its role in a specific social or disciplinary context. In different contexts this role may change.<sup>29</sup>

## 9 Implications for Teaching Physics

Following our discussion we are in a position to assert concerning the potential merits of using TE for educational purposes.

Firstly, as was mentioned, TEs and especially those belonging to the history of physics, often point to the essential features of physical theories. For example, Einstein and Infeld incorporated many “idealized experiments” in their famous book (1938), introducing

<sup>29</sup> The feature of variation of the roles played by a particular TE can encourage the genuine understanding of the considered theory. Learning through variation was stated to be an effective teaching strategy (Schecker and Niedderer 1996; Marton and Tsui 2004).

central concepts of physics to the novice learner. This tradition is kept by numerous popular presentations of physics in which TEs reveal the core conceptions of physical theory. TE often makes it through simplified but representative models which keep the focus on the essential aspects of the subject, eliminating technical details, experimental errors and ruling out the impeding factors of a real experiment (heat, friction, etc.). Inclusion of TE does not exclude real experiments but develops in students an ability to appreciate real experiments, to see their rationale despite many non-relevant details which distract the attention of naïve observers. As in a real laboratory, TEs present a necessary prelude to any experimental activity. Skipping over TE and going straight to real experiments often deprives the latter of meaning and value for the learners.

Secondly, the use of TE appeals to imagination of the learners and allows considering situations impossible to reproduce regardless the sophistication of the equipment. After Newton, every physics student launches his satellite by imagining a stone projected horizontally from a mountain with successively increasing velocity (e.g., Newton 1687/1997: 5–7; Ohanian 1989: 223–224). By addressing situation impossible to reproduce TE becomes not only an indispensable tool of teaching (as well as of scientific research), but also introduces into instruction the same elements which so strongly attract young minds to the bizarre fantasies of science fiction. Such are TEs about the tunnel through the Earth, Galileo's giants, Newton's small Moon, and many others.

TE is especially indispensable in the presentation of modern physics: the theory of relativity and quantum mechanics, where real experiments are practically excluded from regular classroom activity and multimedia tools very often fail, suffering from superficial and conceptually irrelevant contents. One should emphasize to the students that historically heuristic TEs guided the construction of both the theory of relativity and the quantum mechanics. The real experiments remained of crucial importance, but in many cases they came only later on, when it was clear to the researcher what to expect.

Heisenberg's microscope, Schrödinger's cat, two slit interference for single particle, length contraction and time dilation, quantum entanglement are normally presented exclusively with TEs at lectures and in books (e.g. Penrose 1997; Hobson 2003). TE opens a unique window to the strange and unknown world of super-large and super-small scales, distant and unfamiliar, engaging the imagination of the youth whose curiosity concerning the organization of nature is often suppressed, if not totally destroyed by the overwhelming and often simply merciless attack of lecturing the dry formalism, abstract and impenetrable without strong ideology and motivation that could be provided in advance by a simple and inspiring TE.

Thirdly, TEs introduce students into the *culture of science*, create its authentic image. Many of the mentioned above TEs are associated with the ethos of science, display its goals, spirit, values and tradition. TE introduces debate, argumentation, the struggle of ideas, inviting the novice to join in and share the intellectual enjoyment already at the early stages of acquaintance with the disciplinary knowledge. Without formalism, or with a very minimum of it, TEs are able to ignite curiosity within the general public, exposing and connecting the learners to the most challenging and basic scientific issues. As Mach put it (1896/1976: 146):

It is often said that enquiry cannot be taught. In a sense this is correct ... for intellectual situations never repeat themselves. However, the examples of great enquiries are very suggestive, and practicing thought experiments after their model ... is bound to be beneficial.

Finally, TEs allow revealing the conceptions held by individuals on the relevant scientific issues. The hypothetico-deductive reasoning associated with TE encourages individuals to articulate their conceptions so valuable for the good teacher who is going to present a particular scientific subject. This feature makes TE a powerful tool for investigating of students' knowledge and design of stimulating instructional tools of a constructivist type (e.g., Clement 1983; Camp and Clement 1994). The cluster of categories suggested in this paper, Thought Experiment, Intuitive Experiment, Theoretical Manipulation and Intuitive Manipulation may serve a range of purposes in educational research and be useful in teaching and assessment of learning and knowledge.

## 10 Conclusion

This paper considered TE as a special theoretical construct, a scientific tool widely used for mediating between theory and experiment. However, TE is much more than a bridge between them. TE is renowned for its function as a tool for developing, clarifying and critiquing of theoretical conceptions.

Nonetheless, although extremely economical, attractive and effective in presenting theoretical contents, TE, at the same time, possesses potential for confusing. Thus it is important to elaborate on the epistemology of TE as a tool of mental simulation of a theory, argumentation and modeling. Addressing the epistemological status of TE might prevent the common misconceptions with regard to the nature of science, clarify the sense in which TE can produce the new knowledge about nature and the intrinsic limitations that TE possesses in this direction due to its inability to surpass the theory within which it was constructed.

It was suggested here that the meaning of TE, as a concept, can be appreciated using a two-dimensional conceptual variation. This approach implies comparison among the conceptual quartet of *Thought Experiment*, *Intuitive Experiment*, *Theoretical Manipulation* and *Intuitive Manipulation*—a cluster of congenerous activities possessing certain similarities, but essentially different. Using TE could be classified by variation of reference to theory versus intuition, on the one hand, and by stating its confirming or falsifying goal, on the other. The same TE may be used in a variety, even opposite purposes.

TEs have provided a crucial contribution to the scientific progress and established its ethos, playing the first violin in the science orchestra. TEs represent the charming and unique nature of scientific activity of which scientists are deeply proud. As such TE deserves to be explicitly introduced into science curricula to promote the creation of a fascinating image of science, attractive in a general intellectual sense.

## References

- Aristotle (C4 BC/1952a) On the heavens, in the works of Aristotle. Encyclopedia Britannica Great Books, Chicago, Bk. II, Chap. 14, 297a, 297b
- Aristotle (C4 BC/1952b) Physics, in the Works of Aristotle. Encyclopedia Britannica Great Books, Chicago, Bk. IV, Chap. 8, 215a
- Arnold VI (1990) Huygens and Barrow, Newton and Hooke. Birkhauser Verlag, Basel
- Arthur R (1999) On thought experiments as a priori science. *Int Stud Philos Sci* 13(3):215–229
- Berkeley M (1721/1965) De Motu. In: Armstrong DM (ed) Berkeley's philosophical writings. Macmillan, New York, pp 251–273
- Bloom B, Englehart M, Furst E, Hill W, Krathwohl D (1956) Taxonomy of educational objectives: the classification of educational goals. Handbook I: cognitive domain. Longmans Green, Toronto

- Bohr N (1949/1959) Discussion with Einstein on epistemological problems in atomic physics. In: Schilpp PA (ed) *Albert Einstein: philosopher-scientist*. Harper, New York, pp 201–241
- Born M (1924/1965) *Einstein's theory of relativity*. Dover, New York
- Born M (1944) *Experiment and theory in physics*. Cambridge University Press, Cambridge
- Bridgman PW (1936) *The nature of physical theory*. Dover, New York
- Brown J (1991) *The laboratory of the mind. Thought experiments in the natural sciences*. Routledge, London
- Brown J (2002) Thought experiments. In: *Stanford encyclopedia of philosophy*. The Metaphysics Research Lab, Stanford University
- Brown JR (2004) Peeking into Plato's Heaven, *Philos Sci* 71:1126–1138
- Buridan J (1357/1959) Questions on the eight books of the physics of Aristotle. In: Clagett M (ed) *The science of mechanics in the middle ages*. The University of Wisconsin Press, Madison
- Calvi I (1956) Military engineering and arms. In: Leonardo da Vinci. Reynal, New York, pp 275–305
- Camp C, Clement J (1994) Preconceptions in mechanics. Kendall/Hunt, Dubuque
- Clagett M (1959) The science of mechanics in the middle ages. The University of Wisconsin Press, London
- Clement J (1983) A conceptual model discussed by Galileo and used intuitively by physics students. In: Gentner D, Stevens AL (eds) *Mental models*. Lawrence Erlbaum Associates, Hillsdale, pp 325–340
- Cohen IB (1993) A sense of history in science. *Sci & Educ* 2(3):251–277 [1950, *Am J Phys* 18:343–359]
- Cohen MR, Drabkin IE (1966) *A source book of greek science*. Harvard University Press, Cambridge
- Copy IM (1972) *Introduction to logic*. Macmillan, New York
- Crombey AC (1959) *Medieval and early modern science*. Doubleday, Garden City
- Cushing J (1994) *Quantum mechanics: historical contingency and the Copenhagen hegemony*. University of Chicago Press, Chicago
- Cushing J (1998) *Philosophical concepts in physics*. Cambridge University Press, Cambridge
- d'Arbo A (1927/1950) *The evolution of scientific thought: from Newton to Einstein*. Dover, New York
- Diamond J (1999) *Guns, germs and steel*. Norton, New York, pp 424–425
- Dijksterhuis EJ (1986) *The mechanization of the world picture, Pythagoras to Newton*. Princeton University Press, Princeton, pp 269–270
- Draper JW (1890) *History of the conflict between religion and science*. Kegan Paul, Trench Trubner, London
- Dugas R (1986) *A history of mechanics*. Dover, New York
- Einstein A (1905/1952a) On the electrodynamics of moving bodies. In: *The principle of relativity, a collection of originals papers on the special and general theory of relativity*. Dover, New York
- Einstein A (1905/1952b) Does the inertia of a body depend upon its energy-content? In: *The principle of relativity, a collection of originals papers on the special and general theory of relativity*. Dover, New York
- Einstein A (1949/1979) Autobiographical notes. In: Schilpp PA (ed) *Albert Einstein: philosopher-scientist*. Harper, New York, pp 3–95
- Einstein A, Infeld L (1938) *The evolution of physics*. Cambridge University Press, Cambridge
- Einstein A, Podolsky B, Rosen N (1935) Can quantum-mechanical description of physical reality be considered complete? *Phys Rev* 47:777–780
- Feather N (1959) *Mass, length and time*. Watson & Viney, Aylesbury
- Feynman R, Leighton R, Sands M (1965) *The Feynman lectures on physics*. Addison Wesley, Reading
- Galilei G (1613/1957) Letters on sunspots. In: Drake S (ed) *Discoveries and opinions of Galileo*. Doubleday, New York
- Galilei G (1632/1967) *Dialogue concerning the two chief world systems – Ptolemaic and Copernican [Dialogo]*. University of California Press, Berkeley
- Galilei G (1638/1914) *Dialogue concerning two new sciences [Discorsi]*. Dover, New York
- Galili I (2001) Weight versus gravitational force: historical and educational perspectives. *Int J Sci Educ* 23(10):1073–1093
- Galili I, Kaplan D (1997) Students' operation with the concept of weight. *Sci Educ* 80(4):457–487
- Galili I, Lehavi Y (2006) Definitions of physical concepts: a study of physics teachers' knowledge and views. *Int J Sci Educ* 28(5):521–541
- Gendler TS (2004) Thought experiments rethought – and reperceived. *Philos Sci* 71:1152–1163
- Giere RN (1984) *Understanding scientific reasoning*. Holt, Rinehart, & Winston, Chicago
- Giere RN (1999) *Science without laws*. The University of Chicago Press, Chicago
- Gilbert J, Reiner M. (2000) Thought experiments in science education: potential and current realization. *Int J Sci Educ* 22(3):265–283
- Gliozzi M (1962) *Storia della fisica. Il volume della Storia della Scienza*. UTET, Torino
- Grant E (1977) *Physical science in the middle ages*. Cambridge University Press, Cambridge
- Harrison ER (1981) *Cosmology. The science of the universe*. Cambridge University Press, Cambridge

- Hecht E (1996) *Physics*. Brooks/Cole, Pacific Grove
- Helm H, Gilbert J. (1985) Thought experiments and physics education – Part 1. *Phys Educ* 20:124–131
- Helm H, Gilbert J, Watts DM (1985) Thought experiments and physics education – Part 2. *Phys Educ* 20:211–217
- Hobson A (2003) *Physics. Concepts and connections*. Pearson, Upper Saddle River
- King P (1991) Mediaeval thought-experiments: the methodology of mediaeval science. In: Massey G, Horowitz T (eds) *Thought-experiments in science and philosophy*. Rowman & Littlefield, pp 43–64
- Koyré A (1943) Galileo and the scientific revolution of the seventeenth century. *Philos Rev* 52(4):333–348
- Koyré A (1968) *Metaphysics and measurement*. Chapman & Hall, London
- Kuhn T (1977) A function of thought experiments. In: Kuhn T (ed) *The essential tension selected studies in scientific tradition and change*. The University of Chicago Press, Chicago, pp 240–265
- Lakatos I (1976) *Proofs and refutations. The logic of mathematical discovery*. Cambridge University Press, Cambridge
- Lakatos I (1978) *Mathematics, science and epistemology. Philosophical papers 2*. Cambridge University Press, Cambridge
- Lattery M (2001) Thought experiments in physics education: a simple and practical example. *Sci Educ* 10(5):485–492
- Leibniz GW (1902/1968) *Discourse on metaphysics*. Open Court, La Salle, pp 29–32
- Losee J (1993) *A historical introduction to the philosophy of science*. Oxford University Press, New York
- Lucretius (1910) *On the nature of things*. Oxford University Press, Oxford
- Mach E (1883/1989) *The science of mechanics: a critical and historical account of its development*. The Open Court, La Salle
- Mach E (1896/1976) On thought experiment in E. Mach. In: *Knowledge and error*. Reidel, Dordrecht, pp 134–147
- Marion F, Tsui ABM (2004) *Classroom discourse and the space learning*. Mahwah, NJ
- Mason SF (1962) *A history of the sciences*. Collier Books, New York
- Matthews M (1994) Thought experiments. In: Matthews M, *Science teaching: the role of history and philosophy of science*. Routledge, New York, pp 99–105
- McAllister J (1996) The evidential significance of thought experiments in science. *Stud Hist Philos Sci* 27(2):233–250
- Mill JS (1892) *A system of logic ratiocinative and inductive. A connected view of the principles of evidence and the methods of scientific investigation*. Routledge, London
- Miller A (1986) *Imagery in scientific thought*. The MIT Press, Cambridge
- Moody EA (1994) Galileo and Avempace: dynamics of the leaning tower experiment. In: Wiener PP, Noland AN (eds) *Roots of scientific thought. A cultural perspective*. Basic Books Publishers, New York, pp 176–206
- Moss JD (1993) *Novelties in the heavens. Rhetoric and science in the Copernican controversy*. The University of Chicago Press, Chicago
- Murdoch JE, Sylla ED (1978) The science of motion. In: Lindberg DC (ed) *Science in the middle ages*. The University of Chicago Press, Chicago, pp 206–264
- Nersessian N (1993) In the theoretician's laboratory: thought experimenting as mental modeling. In: Hull D, Forbes M, Okruhlik K (eds) *PSA 1992, vol 2. Philosophy of Science Association*, East Lansing
- Newton I (1687/1999) *Mathematical principles of natural philosophy*. University of California Press, Berkeley
- Newton I (1728/1997) *A treatise of the system of the world*. Dover, New York
- Norton J (1991) Thought experiments in Einstein's work. In: Horowitz T, Massey GJ (eds) *Thought experiments in science and philosophy*. Rowman & Littlefield, Savage
- Norton J (2004a) Einstein's investigations of Galilean covariant electrodynamics prior to 1905. *Arch Hist Exact Sci* 59:45–105
- Norton J (2004b) Why thought experiments do not transcend empiricism. In: Hitchcock C (ed) *Contemporary debates in the philosophy of science*. Blackwell, London, pp 44–66
- Norton JD (2004c) On thought experiments: is there more to the argument? *Philos Sci* 71:1139–1151
- Ohanian HC (1989) *Physics*. Norton, New York
- Park D (1988) *The How and Why*. Princeton University Press, Princeton
- Pedersen O, Pihl M (1974) *Early physics and astronomy*. McDonald & Janes, London
- Peierls R. (1980) Model-making in physics. *Contemp Phys* 21:3–17
- Penrose R (1997) *The large the small and the human mind*. Cambridge University Press, Cambridge
- Piaget J (1970) *The science of education and the psychology of the child*. Grossman, New York
- Popper K (1934/1968) On the use and misuse of imaginary experiments, especially in quantum theory. In: Popper K (ed) *The logic of scientific discovery*. Harper & Row, New York, pp 442–456



- Ptolemy C (1952) *The Almagest*. Encyclopedia Britannica, Chicago
- Randall JH (1957) The place of Leonardo Da Vinci in the emergence of modern science. In: Wiener PP, Noland A (eds) *Roots of scientific thought*. Basic Books, New York
- Reichenbach H (1927/1958) *The philosophy of space and time*. Dover, New York
- Reiner M, Burko L. (2003) On the limitations of thought experiments in physics and the consequences for physics education. *Sci & Educ* 12:365–385
- Reiner M, M Gilbert J (2000) Epistemological resources for thought experimentation in science education. *Int J Sci Educ* 22(5):489–506
- Russo L (2004) *The forgotten revolution: how science was born in 300 B.C. and why it had to be reborn*. Springer, Berlin
- Schecker H, Niedderer H (1996) Contrastive teaching: a strategy to promote qualitative conceptual understanding of science. In: Treagust D, Duit R, Fraser B (eds) *Improving teaching and learning in science and mathematics*. Teacher College Press, New York, pp 141–151
- Schneer CJ (1960) *The evolution of physical science*. Grove Press, New York
- Shamos MH (1959) *Great experiments in physics. Firsthand accounts from Galileo to Einstein*. Dover, New York
- Shapin S (1996) *The scientific revolution*. The University of Chicago Press, Chicago
- Sorensen R (1992) *Thought experiments*. Oxford University Press, New York
- Stinner A (1990) Philosophy, thought experiments and large context problems in the secondary Physics course. *Int J Sci Educ* 12(3):244–257
- Stinner A (2005) Thought experiments to teach the ideas of Einstein's theory of relativity. In: Paper presented in the eighth international history, philosophy and science teaching conference, University of Leeds, Book of Abstracts, p 89
- Taylor EF, Wheeler JA (1997) *Spacetime physics*. Freeman, New York, pp 125–126
- Tseitlin M, Galili I (2005) Teaching physics in looking for itself: from a physics-discipline to a physics-culture. *Sci & Educ* 14(3–5):235–261
- Westfall RS (1977) *The construction of modern science*. Cambridge University Press, Cambridge

### Author Biography

**Igal Galili** is professor of science education at the Hebrew University of Jerusalem, Israel. Educated in physics, he turned to the area of physics education where his research addresses students' knowledge of physics and its structure, the nature of physics concepts to be taught, physics knowledge structure and the ways of its representation in teaching. This orientation implies addressing the history and philosophy of science, both by teachers and students, as providing conceptual framework of the meaningful and cultural knowledge of the subject. Within this effort, a special framework of discipline-culture was developed and suggested for teaching science. The same framework was used to explain students' conceptual change, the structure of science curriculum, as well as of scientific revolutions.