Post-print of:

Buzzoni, Marco. "Are there Mathematical Thought Experiments?." Axiomathes (2022): 1-16.
https://doi.org/10.1007/s10516-022-09641-3

# Are there Mathematical Thought Experiments?

Marco Buzzoni[1]

Abstract. With reference to an already existing and relatively widespread use of the expression in question, “mathematical thought experiments” (“TEs”) involve mathematical reasoning in which visualisation plays a relatively more important role. But to ensure an unambiguous and consistent use of the term, certain conditions have to be met: 1) Contrary to what has happened so far in the literature, the distinction between logical-formal thinking and experimental-operational thinking must not be ignored; 2) The separation between the context of discovery and the context of justification is to be rejected, at least in one of the main senses in which it was defended by the logical empiricists and Popper (this excludes any position which, ascribing to mathematical TEs only a heuristic role, regards them an intermediate step to attain more traditional forms of rigour); 3) The distinction between mathematical TEs and formal proofs must be regarded as one of degree, and not as a qualitative one, although this distinction may be used in a de facto way for particular or local purposes.

Key words. Thought experiments; Mathematics, proof; rigour; diagrams; James R. Brown; John Norton.

## 1. Introduction

The literature concerning though experiments (hereafter TEs) has almost always assumed without discussion that one can speak of mathematical TEs, in a sense fundamentally similar to that of TEs in

---

1 Dipartimento di Studi Umanistici, Università di Macerata (Italy), e-mail: marco.buzzoni@unimc.it. This paper has greatly profited from conferences and discussions that took place within the FFIUM Project (Formalism, Formalization, Intuition and Understanding in Mathematics). Many thanks to Gerhard Heinzmann, the principal investigator, who made possible a research stay at the “Archives Henri Poincaré” of the Université de Nancy-Lorraine (15.1.2019-15.2.1919). Ongoing discussions with Gerhard Heinzmann led to a change in my preceding position about mathematical TEs (Buzzoni 2011). But I would like to thank also the entire research group of the FFIUM Project both for conferences held by, and for many helpful discussions with, other members of the Project. Many thanks also to Mike Stuart, who has read an earlier draft of this paper and has made many valuable suggestions for its improvement.

the empirical sciences. The first to do so was Mach himself, who in his article "Gedankenexperiment" wrote:

> "We can hardly doubt that thought experiments are important not only in physics but in every field, even in mathematics, where the uninitiated might least expect it. Euler's method of enquiry whose fruitfulness has far outpaced critical assessment, gives the impression of the procedure of an experimenter who is probing a new field for the first time. Even where the exposition of a science is purely deductive, we must not be deceived by the form. We are dealing with thought experiments, after the result had become completely known and familiar to the author. Every explanation, proof and deduction is the outcome of this process."[0]

In this way, Mach anticipated the two theses most prevalent in the critical literature to this day: firstly, that we may assume without discussion that there is an essential similarity between empirical-scientific and mathematical TEs, and, secondly, that the practice of the thought experiment is essentially confined to the moment of mathematical *discovery*, being then destined to be replaced by formal, rigorous demonstrations in the context of *justification*.

Now, one may concede that any similarity between mathematical and empirical-scientific TEs is a first act of understanding, nevertheless one must also ask what are the consequences of neglecting such an epistemologically important distinction as that between mathematical and empirical-experimental knowledge. For this reason, the question concerning the distinctive traits of mathematical TEs remains an important desideratum in today's debate. Moreover, for those who reject (at least in an important sense) the neopositivist and Popperian separation between discovery and justification, the task also arises of explaining in what sense TEs in mathematics can play a role that is not merely heuristic, but also justificatory.

This paper addresses both issues. It is organized as follows. First of all, in Section 2, I shall give a very brief overview of what I consider to be at least one of the specific features of TEs in the empirical sciences and what I consider the most epistemologically important difference between empirical-experimental and mathematical reasoning. It was this difference that led me in the past to the claim that it would be misleading and deceptive to speak of TEs in mathematics (Buzzoni 2011). This conclusion, however, will be

0 Mach 1906, pp. 197-198; Engl. Transl., p. 144. The literature in recent decades has paid little attention to TEs in mathematics and, with rare exceptions, still less to the comparison between empirical and mathematical TEs. A partial exception is the indirect, though fundamental treatment of TEs in mathematics to be found in Lakatos's *Proofs and Refutations* (Lakatos 1963-4), which received relatively more attention (see for example Yuxin 1990; Koetsier 1991; Larvor 1999; Glas 2001a and Glas 2001b; Kühne 2005, pp. 356-366; Sherry 2006; Shaffer 2015; Hertogh 2021. More generally, on mathematical TEs, see above all Witt-Hansen 1976; Mueller 1969; Brown 1999, 2004, 2007, 2011, 2017, and 2021; Glas 2001a and 2001b, Van Bendegem 2003; Buzzoni 2004, 2008, 2011, 2021a and 2021b; Sherry 2006; Starikova 2007; Cohnitz 2008; Starikowa and Giaquinto 2018, Brown 2022, Norton & Parker 2022, Lenhard 2022, Fehige and Vestrucci 2022.

rejected because of a simple but decisive objection: it is a fact that many authors, among them some mathematicians, have spoken about TEs in mathematics, illustrating this meaning of TE with concrete examples. This, however, in spite of undermining my earlier conclusion, does not force one to disavow all the reasons that led to it. Some of these reasons in themselves remain correct, and as such will be preserved and reaffirmed here in Section 3. But they will be used not to deny the existence of TEs in mathematics or to propose abstract rules for their use, but rather to clarify some conditions of possibility of a consistent use of the term in mathematics, to distinguish it from its use in the empirical-scientific sphere, and to guarantee its coherence with respect to certain fundamental philosophical choices, first and foremost the rejection of the separation between the context of discovery and the context of justification. More specifically, it seems to be advisable to regard mathematical TEs not only as a specific kind of TE, in principle different from empirical-scientific TEs, but also as a subset of mathematical demonstrations in which visualisation and diagrams play a relatively more important role, thus avoiding too great a divergence between epistemological reflections and how mathematicians understand their own practice. However, this way of speaking about TEs in mathematics can be defended only by taking into account three crucial philosophical caveats. First, as already mentioned, the distinction between logical-formal thinking and experimental-operational thinking must not be neglected or underestimated. Second, the separation between the context of discovery and the context of justification that was defended by the logical empiricists and Popper must be rejected (along with Brown's Platonic form of insight: see especially Brown 2011 and 2022). Finally, and more decidedly than Brown, also in this respect influenced by his fundamental Platonic conception, I will pose a difference only of degree between picture proofs and more standard formal proofs, between informal rigour and formal rigour (although it was some considerations of this author that led me to this position, especially the brief but very enlightening ones contained in Brown 2017). As we shall see, to prevent a possible misunderstanding, iconic or visual reasoning should not be understood in too narrow a sense, for example, as an a priori distinctive feature of mathematical TEs. Rather, it is a *de facto* characteristic that admits of varying degrees and that, to my mind, can never be entirely absent from any form of mathematical reasoning, no matter what degree of formalism it may have reached.

## 2. Are There TEs in Mathematics?

As already noted, one of the most serious methodological shortcomings of the literature concerning mathematical TEs is that it has almost always assumed without discussion that one can speak of mathematical TEs in a sense fundamentally similar to that in the empirical sciences. There is no space here to provide a general conception of TEs in the empirical-experimental sciences, as I have done elsewhere. For the purpose of this paper, I shall state almost dogmatically one of the most general conclusions of my account of empirical TEs, especially relevant to the present subject, and then ask whether there is anything analogous in mathematics.

My account (cf. Buzzoni 2004 and 2008) may be perhaps considered a 'quasi-Kantian treatment' of TEs that, as rightly recognized by James Robert Brown (2011, p. 202), is intermediate between rationalist-Platonist and empiricist accounts. To introduce the main idea of this account, let us begin with a very general definition of empirical TEs, which should be compatible with many different perspectives on the topic. An empirical TE is a way of arguing which anticipates, at the theoretical, discursive or linguistic level, a specific or concrete hypothetical experimental situation, so that, on the basis of previous knowledge, we are confident that certain interventions on the hypothetical experimental apparatus would modify some of its aspects (or 'variables') with such a degree of confidence that the actual execution of the experiment becomes superfluous (for this general definition, see e.g. Buzzoni 2008, p. 93).

Important differences between competing views on TEs only arise when one tries to define more precisely the nature and justification of that "previous knowledge" on the basis of which TEs, as long as they are taken as sound, reach their conclusions without appealing to real experiments (for example, according to Brown, this is a priori, Platonic knowledge, for Mach it is exclusively empirical knowledge, etc.).

A distinctive feature of my approach is that the relationship between TE and real-world experiment in the empirical sciences is two-sided: reflective-transcendental and empirical. From a reflective-transcendental point of view, TEs are the condition for the possibility of real experiments and, as such, they are not completely reducible to the latter: without the a priori capacity of the mind to reason counterfactually, we would not be able to formulate any particular empirical hypothesis and it would be impossible to plan real experiments that would be capable of testing it. For reasons of space, I cannot dwell further on this aspect (for further details, see Buzzoni 2008).

From an empirical point of view, however, one should accept the general principle of empiricism (formulated by Mill and taken up, among others, by William James and Karl R. Popper: cf. Mill 1863, p. 51; James 1912, p. 42; Popper 1969, p. 54), according to which observation and experimentation are the only sources of evidence relevant for the acceptance or rejection of empirical statements. If one accepts this principle, pure reason cannot (without contradicting itself) provide any content exclusively by its own efforts, spin reality out of a priori truths, independently of experience. In other words, pure reason cannot build an “a priori physics” and the "a priori knowledge" mentioned in the definition of TE given above must be explained by experience, and more precisely by experience understood in operational-experimental terms, i.e., in terms of actions and interactions between our body and the environment around us. In this sense, reason should not add anything of its own to the contents of experience, since what reason would add in this way would be an arbitrarily added content, something unjustifiable in principle, bad metaphysics, which intends to talk about empirical objects without recourse to the only source that can allow this kind of discourse, that is, experience as the result of the actual interaction between our bodies and the natural world around us.

It follows directly from this that the relation between empirical TEs and real world experiments is to be conceived in such a way that empirical TEs should be capable, at least in principle, of being realised, i.e. transformed or translated into real world experiments. In other words, the connection between thought and real world experiment is essential, or, to paraphrase Kant (as well as Lakatos), (empirical) TEs without real experiments are empty and real experiments without TEs are blind.

As already mentioned, I have made this point elsewhere and my purpose is not to develop this point further here but, rather, to use it in order to understand mathematical TEs. In fact, the intrinsic link between TEs and real experiments is already sufficient to distinguish in principle empirical TEs from mathematical TEs. The link between TEs and real world experiments, which is essential for empirical TEs, is missing in mathematics (or more precisely, as we shall soon see, in pure mathematics). Unless one supports a radically empirical conception of mathematics (as John Stuart Mill did, for example), it is very difficult to deny that an epistemologically decisive distinction between mathematics and empirical-experimental reasoning consists in the fact that only in the latter case does the real-causal interaction between our body and some external objects impinge directly on the truth or soundness of our assertions or inferences.

We can reaffirm this using the traditional distinction between applied and pure mathematics. Even if mathematics, in its concreteness, first emerges as applied mathematics (for example, as a technique for dividing arable land fairly), it can then close itself off, isolated in its own logical-inferential observatory, no longer accepting, even if only by methodological decision (and as such always modifiable), fresh experience. As is well known, applied mathematics remains always open to feedback from new experience. Elementary rules of addition do not apply when we add velocities in the order of magnitude of the speed of light; nor do they apply when we add drops of water.[1] On the contrary, pure or formal mathematics is essentially characterized by the capacity of reason, so to speak, to shut oneself inside, to refuse in principle (though *de facto* only provisionally) to take into account new experiences. Applied mathematics (and, more generally, empirical reasoning) is directly influenced by the real-causal interactions between our body and some external objects; on the contrary, the pure mathematician acts as if a particular field of experience were temporarily, but irrevocably closed, and then proceeds to investigate that field 'a priori', that is, clarifying its inner relations, possibilities and impossibilities. This founds the a priori status of mathematical reasoning, in the sense of a priori created, as Poincaré put it, by a "decree" of the mind (Poincaré 1902, p. 2), that is, a conventional act that only remains in force until a decision is taken to call it again into question. In other words, applied mathematics becomes pure mathematics to the extent that, although *de facto* only for a limited time (otherwise there would be no history of mathematics), no causal-experimental feedback is taken into account, i.e., arguments or derivations are protected from any attempt to refute them by means of new experimental data.

Now, these very brief hints are already sufficient for an initial attempt to answer the two questions: firstly, whether it is appropriate to speak of TE in mathematics, and secondly, if so, in what sense and under what conditions.

A first consequence of what I have been saying is that if we define the term "TE" as we have done above for the empirical-experimental sciences, it could be misleading and deceptive to speak of TEs in formal-mathematical reasoning. For the connection with a corresponding real experiment simply vanishes in purely formal reasoning. There can be no actual distinction between TEs and real experiments because of the self-protection of pure mathematics against new and potentially refuting experience: unlike applied mathematics and empirical knowledge, the actual performance of a mathematical

1 In using the mentioned traditional distinction, however, it is important to emphasise that the expression "applied mathematics" should by no means be understood in the sense that at the beginning there would be pure mathematics, which only later applies to the real world. On the contrary, historically, mathematics first emerged as applied mathematics in the sense that it responded from the outset to demands that came from practical-operational interests. For some further details on the epistemological status of mathematics, see Buzzoni 2011.

TE intended to demonstrate or to offer an ‘argument’ for a theorem totally coincides with the proof of that theorem, thus leaving no room for a separate real performance of the experiment. In other words – and taking into account our general definition of TEs – given the closure of the formal field, which cannot in principle use resources other than those already given within it, *the anticipation in thought of the solution of a problem in pure mathematics amounts to its actual solution.* Or, yet in other words, in pure mathematics there is, at least directly, no reality which could possibly unfold differently from the way in which the proof was developed.

From this crucial difference between mathematical and empirical TEs, many others follow. Here I merely mention that a well-known argument of Mach's in favour of using TEs instead of real world experiments loses all value: Mach's argument that experiments that are only thought out are more economical than those that are actually carried out loses all force in the case of pure mathematical reasoning: we simply do not have the second term of comparison. Insofar as the dialectical relationship of contraposition and reciprocal implication between TEs and real ones disappears, there is no room for a performance different from the train of reasoning which took the course it did take and the actual performance of a mathematical TE intended to prove or offer an ‘argument’ for a theorem coincides completely with the proof or argument for it.

However, there is a very simple objection that can be made (and de facto has been made) to this thesis. It is a fact that many authors, among them some mathematicians, have spoken about TEs in mathematics, and have illustrated their meaning of TE with concrete examples (see e.g. Szabó 1958 and 1969; Lakatos 1963-4; Brown 2004, 2007, 2011, 2017, and 2022; Starikova and Giaquinto 2018). Among those who have spoken of TEs in mathematics, special mention must be made of James Robert Brown and, before him, Imre Lakatos. For Brown, in mathematics “there is something analogous to thought experiments – visual reasoning and picture proofs.” (Brown 2007, p. 3). Brown is certainly the one who has put forward both the best arguments for, and the best examples of, mathematical TEs, understood as arguments that conclude in a cogent way on the basis of "visual reasoning", simple diagrams, or, in any case, situations that can be easily visualised (an important feature of empirical TEs too) (cf. e.g., Brown 1997, p. 177, and Brown 2022). Here is one of his (numerous) examples:

$$1 + 2 + 3 + \ldots\ldots + n = n^2/2 + n/2$$ (see Fig. 2)

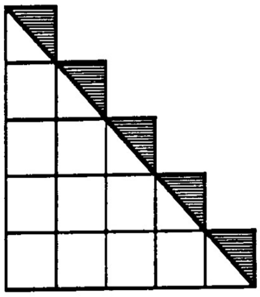

Fig. 2 (Brown 2011, p. 91)

As Brown rightly notes, it must be decided on a case-by-case basis whether the use of picture proofs or more standard proofs provides better or worse insight into the truth of a conclusion. Perhaps mathematical induction provides a "better" and "more explanatory" characterization of the natural numbers than picture proofs. But at least in some cases, such as the diagram in Fig. 2, in which it must be shown that the value of the sum "$1 + 2 + 3 + ........ + n$" is equal to "$n^2/2 + n/2$", the situation seems clearly reversed (Brown 1997, p. 177; on this point see also Brown 2017 and 2022).

The importance of this objection is greater than it might seem at first glance, since I too am among the many who believe (for reasons that cannot be made explicit here for lack of space) that a philosophy of mathematics which disregards the self-understanding of mathematicians at work would be a sterile enterprise, a sort of Wittgensteinian "idle wheel" in human culture. However, the soundness of this objection is not necessarily in tension with my first general condition for clarifying the status of mathematical TEs, that is, that we should not neglect fundamental differences between logical-mathematical and empirical-experimental reasoning. On the contrary, we ought to follow Brown's insistence on the visual or pictorial character of mathematical TEs while yet holding firmly to the distinction in principle between mathematical and empirical TEs.[1]

1 Leaving aside (for reasons of space) some differences, I would like to at least mention the broad agreement that exists between the position defended here and Peirce's considerations on mathematics as diagrammatic reasoning. According to Peirce, in fact, the claim that, as his father Benjamin Pearce had stated, mathematics is "the science which draws necessary conclusions" (Peirce 1895, p. 7), is intimately connected with the importance of graphic or pictorial constructions, or – as Peirce expressed it – with the fact that the mathematician "observes nothing but the diagrams he constructs himself" (Peirce 1895, p. 4). As Peirce noted, in responding to the demands of experience (be they those of the engineer, the insurer, the physicist, or the businessman planning the purchase of land), the mathematician does two things: first, he tries to simplify the facts; and secondly, he tries to formulate another simpler but quite fictitious problem, which "is sufficiently like the question he should answer" to serve as a "substitute" for it. This process, which Peirce calls "skeletonization or diagrammatization of the problem", makes possible the "tracing the consequences of the hypothesis", which is the most characteristic aspect of mathematical work (cf.

Moreover, as we shall see, answering the objection just raised will lead to specifying some conditions – in addition to that of not confusing formal and experimental reasoning – under which it is possible to speak coherently and unambiguously of mathematical TEs. For this purpose, let us briefly review a debate between James Robert Brown and John Norton.

## 3. In what sense and under what conditions may we speak of TE in mathematics

According to Brown, many mathematical proofs presuppose that we immediately see the truth of a theorem, in a sense of "seeing" which is essentially similar to Plato's. This is for example particularly clear in the picture proof of the theorem in number theory which states that the sum of the numbers up to n is equal to $(n^2 + n)/2$ and which we have just presented in the diagram of figure 2.

On the contrary, according to Norton,

> "we do not immediately see the truth of the theorem upon being confronted with the figure. At first we are startled and then we give ourselves a silent commentary on the figure until we work it out. This may take more or less time according to our mathematical abilities but will be rapid in the case of someone with a mathematical background." (Norton 1996, p. 352)

On this important issue, I side with Norton (and Peirce) against Brown. Brown is wrong in underestimating the steps of reasoning necessary for an objective, that is, intersubjectively testable, proof. This is true, not only for proofs that make greater use of verbal/symbolic means, but also for picture proofs: in order to be convincing, we cannot neglect the fact that they require methodical steps (including visual steps) to be reconstructed and re-appropriated in the first person.

Now, the fact that a reconstruction in the first person is necessary in order to have intersubjectively testable reasoning is intimately connected with one important sense in which the neopositivist distinction between the context of discovery and the

---

Peirce 1898, pp. 212-213; as is well known, Peirce's reflections on diagrammatic reasoning have been often taken up in the debate of the last decades on the graphical or diagrammatic character of mathematical proofs (see e.g. Hessler and Mersch (eds.) 2009, Pombo and Gerner (eds) 2010, Panza 2012, Giardino 2013, Meynell 2018, Chapman et al (eds) 2018). In our terminology, the process of "skeletonization" or "diagrammatization" is a process of idealization subordinated to the main purpose of drawing valid consequences within an inferential symbolic space shielded from the real causal interactions present in the natural world.

context of justification must be rejected. This separation, at least in one important sense, is a kind of *contradictio in adjecto*, since it assumes that a kind of untestable insight plays a role in scientific knowledge, which, on the contrary, must always be in principle intersubjectively testable knowledge. Intersubjective controllability in principle is the most general requirement of scientific objectivity, and is no less valid for mathematical knowledge than it is for empirical-experimental knowledge. On this point there is a basic analogy between a mathematical test and an empirical one (which in no way detracts from the difference in principle highlighted above). In the last analysis, a mathematician can test and justify a theorem only by reproducing in the first person the same mathematical steps that were taken in constructing it; in the same way, an empirical hypothesis can be tested by reproducing the same experimental steps that confirmed it in the first place.

From this point of view, by separating discovery from justification, the logical positivists and Popper neglected a *genetic-reconstructive* (and, in the last analysis, historical) aspect that is inseparable not only from empirical, but also from logical and mathematical justifications. They failed to realize a sense in which a certain 'genetic' attitude is essential for justification in a very general sense. If we want to test the truth of a statement, we must adopt a *genetic-reconstructive* (or *genetic-historical*) attitude and retrace the operations performed and communicated by those who first obtained a certain result by means of those operations. On pain of losing the characteristic of being intersubjectively controllable, we cannot forget to check not only all answers, but also all questions, by a personal reconstruction of reasons understood as methodical steps that indicate a goal set at the end of a path to be travelled. In other words, we must suspend a priori assumptions or commitments and reconstruct the development and internal process that leads to our final understanding of the truth-value of the statement in question. In this sense, there is no point in time after which it is possible to disregard in principle the context of discovery[0].

At this point we are able to set forth a second condition for a use of the expression "TE" in mathematics that is coherently tenable. To the first condition, that the distinction between logical-formal and experimental-operational reasoning must not be underestimated or neglected, we may now add a second one, which is that, at least in an important sense, we have to reject the dichotomy between discovery

0 There is also a sense in which the neopositivist distinction between discovery and justification must be preserved, but this is not what I intend to dwell upon now. For a more detailed treatment of this point, see Buzzoni 2008, above all Ch. 1, Sections 4-6, and Ch. 3, Sections 4-6. For more historical details on the discovery/justification distinction, see Schickore and Steinle (eds) 2009, above all Part I and Part II.

and justification. Once we recognise, at least to a certain extent in line with Norton and against Brown's Platonism, that the intersubjective controllability in the first person of all our reasoning is a hard-to-renounce requirement of any rational discourse, there is no reason in principle to reject picture proofs in mathematics, provided that we extend this same demand for intersubjectivity to them. On the contrary, if these two conditions are satisfied, we may use the expression "mathematical TE" to characterize a particular kind of mathematical proof, in which visualization plays a relatively more important role.

Finally, from the second condition set forth above – namely the rejection (in an important sense) of the separation between discovery and justification – a third one follows: between mathematical TEs and more formal proofs there can only be a difference of degree and not of principle.

To argue briefly in favour of this thesis – and at the same time to reaffirm the second – I will examine Van Bendegem's effort to understand TEs in mathematics, not unlike here, as proofs in which visualisation plays an important role. It might be said that Van Bendegem takes up again Brown's point of view, though in a Lakatosian rather than in a Platonic spirit:

> "If we equal facts for the scientist to proofs for the mathematician – he writes – then a mathematical thought experiment is to be something like an "imaginary" proof that should help us to understand the proofs we are looking for. Hence the idea that mathematical thought experiments should provide insight into either what a proof could look like, or why it is convincing, explanatory, in short, why it functions as a proof." (Van Bendegem 2003, pp. 9-10)

Here, we must distinguish two different theses, one of which must be rejected, while the other, though with some qualifications, should be accepted and interpreted as our third condition.

In the just mentioned passage Van Bendegem says that mathematical TEs, which are visual proofs based on diagrams, provide insight into "what a proof could look like". Now, if this assertion were to be understood in the sense that mathematical TEs are only a means to discover mathematical proofs, i.e., that TEs are not persuasive arguments, but mere heuristics to find them, this would be coherent with the fact that Van Bendegem (and after him Starikova and Giaquinto 2018: see § 4) distinguishes *in principle* between mathematical TEs and mathematical proofs.

However, if so interpreted, Van Bendegem's claim would presuppose in the formulation itself the dichotomy between discovery and justification. In this sense, on the basis of the preceding considerations, we should reject this claim without hesitation. It is no accident that this same thesis was defended by Hempel for empirical TEs, which according to him were useful only until a rigorous and testable formulation was found. Hempel is indeed simply applying the distinction between the irrational moment of discovery and the objective moment of justification when he writes that TEs “may *suggest* hypotheses, which must then be subjected, however, to appropriate objective tests.” (Hempel 1965, p. 165).

But it is interesting to note that the second part of Van Bendegem's statement seems to point in the opposite direction. Here he says that a TE in mathematics helps to better understand "why" a proof is "convincing, explanatory, in short, why it functions as a proof". This is a very important point, with which I agree, insofar as this is understood to mean that, as I stated earlier, a proof or an argument functions as a proof or as an argument only if it provides the reasons for believing in its conclusion in the form of steps to be reconstructed and re-appropriated in the first person. This would be entirely in agreement with the sense in which we have rejected the distinction between discovery and justification. In this sense, the fact that TEs in mathematics may de facto provide insight into "what a proof could look like" strictly speaking means that they already constitute a necessary, albeit minimal, element of justification. But unfortunately, even if this would completely overcome the separation between discovery and justification, it would not be enough to make van Bendegem's position coherent, due to the fact that, as already mentioned, he distinguishes *in principle* between mathematical TEs and mathematical proofs.

However van Bendegem's position is to be understood with reference to the discovery/justification distinction, our brief discussion shows that, to be consistent with the rejection of the hiatus between an irrational process of discovery and an intersubjectively testable justification, we must assume no qualitative difference or epistemological break, but only a difference of degree between the visual reasoning of TEs (relatively richer in diagrams) and formal proofs (relatively poorer in intuitive-diagrammatic elements). On the contrary, to ascribe to mathematical TEs a heuristic or preparatory role with respect to formal proofs in the most standard sense is to bring back through the window the neopositivistic separation between the context of discovery and the context of justification, which was earlier thrown out the door.

Many would protest that we should distinguish between, on the one hand, arguments that only provide an "understanding" of a mathematical proof (and that still retain important links with the imaginative-iconic or experiential-intuitive moment of empirical knowledge), and, on the other hand, logically formal proofs, in which the individual steps have been individually inspected and checked in the light of primitive concepts, postulates and admitted rules of inference (cf. e.g. Heinzmann 2015, pp. 116 ff., where, strictly speaking, the author distinguishes in general between “mathematical understanding” and “mathematical explanation”; see also Heinzmann 2022).

But what can it mean to mathematically ‘understand' something, if not, precisely, to reconstruct in the first person the conceptual steps that have led from certain premises to certain conclusions, in order to be sure that they are based on previously accepted knowledge? Can I really say that I have fully understood the meaning of a theorem if some passages of its demonstration are still obscure or ambiguous? Surely not. One may certainly speak of a partial understanding, but in this case too, we are taking into account some possible reason in favour of it, since a hypothesis that does not have at least prima facie a certain plausibility would not even be advanced as a hypothesis. It is true that understanding, like justification, is not an all-or-nothing thing, it is rather a functional relationship, and what is understood or justified at a certain point in time may later appear to be in need of deeper understanding or justification: as Poincaré said, there are not “solved problems and others which are not; there are only problems *more or less* solved” (Poincaré 1910, p. 86). But in all cases, understanding and justification cannot be separated. It is not possible to understand something if one does not have some reason, even a minimal one, to justify it (in no matter how minimalist a sense of 'justifying').

From this point of view, any *qualitative* difference between, on the one hand, arguments that extend our 'understanding' of a mathematical claim and, on the other hand, verbal-symbolic, formal, and properly rigorous proofs of this claim vanishes. This difference is consistent with a Platonic position that distinguishes between, on the one hand, an 'understanding' by means of an intuitive (noetic, in Plato’s parlance) type of knowledge and, on the other hand, a discursive (dianoetic, in Plato’s way of speaking) argumentation. But this is exactly the sense in which the distinction between discovery and justification must be rejected. To maintain that TEs in mathematics prepare future formal proofs, on closer inspection tacitly assumes an intuitive and creative capacity that would be at work only at the initial moment of the discovery, being still incapable at that moment of providing the authentic justification that only complete formalization could provide. In a word, this tacitly assumes the untenable sense in which the logical positivists distinguished discovery from justification, psychology from logic, history from rational reconstruction.

There is one final objection to be addressed. It might be objected that to regard mathematical TEs as mathematical reasoning in which visualisation plays a significant role, would mean to accept a too restricted use of both TEs and proofs in mathematics (I owe this objection to a personal communication from Gerhard Heinzmann; but see also Heinzmann 2022).

There are two replies to this objection. First, the objection assumes a too restricted use of visual or diagrammatic reasoning. In fact, what is iconic, visual or diagrammatic can only be distinguished from what is conceptual by means of a reflective activity (proper to philosophical analysis), which distinguishes in thought what is connected in our concrete imaginative activity – be it scientific or not.[1] In any kind of concretely real experience, universals must always be exemplified. It is easy to show that this is true in everyday empirical knowledge. Thus, for example, seeing a road sign of a dangerous curve might remain in my mind a mere image without any clear conceptual (that is, prospective or anticipatory) character, but, if connected to my knowledge of the road traffic regulations, it is most likely to give rise to conditional forms of prediction (i.e. implicit or explicit TEs). For example, it may arouse in my mind the idea of hypothetical driving behaviours and their possible empirical consequences (for example, "if I don't slow down here, I'll end up off the road"). But this is true in general, and even in the most abstract or pure mathematics, some intuitive or imaginative elements are to be connected with the use of linguistic symbols. This implies that no mathematical demonstration, however formalised, can be completely devoid of an 'intuitive' element, however minimal, i.e., an image, diagram or icon. According to Weyl (who in turn reports what was narrated by O. Blumenthal), Hilbert seems to have said about his own axiomatic system that "[i]t must be possible to replace in all geometric statements the words point, line, plane, by table, chair, mug." (Weyl 1944, p. 635) I agree entirely, but this proves just the opposite of what it is usually assumed to prove. The essential point is that, in any case, we need some particular image, no matter whether it is lines and points or tables and chairs: without an intuitive anchorage (and more precisely, without an intuitive-operational anchorage, linked to our agency), our symbolic language, given the nature of the symbol as such, simply could not exist.

1 Because of this need to distinguish and at the same time connect the conceptual and the concrete-sensible aspects of imagination, the distinction drawn by Stuart 2019 between, on the one hand, imagination as an ability ($imagination_0$), and, on the other hand, two kinds of imagination as different particular manifestations or "uses" of it ($imagination_1$ and $imagination_2$, to which, in my opinion, empirical research could add many more), is important. Stuart sees his distinction as qualitative (or 'categorical'), and it is in this sense that I accept and use it here. On this point see also Savojardo 2022, who rightly insists on the connection between this distinction, the rejection of the separation between discovery and justification, and the notion of embodied cognition.

So far for the first reply to the objection that to reserve the expression "TE" for arguments or demonstrations where a large part of visual reasoning is involved would mean a too restricted use of both TEs and proofs in mathematics. Let us now turn to a second reply. Up to this point, I have defended my proposal by making explicit certain conditions necessary for its internal philosophical coherence. But of course, any terminological choice should also depend on the usages actually established within a linguistic community. Now, even from this point of view, my proposal is easily defensible.

A first reason for using "TE" to refer to aspects of mathematical reasoning in which visualisation plays a significant role is that most authors, also of different tendencies – from Brown (2011, p. 17) to Norton (1991, p. 130) to exponents of the mental modelling approach to TEs (cf. e.g., Miščević 2021, ch. 3.5) – considered perspicuity and visualisation as one of the hallmarks of TEs not only in the empirical sciences, but also in any other field, including philosophy.

A second important reason lies in the fact that several authors, and among them some of great significance, have used the term TE with an essential reference to visual reasoning. In addition to the aforementioned Brown, special mention must also be made of Lakatos. In *Proof and Refutations,* discussing a case in which the importance of visual reasoning is crucial, Lakatos explicitly quoted the expression "thought-experiment" from Szabó (1958), who had claimed that the expression "thought-experiment" (which he proposed as corresponding to the Greek term *deiknymi*) was "the most ancient pattern of mathematical proof" (see Lakatos 1963-1964, I, p. 10 fn.). Now, with the term "deiknymi" Szabó means something very similar to Lakatos's "informal demonstration" (for this claim see also Máté 2006), although, unlike Lakatos, he used it in a negative rather than a positive sense, since he designated by this expression a form of argumentation that, in the pre-Euclidean Greek mathematics, provided only "evidence" that was "of a purely empirical-visual kind" (*rein empirisch-anschaulicher Art*) (cf. Szabó 1960, p. 40; see also Szabó 1969, p. 249. As a paradigmatic example of "thought-experiment" Szabó cited the demonstration in Plato's Meno, Sections 82b-85e, in which a slave, guided by Socrates's questions and by various drawings, manages to prove a simple geometric theorem).

Finally, among the authors who used the term TE with an essential reference to visual reasoning in mathematics, we ought to mention Edmond Goblot (1858–1935), unjustly neglected today. He insisted on the inevitable presence, in mathematical TEs, of a "singular", "intuitive" or, as we might add, iconic element. Goblot interpreted mathematics in a constructive and operational way, and claimed that in a "Gedankenexperiment" (*expérience mentale)* – a term which, in explicit opposition to Mach, he used only for mathematical demonstrations – an "intuitive" element is the necessary starting point of any mathematical construction. If one cannot demonstrate except by operating, mathematical proofs require an intuitive element in the form of a singular figure, both because operations cannot be performed, even mentally, without such an element, and also because their results cannot be "ascertained" *(constate)* except on the basis of it. All these features of Goblot's theory of TEs in mathematics are well summarised in the following passage:

"The demonstration consisted […] in an *operation* and in the *ascertainment [constatation]* of the result obtained. It goes without saying that this is not a manual operation [...] but a mental operation, and that it is not a question of a *physical* ascertainment *[constatation],* such as that one could make with measuring instruments, but a *logical* ascertainment. [...] The fact is that one cannot demonstrate except by operating; but an operation (construction, superposition, rotation, etc.) cannot be performed, even mentally, and its result cannot be ascertained [*constaté*] except on the basis of a singular figure" (Goblot 1918, 263-64. For more details on Goblot's account of mathematical TEs, see Buzzoni 2021b).

In short, to the much better known authors mentioned above who used "TE" to designate in mathematics a mode of reasoning connected to the visual, iconic or diagrammatic element, we may add Goblot, an author who dedicated a good part of his theoretical efforts to understanding mathematical demonstrations as TEs.

## 4. Conclusion

Using an already existing and relatively widespread use of the expression in question, "mathematical TEs" may be regarded as mathematical reasoning in which visualisation plays a relatively more important role. However, certain conditions have to be met to ensure an unambiguous and consistent use of the term "TE". These conditions may be summarized as follows:

1) The distinction between logical-formal thinking and experimental-operational thinking must not be neglected or underestimated. If one takes empirical TEs as a paradigmatic instantiation of the use of the term, we should say that, insofar as the dialectical relation of unity and distinction between thought and real experiments disappears, this expression could be misleading in pure mathematics, where the solution given to a problem in thought alone already amounts to the

actual solution. This, however, should not discourage us from speaking of TEs in mathematics, as long as one is clear that the conditions of use for that term, while similar, is also clearly distinct from the experimental-scientific one. For this purpose, the first and most important condition is to always bear in mind the difference between mathematical and experimental-scientific TEs.

2) A second condition for speaking coherently and unambiguously of mathematical TEs is to reject the separation between the context of discovery and the context of justification, at least in one of the main senses in which it was defended by the logical empiricists and Popper. This excludes any position which, ascribing to mathematical TEs only a heuristic role, regards them an intermediate step to attain more traditional forms of rigour. Such a position would implicitly assume the existence of a sphere of irrational understanding or insight that would not be subject to the principle of intersubjective controllability in the first person.

3) A third condition is that the distinction between mathematical TEs and formal proofs must be regarded as one of degree, and not as a qualitative one, however important this distinction may be de facto for particular or local purposes. Between the formal proofs and the so-called pictorial or visual ones there is only a de facto and historical difference. It is a distinction between different ways or modalities of doing mathematics in the never closed set of possible mathematical reasoning.

## References


Brown J.R. 1997. Proofs and Pictures. The British Journal for the Philosophy of Science, 48, pp. 161-180

Brown J.R. 1999: Philosophy of Mathematics. An Introduction to the World of Proofs and Pictures. Routledge, New York (quotations are from the 2008 edition).

Brown J.R. 2004. Peeking into Plato's Heaven. Philosophy of Science, 71, pp. 1126-1138.

Brown J.R. 2007. Thought Experiments in Science, Philosophy, and Mathematics. Croatian Journal of Philosophy, 7, 3-27.

Brown J.R. 2011. The Laboratory of the Mind: Thought Experiments in the Natural Sciences. 2nd edition, Routledge, London.

Brown J.R. 2017. Proofs and Guarantees. The Mathematical Intelligencer. 39:47-50 (doi 10.1007/s00283-017-9730-1).

Brown, J.R. 2022. Rigour and Thought Experiments: Burgess and Norton. Axiomathes. https://doi.org/10.1007/s10516-021-09567-2 (this issue)

Buzzoni M. 2004. Esperimento ed esperimento mentale, Angeli, Milano.
Buzzoni M. 2008. Thought Experiment in the Natural Sciences. An Operational and Reflexive-Trascendental Conception, Königshausen+Neumann, Würzburg.
Buzzoni M. 2011. On Mathematical Thought Experiments. Epistemologia. An Italian Journal for Philosophy of Science, 34, pp. 5-32.
Buzzoni M 2021a. Mathematical vs Empirical Thought Experiments. In: C. Calosi, P. Graziani, D. Pietrini and G. Tarozzi (eds), Experience, Abstraction, and the Scientific Image of the World, pp. 215-228. FrancoAngeli Editore, Milan.
Buzzoni M 2021b. A Neglected Chapter in the History of Philosophy of Mathematical Thought Experiments: Insights from Jean Piaget's Reception of Edmond Goblot. HOPOS, 11, pp. 282-304.
Chapman P., Stapleton G., Moktefi A., Perez-Kriz S., Bellucci F. (eds) 2018. Diagrammatic Representation and Inference. Springer, Cham.
Cohnitz D. 2008. Ørsteds 'Gedankenexperiment': eine Kantianische Fundierung der Infinitesimalrechnung? Ein Beitrag zur Begriffsgeschichte von 'Gedankenexperiment' und zur Mathematikgeschichte des frühen 19. Jahrhunderts". Kant-Studien, 99, 407-433.

Fehige Y, Vestrucci A (2022) On Thought Experiments, Theology, and Mathematical Platonism. Axiomathes . https://doi.org/10.1007/s10516-022-09627-1 (this issue)
Giardino V. 2013. Towards a diagrammatic classification. The Knowledge Engineering Review 28: 237–248.
Glas E. 2001a. The 'Popperian Programme' and Mathematics. Part I: the Fallibilist Logic of Mathematical Discovery. Studies in History and Philosophy of Science, 32, 119–137.
Glas E. 2001b. The 'Popperian Programme' and Mathematics. Part II: From Quasi-Empiricism to Mathematical Research Programmes. Studies in History and Philosophy of Science, 32, 355–376.
Goblot E. 1918. Traité de logique. Colin, Paris.
Heinzmann G. 2015. Pragmatism and the Practical Turn in Philosophy of Mathematics: Explanatory. In E. Agazzi, G. Heinzmann (eds), Pragmatism and the Practical Turn in Philosophy of Sciences, pp. , 113-129. Angeli, Milan.
Heinzmann G. 2022 Mathematical Understanding by Thought Experiments, this issue.
Hertogh C. P. 2021. Computation and Visualization Thought Experiments after Lakatos's Heuristic Guessing Method. In:

Philosophical Approaches to the Foundations of Logic and Mathematics (Poznań Studies in the Philosophy of the Sciences and the Humanities, Band 114), pp. 271–297. Brill, Leiden.
Hessler M. and Mersch D. 2009 (eds.). Logik des Bildlichen. Zur Kritik der ikonischen Vernunft. Transcript, Bielefeld.
James W. 1912. Essays in Radical Empiricism. Longman, London.
Koetsier T. 1991. Lakatos' Philosophy of Mathematics. A Historical Approach. North Holland, Amsterdam.
Kühne U. 2005. Die Methode des Gedankenexperiments, Suhrkamp, Frankfurt a.M.
Lakatos I. 1963-4. Proofs and Refutations. The British Journal for the Philosophy of Science 14: I, pp. 1-25; II, pp. 120-139, III, pp. 221-245, IV, pp. 296-342.
Larvor B. 1999. Lakatos' Mathematical Hegelianism. The Owl of Minerva 31, pp. 23-54.
Lenhard J 2022. Proof, Semiotics, and the Computer. On the Relevance and Limitation of Thought Experiment in Mathematics, Axiomathes, 32, n. 6 (this issue).
Mach E. 1906. Über Gedankenexperimente, in: Erkenntnis und Irrtum, Barth, Leipzig, 1905, 5th ed. 1926, pp. 183-200; Engl. transl. by T.J. McCormack, On Thought Experiments, in: Knowledge and Error, Reidel, Dordrecht and Boston, 1976, pp. 134-147.
Máté A 2006. Árpád Szabó and Imre Lakatos, Or the relation between history and philosophy of mathematics. Perspectives on Science 14, pp. 282-301.
Meynell L. 2018. Images and Imagination in Thought Experiments. In Kantian Accounts of Thought Experiments, ed. J. R. Brown, Y. Fehige, and M. Stuart, pp. 498-511. Routledge, New York.
Mill J.S. 1863. Utilitarianism. Parker, London.
Miščević N. 2021. Thought Experiments. Springer, Cham.
Mueller I. 1969. Euclid's Elements and the axiomatic method. British Journal for Philosophy of Science, 20, pp. 289–309.
Norton J. 1991. Thought Experiments in Einstein's Work. In Horowitz T. and Massey G.J. (eds) 1991, Thought Experiments in Science and Philosophy, pp. 129-148. Rowman & Littlefield, Savage (MD).
Norton J. 1996. Are Thought Experiments Just What You Thought? Canadian Journal of Philosophy, 26, 333-366.
Norton J.D. and Parker M.W. 2022. An Infinite Lottery Paradox. Axiomathes. https://doi.org/10.1007/s10516-021-09556-5 (this issue)
Panza M. 2012. The twofold role of diagrams in Euclid's plane geometry. Synthese 186: 55–102.

Peirce 1895. The Nature of Mathematics. Quotations are from Philosophy of Mathematics. Selected Writings, ed. by Matthew E. Moore, pp. 3-9. Indiana University Press, Bloomington and Indianapolis, 2010.
Peirce 1898. The Logic of Mathematics in Relation to Education. Educational Review 15, pp. 209-216.
Poincaré H. 1902. La Science et l'hypothèse. Paris: Flammarion. Quotations are from Science and Hypothesis. The Complete Text, Engl. Transl. by Mélanie Frappier, Andrea Smith, and David J. Stump, Bloomberg, London, 2018.
Poincaré H. 1910. The Future of Mathematics. The Monist, 20, pp. 76-92
Pombo O. and Gerner A. (eds) 2010. Studies in Diagrammatology and Diagram Praxis. College Publications, London.
Popper K.R. 1969. Conjectures and Refutations, 3d ed, Routledge & Kegan Paul, London.
Savojardo V. 2022 forthcoming. Scientific embodied imagination as a link between creativity and method in the empirical sciences.
Schickore J. and Steinle F. (eds.) 2006. Revisiting discovery and justification: historical and philosophical perspectives on the context distinction. Dordrecht: Springer.
Shaffer M.J. 2015. Lakatos' Quasi-empiricism in the Philosophy of Mathematics. Polish Journal of Philosophy 9, pp. 71-80.
Sherry D. 2006. Mathematical reasoning: induction, deduction and beyond. Studies in History and Philosophy of Science, 37, 489-504.
Starikova I. 2007. Picture-Proofs and Platonism. Croatian Journal of Philosophy, 7, 81-92.
Starikova I. and Giaquinto M. 2018. Thought Experiments in Mathematics. In: J.R. Brown, Y. Fehige, M. Stuart (eds.). The Routledge Companion to Thought Experiments, pp. 257-278. Routledge, New York.
Stuart M.T. 2019. Towards a dual process epistemology of imagination. Synthese https://doi.org/10.1007/s11229-019-02116-w
Szabó Á. 1958. 'Deiknymi' [Greek δείκνυμι] als Mathematischer Terminus fur 'Beweisen'. Maia, N.S., 10, pp. 1-26.
Szabò Á. 1969. Anfänge der griechischen Mathematik. De Gruyter, Oldenbourg. Quotations are also from the Engl. Translation, The Beginnings of Greek Mathematics. Reidel, Dordrecht 1978.
Van Bendegem J.P. 2003. Thought Experiments in Mathematics: Anything but Proof. Philosophica, 72, pp. 9-33.
Weyl H. 1944. David Hilbert and his mathematical work. Bulletin of the American Mathematical Society, 50, 612-654.

Witt-Hansen J. 1976. H.C. Ørsted, Immanuel Kant, und the Thought Experiment. Danish Yearbook of Philosophy, 13, pp. 48-65.

Yourgrau W. 1967. On models and thought experiments in quantum theory. Monatsberichte der Deutschen Akademie der Wissenschaften zu Berlin, 9, pp. 865–874.

Yuxin Z. 1990. From the Logic of Mathematical Discovery to the Methodology of Scientific Research Programmes. The British Journal for the Philosophy of Science, 41, pp. 377-399.