

Phenomenology and Thought Experiments

Thought Experiments as Anticipation Pumps

Harald A. Wiltsche
Department for Philosophy
University of Graz, Austria
harald.wiltsche@uni-graz.at
www.haraldwiltsche.com

The aim of this paper is to present an outline of a phenomenological theory of thought experiments (henceforth: *TEs*). In doing so, I am dealing with a topic that is currently starting to receive increased attention from philosophers with phenomenological leanings. However, since no serious attempt has been made to tackle the issue in a systematic fashion, I will not merely review existing phenomenological work on *TEs* (such as Mohanty 1991; Kunjundzic 1995; Froese & Gallagher 2010; Fehige & Wiltsche 2013; Wiltsche 2013; Hopp 2014). For the most part, my paper is programmatic: its aim is to suggest some basic directions in which a phenomenological theory of *TEs* should be developed.

The paper is structured as follows. I will begin by saying a word or two on what phenomenology is and on why I believe that it can contribute to the ongoing debate on *TEs*. I shall then introduce five phenomenological concepts that will prove crucial for an understanding of *TE*-reasoning: *fulfillment*, *frustration*, *horizon*, *anticipation* and *background knowledge*. My strategy is to first give a brief sketch of how knowledge acquisition works in the perceptual realm. It is against this backdrop that I will then outline how knowledge is generated in the realm of *TE*-reasoning. It is one of my main theses that *TEs* should be understood as *anticipation pumps*.

Before I begin, however, two qualifying remarks are in order. The first concerns my use of the term “phenomenology”: Just as there is more than *one* analytic method, there is no general agreement within the phenomenological community on what *the* phenomenological method precisely is. It is thus important to bear in mind that the following portrayal of phenomenology reflects my own background in (a liberal reading of) Edmund Husserl’s philosophy. Secondly, much of what I will have to say about *TEs* concerns *TE*-reasoning in science and not *TE*-reasoning in philosophy. To a certain extent, this limitation again reflects my own background in the philosophy of science. However, since I believe that scientific and philosophical *TEs* differ only in degree and not in kind, I am confident that the proposed framework can also be applied to non-scientific *TEs*.

1. Getting Phenomenology off the Ground

When phenomenology entered the philosophical stage in the early 20th century, Neo-Kantianism was still in full swing. Otto Liebmann had coined the unifying motto “Back to Kant!” under which the Neo-Kantians sought to overcome the dispute between materialism and idealism that had dominated much of the philosophical discourse in Germany and Austria up until the mid-19th century. It was against this motto that the first generation of phenomenologists directed their own slogan “Back to the *things themselves!*”. Yet, it wasn't particularly Kant of whom Husserl and his followers were critical. The first wave of phenomenologists disapproved of any kind of philosophy that

looks at its problems through the glasses of pre-established theories, systems or schemes. Instead of forcing problems into a particular (and potentially artificial) theoretical mould, phenomenologists were (and still are) driven by a deep respect for the *phenomena*, i.e. the things exactly as they are given in experience. On a phenomenological view, many philosophical problems could be solved – or even better: made to evaporate – if we resisted the temptation to interfere with ready-made theoretical schemes and put more effort in a faithful description of the phenomena.

Paradigmatic for this approach is Husserl's sixth Logical Investigation (Husserl 2001b, 177-348): The aim of this book-length treatise is to tackle one of the most fundamental issues in philosophy, namely to understand the concepts "truth", "knowledge" and "knowledge acquisition", as well as the relations between them. On Husserl's view, the trouble with many existing approaches to this issue is that they are infected with pre-established theoretical schemes already at the level of the formulation of the problem. For instance, a common way to start is to ask how a self-enclosed subject is able to reach the external world in order to attain justified true beliefs about reality by relying on acts of visual, auditory or olfactory perception. However, built into this question are a number of far-reaching assumptions such as common-sense realism or the view that a gap between the interiority of the mental and the exteriority of the world in fact exists. Husserl accepts that these and similar other assumptions may be useful or, in some cases, even indispensable in certain practical contexts. Yet, since the job of philosophy is not to simply adopt existing assumptions, but rather to scrutinize them, philosophical analysis must proceed from a standpoint that is not already contaminated with pre-established theoretical schemes. It is for this reason that Husserl actively "brackets"¹ all kinds of metaphysical, scientific and commonsensical assumptions and starts with a careful and unbiased *description* of the types of conscious experiences in which different types of objects become present from a first-person point of view. To be sure, nothing prevents the phenomenologist from eventually taking sides in traditional metaphysical disputes about, say, dualism, realism or physicalism. But whatever she will end up saying on such matters must result from a description of the phenomena and not from the mere stipulation of basic maxims or principles.

As I have pointed out, phenomenologists are driven by a deep respect for the phenomena, i.e. for the things as they appear in conscious experience. This general tenet – which, in my view, captures much of what makes a phenomenological account truly phenomenological² – also yields important consequences for how phenomenologists

¹ The notion of "bracketing" comes from Husserl's mathematical background, specifically from the concept of absolute value (the notation " $|x|$ " was introduced by Husserl's teacher Karl Weierstrass). The absolute value $|x|$ of a real number x is the non-negative value of x without regard to its sign. So, for instance, 3 and -3 have the same absolute value 3. Hence, to bracket a natural number means to be indifferent with regard to a number's property of being positive or negative and to focus on its magnitude instead. Accordingly, to bracket assumptions neither means to endorse, nor to deny them. It merely means to focus on what is given independently from all assuming.

² It should be emphasized that it is mainly for the sake of brevity that I am working with a rather restricted conception of phenomenology here. For the purposes of this article, I am conceiving of the phenomenological method roughly along the lines of Charles Siewert's "plain phenomenology". On this view, one is doing phenomenology if (1) one explains mental distinctions, if (2) one shows how such explanations have significant theoretical consequences, if (3) one's explanations relies on a source of warrant special to some *first-person* applications of the distinctions explained and if (4) one does not assume that such first-person warrant as one relies on is derived from third-person evidence (cf. Siewert 2007, 202). However, I'd also like to stress that this understanding of phenomenology does not do full justice to the all the subtleties of Husserl's account. Readers interested in a more comprehensive

should approach the issue of scientific TEs. Let me begin by briefly summarizing how the discussion is usually framed: It is beyond dispute that TEs are part of scientific practice in core disciplines such as physics or biology. Most commentators also agree that at least some scientific TEs are successful in generating knowledge about the empirical world. But how is this possible? Since, by definition, they are experiments *in thought*, TEs do not seem to establish direct cognitive contact with the world. This makes the success of TEs indeed puzzling: How can we learn something about the world by merely thinking about it? It has become common to refer to this question as the *paradox of TEs* (c.f. Horowitz & Massey 1991, 1).

Much of the contemporary discussion about scientific TEs is devoted to the task of solving this paradox. Hence, the existence of apparently successful TEs is usually treated as the explanandum to which suitable theoretical explanations are directed. On one such theory, the success of (some) TEs is explained by means of a Platonist construal of the laws of nature and by the ancillary epistemological thesis that TEs allow us to directly grasp these laws with our mind's eye (e.g. Brown 2004, 2011). Another theory explains the success of TEs by declaring them to be arguments in disguise. On this view, scientific TEs lead us to knowledge because they are, in reality, chains of inductive and/or deductive inferences (e.g. Norton 1996, 2004). On a third popular account, the success of TEs is explained by means of a psychological theory according to which the manipulation of mental models allows us to trigger knowledge formations that aren't cognitively available otherwise (e.g. Mišćević 1992, Nersessian 1993).

A good deal of ink has been spilled over these proposals and I will add to this discussion below. For the moment, however, I want to emphasize that much of the plausibility of each of these theories depends on quite substantial metaphysical, epistemological and ontological assumptions whose scope goes well beyond their actual subject matter. Take, for instance, Platonism: The plausibility of a Platonist explanation of TEs largely depends on our willingness to accept a rationalist epistemology, an ontology that includes universals and the view that the laws of nature are contingent necessitation relations between universal properties. If one accepts these assumptions, then Platonism is a perfectly good explanation for the existence of TEs in science history as well as of their apparent success. If, on the other hand, one thinks that we should be parsimonious in ontological and metaphysical matters and that sense perception is the only source of knowledge about the world, then one will likely opt for one of the empiricist alternatives to Platonism. Hence, although it would be too strong of a claim that the available theories on TEs are evaluated *only* on the basis of their commitments to certain metaphysical and epistemological stances, it is at least safe to say that these commitments play a decisive role in how the discussion normally evolves. And, supposedly, it is also due to these commitments that certain strands of the contemporary debate on TEs appear to be a mere sideline of the age-old battle between rationalism and empiricism.

Now, my point is not to deny that the TE-debate may have implications for the quarrel between rationalists and empiricists. Since it is at least initially plausible to regard TEs as *prima facie* cases of knowledge acquisition *minus* sense perception, the onus seems to be on philosophers with empiricist leanings to come up with a deflationary account on TE-reasoning. Rather, my point is that, if we wish to address TEs phenomenologically,

treatment may either refer to Crowell (2006) for a highly instructive introductory essay or to Smith (2007) for a more detailed, book-length study.

the issue should be approached in a way that is not already contaminated with pre-established assumptions concerning metaphysical, ontological and epistemological matters. To construct one's theory on the basis of *either* rationalist *or* empiricist intuitions not only clashes with the phenomenological demand to proceed from an unbiased description of the phenomena; it also harbours the danger of begging the question against the respective opponent, of distorting the subject matter before the actual analysis can even begin and of ending up in a stalemate between irreconcilable epistemological and metaphysical systems.

Hence, what sets phenomenology apart from other theoretical alternatives is, first of all, its point of departure: Instead of treating TEs as explananda that call for theoretical explanations whose credibility depends on more general metaphysical, ontological and epistemological intuitions, phenomenology actively brackets such intuitions and proceeds from an unbiased description of the *phenomenon* of TE-reasoning instead. In other words: The main objective, at least at the initial stage, is not to construct a theory that seeks to bring TEs in line with certain pre-established standpoints. The main objective of a phenomenological account is rather to give a faithful description of the *actual performance* of TEs from a first-person perspective and to go on from there.

2. Filling Up the Phenomenological Toolbox

As I have pointed out, phenomenology differs from other theories in its basic approach: A phenomenological account on TEs brackets all kinds of metaphysical, epistemological and scientific assumptions and proceeds from a faithful description of the *phenomenon* of TE-reasoning. However, what does this mean exactly? Where do we have to start if we wish to describe the performance of TE-reasoning from a first-person perspective? How do we identify certain mental episodes as TEs? And how could we possibly tell if these mental episodes embody instances of successful knowledge acquisition?

In the face of these questions, my strategy in the remaining parts of this paper will be as follows: Since I am interested in the question of whether TEs fall under the rubric of successful knowledge acquisition, I will have to say a few words on a phenomenological theory of knowledge first. It is mainly through the discussion of examples that I will introduce five concepts that are crucial for a phenomenological understanding of knowledge acquisition in the perceptual realm. These concepts will form the background against which I will then take a closer look at the imagination,³ the medium in which TEs are performed. My aim is to highlight the similarities as well as the differences between perception and imagination in order to elucidate how knowledge is generated through the use of TEs.

³ "Imagination" is an umbrella term that is notoriously hard to define (cf. e.g. Kind 2013). In what follows, I will take it to denote a type of mental event in which we are intentionally directed towards imagined objects or scenarios and not, for instance, towards propositions. Hence, I am distinguishing between acts of imagining (in which imagined objects or scenarios are given in a quasi-sensory manner) and acts of conceiving (in which the objects of our intentional directedness are propositions). It should be noted, however, that my terminology departs from Husserl's who uses the notion "phantasy" in order to denote what I call "imagination".

2.1 Fulfillment and Frustration

What is knowledge? When are we dealing with cases of successful knowledge acquisition? In line with the general tenets of the phenomenological research program, these questions, too, are to be answered by way of a description of those types of conscious experience in which the relevant distinctions become evident. Take the following two cases as a starting point: 1) I am judging that my bike is in the office while I am still in the cafeteria. 2) I am judging that my bike is in the office while I am standing right in front of it. Phenomenologically construed, these two cases have a lot in common. To begin with, both experiences are experiences *of the same type*, namely experiences of judging (in contrast to experiences of doubting, imagining, fearing etc.). In addition, both experiences also have the same intentional object, namely my bike in the office. But, clearly, there is a striking difference as well: While the judgement about my bike's whereabouts may just be a wild guess as long as I am still in the cafeteria, the direct acquaintance with my bike (i.e. my standing right in front of it) warrants the judgement about my bike's location beyond all reasonable doubt.

The acknowledgement of this difference allows us to make a first step towards the notion of *fulfillment*, one of the main pillars of phenomenological epistemology: If – like in the first case – I am directed towards an object in its absence, then my intention towards the object is *empty*. If, on the other hand, I am directed towards an object in its actual presence – if the object is given in its “bodily’ selfhood” (leibhaftige Selbstheit) (Husserl 1983, pp. 9-10; translation modified) –, then my intention towards the object is *fulfilled* by the presence of that very object. Fulfillment, i.e. the congruence between the object as it is emptily intended and the object as it is intuitively given, is the ideal limit towards which our judging strives, “the measure of its success or failure” (Crowell 2006, 14). If this kind of congruence not only takes place, but is also registered by a cognizing subject, then we are dealing with what Walter Hopp aptly calls “knowledge at its best” (Hopp 2011, chapter 7).⁴

Let me summarize: *Fulfillment* takes place if there is a relation of congruence between an object as it is emptily intended and the object as it is intuitively given. Conversely, *frustration* occurs when the intuitive experience of an object is non-congruent with the empty intention towards that same object. In an ideal case of fulfillment, I realize that the object is exactly like I thought it would be. In cases of frustration, the intended object turns out to be different than initially thought. On a phenomenological view, all epistemic concepts such as evidence, justification or falsification ultimately lead back to the basic concepts of fulfillment and frustration.

2.2 Horizon and Anticipation

The concepts of fulfillment and frustration are crucial for the phenomenological understanding of knowledge and knowledge acquisition. Fulfillment, i.e. the congruence

⁴ It must be noted that the concept of fulfillment is not restricted to the realm of perceptual experiences. Although it is true that veridical perceptions are model cases of fulfillment (cf. e.g. Husserl 1983, 5-6, 82-83, 154, 327), fulfillment is a functional concept that goes along with a functional object concept. This is to say that fulfillment takes place whenever an empty intention towards an object is in congruence with a fulfilling intention towards the same object. However, as Husserl makes clear, “an object [...] may as readily be what is real as what is ideal, a thing or an event or a species of a mathematical relation, a case of being or a what ought to be” (Husserl 2001a, p. 145).

between an object-as-intended and the object-as-given, is the ideal limit to which all of our knowledge-claims aspire. On closer inspection, however, this turns out to be only part of the story, especially – but not only – with respect to the acquisition of empirical knowledge. Consider again the example of the veridical perception of my bike. At first glance, the direct perceptual givenness of my bike seems to clearly fulfill the judgment that my bike is in the office. But a more accurate description reveals that what is really experientially given in this situation is not simply *my bike*, but only *one single profile of my bike*, its current front side. To be sure, I could alter my position and make the current backside the new front side, and vice versa. But this doesn't change the fact that my bike is always given *in perspectives* and that, more generally, things always and necessarily have more parts, functions and properties than can be actualized in a single intentional act. My bike – as it is intended – is *transcendent*, not only in the sense that it can be seen from indefinitely many more perspectives than I can take up at a given point in time. It is also transcendent in the sense that it has, for instance, a momentarily hidden internal structure, a history, certain practical functions or many properties that aren't in the center of attention right now.

So, a closer look at how things appear to us reveals that our intentions towards these things always “transcend” or “go beyond” the actual experiences that give rise to them. As the example of my bike shows, there is a describable discrepancy between what is meant through a particular intentional act (my bike over there) and what is experientially given (my bike's facing side with its momentarily visible features). Phenomenologically construed, this discrepancy does not represent a problem that must be somehow remedied, e.g. by proposing a theory that explains how a number of seemingly disconnected profiles add up to a homogeneous thing to which we then attribute these profiles. The fact that our intentions towards things always transcend the sphere of intuitive givenness is rather to be treated as a phenomenologically discoverable feature of experience itself: Intending is, as Husserl puts it, always and necessarily an “*intending-beyond-itself*” (Husserl 1960, 46). In being intentionally directed towards material things, we “know” that there is more to them than is revealed in one single glance. This “knowing” is no matter of inferential belief or judgement over and above the experiences in which things are perspectively given; it is rather an essential part of any such experience.

The important lesson to draw from these considerations is that “[e]ach individual percept is a mixture of fulfilled and unfulfilled intentions” (Husserl 2001b, 221). Or, to put it in an alternative terminology: Intentional experiences are always embedded in implicit *horizons* of intentions that are momentarily unfulfilled, but that could be fulfilled in the course of further acts. Even though I can now only see my bike's facing side with its momentarily visible features, my bike appears to me as something that could be explored more fully. I “know” that I could alter my vantage point and explore its momentarily hidden back side. I “know” that I could look more closely and explore its surface in more detail. I “know” that I could cut the frame in half and explore its internal structure. It is these and indefinitely many other potentialities that add up to the implicit horizon against the background of which singular intentions towards things always stand out.

Phenomenological descriptions reveal that intentional acts towards things always point to implicit horizons of empty intentions. Hence, on a phenomenological view, experience is never exhausted by what is actual; experience is always already saturated with

implicit references to future experiences that are possible insofar as they could be actualized in the course of further acts. However, in order to get a more complete understanding of the phenomenological notion of horizon, two more aspects have to be stressed: The first thing to note is that an act's horizon is, as I have already indicated, *implicit* to the act itself. This is to say that the horizon is no theoretical construct that is *retrospectively* ascribed to the initial act. An act's horizon is rather *co-given* with the initial act even though we usually aren't aware of this. That my bike has a backside isn't something that can only be asserted after I have changed my vantage point. It is also not something that is the product of some sort of inferential process. Rather, it is something that belongs to the very meaning of being intentionally directed towards material things.

The second important aspect is this: On the basis of what I have said so far, one could define horizons as sets of empty intentions against the background of which particular fulfilled intentions always and necessarily stand out. But this definition is somewhat misleading: Although it is correct to say that a horizon consists of empty intentions and thus can be described as a "halo of emptiness", it is crucial to stress that "this emptiness is not a nothingness", but rather that "the sense of this halo [...] is a *prefiguring* that *prescribes a rule* for the transition to new actualizing appearances" (Husserl 2001c, 42; my emphases). What Husserl is saying here can be elucidated with the help of my earlier example: If I perceive my bike, my intention towards the bike's facing side is conjoined with a horizon of empty intentions and thus with the anticipation that there is *more* to the bike than is revealed in one single act. However, this *more* is far from being indeterminate: In perceiving my bike, I implicitly anticipate concrete courses of experiences that are compatible with what was originally intended. The perception of the greyish blue of my bike's facing side, for instance, comes with the implicit anticipation that the momentarily hidden backside will exhibit the same color as well. To be sure, it is possible that this anticipation is frustrated by future experiences. But if I were to find out that my bike's backside isn't greyish blue, but coated with diamonds and rubies, then I would probably start to wonder whether I am really dealing with *my* bike at all.

So, typically, horizons are not indeterminate in the sense that a given act points to the entirety of acts that are logically compatible with the initial act. Horizons are rather *structured*: they prescribe implicit rules of anticipation that restrict how things could appear if the corresponding experiences were still to qualify as experiences of the same thing.

2.3 Background Knowledge

The point of the previous considerations is that our experience of things is "thoroughly interwoven with anticipations" (Husserl 2001c, 47). Things are always and necessarily given in horizons and these horizons prescribe rules through which we anticipate the course of future experiences. Under normal circumstances, these rules remain implicit and largely unnoticed. It is mainly in cases of frustrated anticipations that their existence even becomes manifest.⁵

⁵ Here is an example that illustrates the issue: Imagine you are asked to bring a shopping bag into the kitchen. You go to the car, you see the bag in the trunk, you get ready to lift it and – Oops! – your arms go up way too easy and you almost topple over. What happened? Well, the bag is significantly lighter than you anticipated because it is filled with paper towels. What is remarkable about this case is, first of all, the funny feeling that you had when you lifted the bag. Since you have lifted many light items before, this

Given the view outlined so far, an obvious question arises: Where do the aforementioned anticipations come from? How are they generated? And what is their status? As David Woodruff Smith and Ronald McIntyre have pointed out (Smith & McIntyre 1982, chapter 5; Smith 2004, chapter 5), both horizons and the rules of anticipation that are given through them are dependent on highly heterogeneous stocks of background knowledge that, in their totality, make up our fundamental background image of the world. In large part, this background knowledge is empirical and thus highly contingent in nature. Just think of the earlier example of my anticipation that my bike's backside will exhibit the same color as its facing side: Clearly, my anticipation in this case depends on the fact that I've had countless previous encounters with my bike. Hence, a more determinate horizon with more determinate anticipations is co-given with *my* experiences of *my* bike than would be co-given with *your* experiences of *my* bike.

Other parts of our background knowledge are not empirical, but theoretical in nature: Psychologists conducted experiments in which students were asked to observe the falling of two objects (a metal sphere and a plastic sphere of the same diameter) and to record their observations (Gunstone & White 1981). Students who initially held that heavier objects fall faster were much more likely to report observations that supported their theory. Experiments such as these suggest that our anticipations of further courses of experience are sometimes strongly influenced by theoretical components of our background knowledge.

Finally, as phenomenologists typically stress, there is a third type of background knowledge that belongs to neither of the two aforementioned categories: Take, for instance, the anticipation that my bike will exhibit, not a particular color on its backside, *but rather some kind of rear side at all*. Since we fail to even imagine a material thing that does not exhibit a backside *of some sort*, phenomenologists claim that the corresponding anticipation is not governed by a contingent piece of background knowledge; it is rather governed by a "*necessary la[w]* which determine[s] what must necessarily belong to an object in order that it can be an object of this kind" (Husserl 1973, 352). That material things, unlike, say, geometrical objects, are necessarily given in perspectives is determined by "a universal essence which [...] *prescribes an intellectually seen generical rule* for every particular object becoming intended to in multiplicities of concrete mental processes" (Husserl 1983, 341).

3. Perception and Imagination: The Return of the Paradox

My approach in the previous sections was to use examples from the perceptual sphere in order to introduce five phenomenological key concepts. Yet, on closer inspection, the applicability of these concepts is not restricted to the perceptual realm. Consider, for example, "horizon" and "anticipation": As I have pointed out, material things are necessarily given in horizons through which we anticipate further courses of experience. Horizontal givenness, however, is by no means exclusive to perception: "Imaginations [...] present their objects within exactly the same horizons" (Husserl 1973, 169). Imagine,

feeling is certainly not a consequence of the fact that you lifted something light. What makes the experience stand out is rather how it compared to your (frustrated) anticipations. What is more, cases like these also highlight the fact that many of our anticipations correspond to pieces of non-propositional, practical and sensorimotor knowledge about how things are done and how we use our bodies to do them (cf., e.g., Husserl 1997; Noë 2004).

for instance, Bart Simpson writing “No one cares what my definition of ‘is’ is” on Mrs. Krabappel’s blackboard. If you do so, then the object of your attention is a yellow cartoon character that is scribbling on a blackboard. Yet, co-given with this initial experience is a set of empty intentions through which you anticipate further features of the imagined scenario. Although they aren’t in the centre of your attention right from the start, you automatically anticipate further features such as the spikiness of Bart’s hair, the wall behind the blackboard or the color of Bart’s pants. If the focus of attention shifts and objects from the margin move to the centre, certain other objects fade into the background, without, however, disappearing completely. Like in the previous examples from the perceptual sphere, you “know” that there is more to the imagined objects than can be grasped in one single act.

Horizontal givenness is not the only commonality between perception and imagination. Remember the crucial distinction between empty and fulfilled intentions: If I am merely thinking of my bike, then my intention towards my bike is empty. If, on the other hand, I perceive my bike, my intention towards the intended thing is intuitively fulfilled by the thing’s presence. But is perception the only means by which the empty intention towards my bike can be fulfilled? Not at all. If I imagine my bike, then this, too, fulfills the mere thought of my bike – what was empty before (the bike as the object of my thought) is now fulfilled by the quasi-experience of my imagined bike. Hence, perception and imagination reveal a close parallelism not only with respect to the structural characteristic of horizontal givenness: Since “[t]he objective intention directed toward the imagined object has its filling in the experienced phantasms, just as the objective intention in perception has its filling in sensations” (Husserl 2005, 93; translation modified), the parallelism between perception and imagination involves the crucial concept of fulfillment too.

There are, as we have seen, important essential features that are common to perception and imagination. But, of course, there are crucial differences as well. The most obvious concerns the fact that, unlike imagination, perception is an “*originally presentive* mode [of consciousness]” (Husserl 1982, 327). What this means can again be illustrated with recourse to the earlier example: If I perceive my bike standing in my office, then this act is *self-giving* with regard to actual things (my bike) and states of affairs (that my bike is standing in my office). Thus, it is not only the case that the perceptual givenness of my bike intuitively fulfills the thought that my bike is in my office. Even more importantly, the perceptual act “gives its object itself in the flesh” (Husserl 2001c, 140) and thus exhibits a particular quality that is discussed in the contemporary analytic literature under labels such as “presentational feel” (Foster 2000, 112), “scene-immediacy” (Sturgeon 2000, 24) or “presentational phenomenology” (Chudnoff 2013, chapter 1.2). On a phenomenological view, this quality is a main reason why we take perceptual acts to confer justification on empirical beliefs. I see my bike leaning against the bookshelf and it instantly seems to me that what I see is a truthmaker for the proposition “My bike is in the office”. The fact that the visual experience of my bike instantiates the quality of having a “presentational feel” explains why this is so.

But now compare this with the case of imagination: If I imagine my bike standing in the office, then this, as I have pointed out, also counts as a case of fulfillment: What was empty before is now fulfilled by the quasi-sensory experience of my imagined bike. However, clearly, the quasi-experience of an imagined bike is no justifier for the belief that my bike is in my office. Phenomenologically construed, this is because imaginative

quasi-experiences do not instantiate the property of having a “presentational feel” with respect to actual things and states of affairs. While “[p]erception makes a present reality appear to us as present and as a reality [...], [imagination], on the other hand, lacks the consciousness of reality in relation to what is [imagined]” (Husserl 2005, 4). It is for this reason that the imagination does not count as a direct source of justification of empirical beliefs.

Perception is an act that gives its object as actual. An act of imagination gives its object as non-actual. This, in a nutshell, is the reason why the imagination does not count as a direct source of empirical justification. But there is a second, even more fundamental difference that concerns the conditions under which anticipations and expectations⁶ are frustrated in each respective realm. Here is an example that illustrates the point: Suppose that I perceive the greyish blue of my bike’s facing side. And suppose furthermore that – for whatever reasons – the horizon in which this perception is embedded comes with the anticipation that my bike’s backside is coated in pure gold. What will happen? Since a bike that is both mine *and* ridiculously expensive isn’t part of the inventory of the actual world, this anticipation will be frustrated by further perceptual experiences. And, obviously, there is nothing I can do about this: By and large, our perceptual experiences are not under our voluntary control; their occurrence as well as their content is systematically fixed by our external environment.

But things are entirely different in the imagination. Suppose that I imagine my bike. And suppose furthermore that the horizon in which my imagined bike is given comes with the anticipation that its backside is coated in pure gold. Obviously, nothing prevents me from imagining my bike in a way so that the anticipation of a gold-coated backside would be intuitively fulfilled. Generally speaking, we know that we can do all sorts of crazy things in the sphere of imagination: I could imagine my bike as almost transparent or tall like a mountain. I could imagine that it is accelerated to 90% of the speed of light. Or I could imagine that my bike transmutes into a honeybee and flies away. It is in my freedom to project these and indefinitely many other scenarios with the help of my imaginative capacities. Hence, while the persistency of anticipations and expectations is strictly regulated by the external world in the perceptual sphere, the imagination seems to be distinguished by “its *optional character* [and] therefore, speaking ideally, [by] its unconditioned arbitrariness” (Husserl 2005, 642). “[I]t remains”, as Husserl puts it elsewhere, “within the province of our freedom to allow the indeterminateness of [...] horizons to be quasi-fulfilled in an arbitrary way by imagining.” (Husserl 1973, 171)

Given this rough-and-ready comparison, one could come to the following conclusion: Imagination is like perception in many ways. But beneath the surface of similarities, two crucial differences remain: First, perception gives its objects as actual. The objects of imagination, on the other hand, exhibit the “characteristic of inactuality” (Husserl 2005, 320). Secondly, perceptual anticipations and expectations are fixed by the external

⁶ A word on terminology: Anticipations are non-independent parts of perceptual experiences. If, for instance, I enter the kitchen and flip the light switch, the experience of the flipping of the switch comes with the tacit anticipation that the light will go on. This anticipation is neither an act in its own right nor is it consciously experienced in the normal course of events. However, anticipations can become the basis for separate acts of *expectation*. If, for instance, I have just repaired the light switch and check it for the very first time, the flipping of the switch may come with the expectation that the light will go on. In this case, the expectation is an act in its own right and has the status of an explicit prognosis of future events.

world. Imaginative anticipations and expectations, on the other hand, seem to be essentially unrestricted.

If we take this result as conclusive, then this also has serious consequences for the phenomenological understanding of TEs. In particular, the paradox of TEs seems to return within the framework of phenomenology in its original force: Scientific TEs are supposed to justify beliefs about the empirical world. But since TEs rely on our imaginative capacities, and since the imagination is no direct source of empirical justification, we are in need of a plausible story about how TEs could possibly achieve this aim. But whatever story we tell, we will have to confront the following worry: The principal reason for regarding physical experiments as epistemically significant is that our expectations towards them can be and quite regularly are frustrated by perceptual experiences whose occurrence and content is not controlled by us. To put it bluntly: When we perform physical experiments, reality sometimes “kicks back” at us. And if it does, we learn something utterly important, namely that the world is different from what we had expected. But what could possibly “kick back” at us in TEs? Since TEs exploit our imaginative capacities, and since, apparently, these capacities are characterized by the unconditional freedom to imagine whatever we like, the conditions of fulfillment and frustration in TEs appear to be completely under our control. If that is the case, however, then it is hard to see how TEs could possibly qualify as a serious method of knowledge acquisition. Viewed in this light, TEs are similar to children’s games. Of course, they are fun to play. But since their rules can change erratically at any time during the process, they are also somehow pointless from an adult’s perspective.

4. Three Types of Constraints

Here is the upshot of the previous section: In order to solve the paradox of TEs, we need to tell a story about why the imagination can be regarded as a source of stable evidence. But there is an obstacle to any such story: One of the principal reasons for regarding perception as epistemically valuable is that what is perceived is not entirely up to us. The occurrence as well as the content of our perceptual experiences is systematically fixed by our external surroundings. Imaginings, on the other hand, seem to lack external constraints: Apparently, we are free to imagine whatever we like, whenever we like.⁷ Hence, the conditions of fulfillment and frustration appear to be entirely under our control in the sphere of the imagination. But if this is true, then it is unclear how the imagination could possibly generate stable evidence.

On the basis of what has just been said, a skeptical attitude towards the epistemic value of imagination (and, consequently, of TEs) is fuelled by the view of imagination as being essentially unrestricted. But is this view correct? In what follows I will try to convince you that it isn’t. My aim in this section is to go through a couple of examples in order to make clear *that*, *how* and *to what extent* the course as well as the content of our imaginings are subject to three different types of external constraints. Getting clear on

⁷ Sartre seems to follow this line of reasoning when he comes to the conclusion “that it is impossible to find in the [mental] image anything more than what one puts into it; in other words, the [mental] image teaches nothing” (Sartre 2004, 103). But similar views can also be found outside of phenomenology: Consider, for instance, Ludwig Wittgenstein’s remark that “[i]t is just because imaging is subject to the will that it does not instruct us about the external world” (Wittgenstein 1980, §80). Or take Alan White’s more recent assessment that “one can’t be surprised by the features of what one imagines, since one put them there” (White 1990, 92).

the nature of these types will finally bring us closer to a phenomenological understanding of how knowledge is generated through TEs.

Suppose, that I ask you to imagine Bart Simpson writing “I will not use abbrev.” on Mrs. Krabappel’s blackboard. If you do so, then the object of your attention is a yellow cartoon character that is scribbling on a blackboard. Yet, co-given with this initial experience is the anticipation of further features of the imagined scenario: Although these features aren’t in the centre of attention right from the start, you automatically anticipate further aspects such as the spikiness of Bart’s hair, the purple wall behind the blackboard, Bart’s depraved behaviour or his red pants. But wait! While the sequence of imaginings evolves in more and more detail, you suddenly realize that something isn’t right. Even though you can’t quite figure out what it is, the whole scenario doesn’t feel like it should. It is only through careful reflection that you finally notice what bothered you: Bart’s pants are blue, not red!

What happened in the situation just described? To begin with, by following my instruction to imagine Bart Simpson, you immersed yourself in the quasi-world of “The Simpsons” and thus accepted certain limitations to your imagining. The concept “Bart Simpson” contains what has become known to you and your epistemic community about the kind of object in question. Hence, by applying the concept “Bart Simpson” in order to determine what you were about to imagine, you activated certain parts of your background knowledge that, in turn, motivated a horizon through which further quasi-experiences were anticipated. It is of course true that the imagination, unlike perception, gives you the freedom to produce quasi-experiences that are at odds with what is contained in the initial concept: You could either consciously decide to imagine quasi-experiences that are not compatible with the concept “Bart Simpson”. Or you could just inadvertently fail to stay within the boundaries of what the concept prescribes. But, in any case, if you imagine a white cartoon character that has a yellow bill and wears a sailor shirt with a red bow tie, then you obviously failed to immerse yourself into the quasi-world of “The Simpsons”.

There are further lessons to draw from this example: First, it underscores the fact that normally, when we employ our imaginative capacities, we are not imagining series of disconnected figments. Rather, most imaginative processes consist of sequences of related imaginings that, in their temporal succession, form a coherent whole. And, given the example above, it is easy to see where the coherence is coming from: it is ensured by the background knowledge that enters the imagination through the concepts with which we determine what we intend to imagine and from which our imaginative processes take their basic direction. This is not only true when we are asked to imagine Bart Simpson. It is also true when we are using our imaginative capacities to decide whether the sofa at Ikea will fit through the hallway at home. In both of these cases, the concepts through which we determine what we intend to imagine refer to chunks of background knowledge that, in turn, put constraints on how our imagining can evolve. To be sure, we could always break the coherence by imagining quasi-experiences that stand in no relation to what has been imagined before. But, given our practical ends at Ikea, imagining a sofa that suddenly transforms into a swarm of butterflies just isn’t the rational thing to do.

What the example of Bart Simpson also shows is that, secondly, frustration is by no means impossible in the sphere of the imagination. Other than in the perceptual realm,

however, anticipations and expectations are not frustrated by quasi-experiences alone. Imaginative anticipations are frustrated by way of a comparison between quasi-experiences and the background knowledge that enters the imagination through the concepts with which we determine the course of our imagining. This is exactly what happened in the example above: Initially, the content as well as the course of our imaginative endeavours were determined by the concept “Bart Simpson”. And it was through reflection on the background knowledge to which the concept refers that certain quasi-experiences could be singled out as being at odds with what the concept prescribes. Hence, there turns out to be a close relationship between the conditions of frustration on the one hand and the concepts that determine the course of our imagining on the other. Given this relationship, we can formulate an important conclusion: If we want to learn about the conditions of fulfillment and frustration in a given imaginative process, we have to look closely at the concepts through which the horizontal anticipations are determined.⁸

What I have tried to do so far in this section is take some first steps towards undermining the view according to which the imagination is epistemically inept due to the lack of external constraints to the conditions of fulfillment and frustration. Following my analysis, the imagination indeed gives us the freedom to produce random quasi-experiences that aren’t related to earlier imaginings. However, it is also possible to use our imaginative capacities differently: We can choose to immerse ourselves in a quasi-world by staying within the boundaries that are prescribed by the concepts through which we determine what we actually wish to imagine. And if we do so, the conditions of fulfillment and frustration are externally fixed; they are fixed by the background knowledge that enters the imagination through the concepts that give our imaginings their basic direction.

Assume for the moment that the foregoing analysis is correct. Even if it is, however, one could still wonder whether it really captures what is truly characteristic of scientific TEs. One could argue as follows: It may be true that the conditions of fulfillment and frustration are relatively well defined as long as we use the imagination in a fairly conservative fashion by sticking to concepts such as “Bart Simpson” or “bike”. Concepts like these prescribe clear rules of anticipation because they refer to rather uncontroversial chunks of empirical background knowledge that have been acquired through common experiential sources. In the case of “Bart Simpson”, for instance, we know what to imagine because we have spent many hours watching the respective TV show. But isn’t it characteristic of TEs to employ our imaginative capacities in a much more exceptional manner? Of course, at first glance, scientific TEs also seem to utilize fairly conventional concepts such as “bucket”, “car”, “train”, “cat” or “tower”. But the point of most scientific TEs is to imagine these objects under conditions to which the more common parts of our background knowledge do not apply. Our common background knowledge may prescribe clear rules of anticipation if, for instance, we imagine a car that is rushing towards a garage with 100 km/h. But what determines the

⁸ This, of course, is not only true of the imagination. Suppose that you see an object that looks like a barn. If you apply the concept “barn” to the seen object, then you will anticipate the object to have a backside that resembles its facing side. Accordingly, the experience of a backside that doesn’t resemble the facing side (for instance, a backside that is characteristic of fake barns) will frustrate your initial intention. If, on the other hand, you apply the concept “fake barn” to the seen object, a backside that doesn’t resemble the facing side is exactly what you anticipate. Hence, a close relation between concepts and the conditions of fulfillment and frustration is not only to be found in the sphere of the imagination, but also in the perceptual realm (cf. also Hopp 2011, chapter 2.1).

conditions of fulfillment and frustration if we are imagining a car rushing toward a garage *with 90% of the speed of light?*

Even a cursory look at the practice of TE-reasoning in science reveals that TEs indeed employ our imaginative capacities in a quite exceptional way. In order for most scientific TEs to be performed, we have to immerse ourselves in quasi-worlds that differ from our actual world in more or less drastic ways. As the following examples show, “immersion” consists in the active bracketing and/or modification of certain parts of our background knowledge: Immersion into the quasi-world of Galileo’s ship consists in bracketing the background knowledge according to which ships at sea are always subjected to rocking motions. Immersion into the quasi-world of Stevin’s chain consists in bracketing the background knowledge according to which objects moving down inclined planes are always subjected to kinetic friction and air drag. Immersion into the quasi-world of Newton’s bucket consists in bracketing the background knowledge according to which the material universe consists of more objects than just a water bucket and a rope.

As we shall see in the next section, the ability to be selective with regard to the background knowledge that determines the content as well as the course of our imaginings is indeed crucial for TEs in science. In the present context, however, this ability is still a reason for concern. Our problem was this: In standard cases of imaginative activity, the conditions of fulfillment and frustration are determined by the relevant parts of our background knowledge that have been acquired in the course of previous encounters with the world. But what determines the conditions of fulfillment and frustration in cases where the imagination takes us well beyond our ordinary experiential grasp?

The obvious answer is that much of what we anticipate and expect in scientific TEs is determined by the *theories* that work in the background of these TEs. The aforementioned car/garage-TE from Special Relativity is a case in point: If we are asked to imagine a car that is rushing towards a garage with 90% of the speed of light, then, obviously, our anticipations and expectations aren’t guided by previous perceptual encounters with this kind of scenario. What determines the course of our imaginings is rather a theory according to which the measured length of an object decreases noticeably at velocities close to c .⁹ The point is even more obvious considering examples such as Heisenberg’s gamma-ray microscope or EPR: In cases like these, not only the course, but even the content of our imaginings is almost entirely determined by theoretical components of our background knowledge. To put it bluntly: We just wouldn’t know what to imagine if we didn’t know a thing about quantum mechanics.

Many constraints that impinge on our imaginative activities are due to empirical components of our background knowledge. Others are due to theoretical parts of our background knowledge. This, in a nutshell, is the upshot of this section. However, phenomenologically construed, there is yet a third type of constraints that are neither empirical nor theoretical in nature. Here is an example that might help to make my

⁹ On closer inspection, the car/garage-TE is particularly interesting because it is not at all clear which part of Special Relativity should actually determine what we are supposed to imagine in this TE: On the one hand, Special Relativity tells us that objects moving with velocities close to c are Lorentz-contracted. On the other hand, we know since the 1930ies that Lorentz-contracted objects would not appear contracted, but rotated (Lampa 1924; Terrell 1959; Penrose 1959). This raises the question as to whether it is relevant for the success of TEs to imagine objects realistically (cf. Brown 2013).

point: Suppose that I ask you to imagine a brick that is thrown against a windowpane. And suppose furthermore that I ask you to bracket everything you know about the behaviour of bricks and windowpanes. Hence, what I ask you to imagine is a quasi-world in which bricks and windowpanes look like actual bricks and windowpanes, but in which their behaviour is absolutely unpredictable. Now the question is this: Are your anticipations and expectations concerning the behaviour of the brick and the windowpane indeterminate? Are the conditions of fulfillment and frustration entirely up to you?

At first sight, this question might seem odd. If you are serious about bracketing your background knowledge about bricks and windowpanes, then, apparently, there is nothing left that could possibly determine your anticipations and expectations in the scenario described above. You could imagine a quasi-world in which windowpanes are shattered upon impact. But you could also imagine a quasi-world in which it is the bricks that crumble to dust whenever they hit a windowpane. Or you could imagine a quasi-world in which windowpanes transmute into sprays of flowers whenever they are hit by bricks. In short: After bracketing your background knowledge, no anticipation is too outlandish to be in principle unfulfillable by matching quasi-experiences. Or so, at least, it seems.

In fact I think this view is descriptively wrong. There are restrictions to what we can or cannot imagine, even after we have bracketed our entire background knowledge about the objects in question. Consider, for instance, the anticipation of a windowpane that is completely shattered and completely unshattered, both at the same time. As you can easily verify by trying to actually imagine this scenario, it is impossible to intuitively fulfill this anticipation. Or consider the anticipation of a brick that is red and blue all over. Here too, we fail to imagine a fulfilling quasi-experience. Or consider the anticipation of a brick that is seen from all sides at once or whose redness can be smelled. Again, it is impossible to imagine a quasi-world in which these anticipations would be fulfilled by matching quasi-experiences. In all of these cases the problem is not that a particular quasi-experience contradicts a rule of anticipation that suggests itself in the light of previous encounters with the actual world or in the light of a theory. Rather, the problem is that certain quasi-experiences are impossible to imagine, no matter how hard we try. In all of the aforementioned cases we have to realize that the “freedom [of voluntarily producing quasi-experiences] is limited insofar as essential laws of possible *quasi*-fulfillment are inherent here within the boundaries of the unity of an identical possible objectivity understood as intentional and still indeterminate” (Husserl 2005, 671). Hence, the content and course of our imaginings is not only determined by empirical and theoretical components of our background knowledge. Phenomenologically construed, our imaginings are also constrained by “*necessary laws* which determine what must necessarily belong to an object in order that it can be an object of this kind” (Husserl 1973, 352).

In a recent paper, Walter Hopp has drawn particular attention to the role this kind of essential knowledge plays in TE-reasoning. One of Hopp’s main theses is “that thought experiments, at their best, are in fact founded on acts of fulfillment in which we intuit universals and the relations among them, and that the actual instantiation of those universals and relations is immaterial” (Hopp 2014, 81). In advancing this claim, Hopp emphasizes the similarities between Brown’s Platonism and a phenomenological theory of TEs. According to Hopp’s analysis, both accounts agree in their acknowledgment of a

knowledge-yielding faculty that is independent of the five senses and that is operative in at least some TEs.¹⁰

What I have tried to do in this section is to undermine the view of the imagination as essentially unrestricted. If my analysis is correct, then our imaginative capacities are subject to three types of constraints: First, there are constraints that are due to the empirical parts of our background knowledge. Secondly, there are constraints that result from the theories that we have incorporated into our background knowledge. Thirdly and finally, there are constraints that are due to “essential laws which govern acts as *intentional* experiences, in all their modes of sense-giving objectivation, and their fulfilling constitution of “true being” (Husserl 2001b, 319).

5. Putting the Pieces Together

With the above findings in place, we can finally turn to the heart of this paper. Drawing on the results from the previous sections, I will now take a first stab at a phenomenological description of TE-reasoning in science. On my view, TE-reasoning is a process that occurs in three stages: the preparatory stage, the performance stage, and the conclusion stage. I will comment on each of these stages in turn.

5.1 The Preparatory Stage

Before we can even begin to perform a TE, we need to know a couple of things. First of all, and most obviously, we must know to what end the TE is performed. This is to say that we need a sufficiently clear grasp of the target thesis that the TE is meant to refute, to corroborate or to clarify. Secondly, and no less importantly, we need information about the kind of scenario we are supposed to imagine. This information is encapsulated in the TE-narrative in which the details of the TE-setup is specified. Of course, TE-narratives differ significantly with respect to detail, complexity and style. Some of them are text-only. Others also make use of visualizations. Some of them are colourful and filigreed. Others are strictly technical and reduced to the bare essentials. In any case, however, TE-narratives pursue a twofold task: On the one hand, TE-narratives tell us something about the objects of which the TE-scenario is composed. On the other hand, TE-narratives specify the characteristics of the quasi-worlds in which TE-scenario is to be imagined. Consider, for instance, the following excerpt from a TE-narrative from a standard physics textbook:

“A man has an $l = 5$ m long garage and buys and $l_0 = 7$ m long car (proper length of the car). He reasons that, if he drives sufficiently fast, the car will fit in the garage due to length contraction (ignore the fact that he is going to ruin his new car by smashing it against the garage wall).” (Faraoni 2013, 25)

¹⁰ It must be noted, however, that Hopp is well aware of the differences between a phenomenological theory of TEs and Brown’s (cf. Hopp 2014, 89-90). While phenomenologists will typically not be at odds with the rationalist part of Brown’s story – particularly with his claim that sense experience is not the sole source of knowledge about the world –, they will most certainly take issue with his construal of the laws of nature and with his Platonist two-world ontology. Hence, building on Davies’ useful distinction between extreme and moderate deflationism on the one hand and (two kinds of) moderate and extreme inflationism on the other (Davies 2007, 37-42), phenomenology positions itself somewhere between (the second kind of) moderate and extreme inflationism.

In order to understand this TE-narrative, we must have different types of background knowledge at our disposal: While concepts such as “car” or “garage” refer to conventional chunks of empirical background knowledge, concepts such as “proper length” or “length contraction” refer to information about the kind of physics that forms the theoretical framework of the TE. It is only if these pieces of information are part of our general background knowledge that we are able to set up the TE accordingly. If, for instance, knowledge about Lorentz contraction is lacking, then we inevitably fail to set up the TE according to the intentions of its presenter.

But TE-narratives not only give us information about the kinds of objects and processes that we need to imagine in order to carry out the experiment. TE-narratives also contain information about the quasi-world in which the TE ought to be performed. While this information is only implicit in many cases, the above-quoted TE-narrative hints at the specifics of the required quasi-world by remarking that no actual car would remain intact under the imagined conditions. Of course, we know that no actual brake system could decelerate the car from velocities close to c . And we know that the car would melt in less than a blink of an eye. But these and similar other components of our background knowledge must be bracketed if we want to immerse ourselves into the quasi-world that suits the TE we are about to perform. Like many other TEs in physics, the car/garage-TE just wouldn't work in a quasi-world that resembles our actual world too closely. The TE only works in an idealized quasi-world in which many factors that determine the behaviour of the actual world are neglected.

5.2 The Performance Stage

The purpose of a TE-narrative is to prime its audience for the performance of the ensuing TE. The main tool for doing so is the concepts of which the narrative is composed. These concepts refer to the kind of background knowledge that is necessary in order for the TE to be set up. It is only if the right kind of background knowledge is already in place that we are in a position to imagine the TE-setup as well as the quasi-world in which the TE ought to be embedded. Hence, I fully agree with David Gooding's observation that “[a] TE becomes possible [only] when a world is sufficiently well-represented that experimental procedures and their likely consequences can be described within it” (Gooding 1992, 281). Whether or not a TE manages to lead us to its desired outcome largely depends on our familiarity with the quasi-world in which the TE ought to be embedded according to its presenter.¹¹

¹¹ This point also has implications for the historiography of TEs. As historically inclined philosophers such as James McAllister (1996) or Paolo Palmieri (2005) have emphasized, it is common in the philosophical literature to make use of the history of TE-reasoning in a rather idiosyncratic, sometimes even “cartoonish” (Palmieri 2005, 223) manner. Not only that case studies are presented without paying sufficient attention to their context; many philosophers also approach the history of TE-reasoning in an overtly “presentist” or “Whiggish” way. In the case of Galileo's famous tower-TE, for instance, it has been objected that it is only if *we today* “look at this thought experiment from an historically distant perspective and with the knowledge of modern physics concerning falling bodies in a vacuum [that] the inference from the contradiction to the ‘right’ conclusion [...] seems to be immediate and untutored by any empirical or logical reasoning” (Brendel 2004, 95; cf., for a similar complaint, Norton 1996, 344-345). What seems right to me about this objection is the almost trivial truth that we cannot presuppose our contemporary background knowledge if we wish to determine the epistemic weight that a TE carried in its original historical context. However, I also believe that the framework of phenomenology is particularly well-suited to draw a meaningful distinction between the epistemic weight a TE carries for us, relative to our modern background image of the world, and the epistemic weight a TE originally carried for a given

Assume that we have successfully completed the preparatory stage in our current example, the car/garage-TE. This is to say that we have a sufficiently clear grasp of the target thesis, we possess the right kind of background knowledge and we understand the specifics of the quasi-world in which the TE-scenario ought to be imagined. We are now in a position to imagine the required TE-setup. If we do so, then the primary object of our intentional directedness is a scenario that consists of an imagined car that is about to rush towards an imagined garage at a speed close to c . However, on the basis of our earlier reflections on the horizontal structure of intentionality, it is clear that the imagined car and the imagined garage is by no means all that is given to us in this situation. Co-given with the quasi-experience of the primary object is an array of empty intentions that point towards possible future states of the imagined TE-scenario. In imagining the TE-setup, as it is laid out in the TE-narrative, we automatically co-intend a horizon that prescribes anticipatory rules concerning the course of further quasi-experiences. These rules of anticipation are motivated by the background knowledge that is necessary for imagining the TE-setup in the first place. And it is through these anticipatory rules that our background knowledge restricts how the imagined scenario could evolve if the corresponding quasi-experiences were still to qualify as experiences of the initial scenario.

The point I am trying to make is this: If we imagine a particular TE-setup, we are not merely imagining a particular arrangement of imagined objects that are embedded in a particular quasi-world. Since the objects of our intentional directedness are necessarily given in horizons, we also co-intend rules of anticipation that restrict how the TE-setup could evolve if the ensuing phases were still to qualify as phases of the initial setup. This way of putting things also makes clear what it actually means to carry out a TE: To perform a TE is to immerse oneself into a particular quasi-world, to imagine a TE-setup within this quasi-world and then to let the TE-setup evolve according to the anticipatory rules that are co-given with each and every quasi-experience of the unfolding TE. Or, to put it in slightly different terms: To perform a TE is to “live through” the anticipatory horizons against the background of which fulfilled intentions towards particular phases of the imaginative process always stand out.¹² With each new phase a new horizon comes to fruition, “a new system of determinable indeterminacy, a new system of progressive tendencies with corresponding possibilities of entering into determinately ordered systems of possible [quasi-]appearances, of possible ways that the aspects can run their course, together with horizons that are inseparably affiliated with these aspects” (Husserl 2001c, 43). In this whole process of “living through” the anticipatory horizons in which each phase of the imaginative process is given, processes of explicit inferential reasoning are mostly absent. This explains one of the most intriguing features of scientific TEs, namely the ease with which they are performed and the apparent effortlessness with which we reach the desired outcome.

According to view developed so far, the way the initial TE-setup pans out is predelineated by the rules of anticipation that come with every new quasi-experience of the unfolding TE. This, however, raises an obvious question: At any point during the performance of a TE, there are far more empty intuitions towards future quasi-

historical community. Cf., for a couple of first steps towards employing the phenomenological tool of the “epoché” in dealing with science history, Arabatzis (2012) and Palmieri (this volume).

¹² This echoes Gooding’s remark that “[t]o explain the force of an experiment it helps to understand it as a process to be worked through, rather than as a logical structure” (Gooding 1992, 283).

experiences than are relevant for the actual outcome of the TE. If we imagine a car rushing towards a garage, for instance, we implicitly anticipate the car to have a backside. Or we anticipate the garage to have a specific color. But, obviously, anticipations such as these are entirely irrelevant for what the TE is supposed to show. Hence, the question arises as to how relevant anticipations are distinguished from irrelevant ones.

I opened this section with the remark that the aim of TE-narratives is to prime their audience for the ensuing TE. In the light of the question that has just been raised, this remark must be further clarified: On the one hand, TE-narratives tell us what we are supposed to imagine. They do so by employing particular concepts that activate particular components of our background knowledge. Anticipatory rules that are motivated by these components then restrict how the ensuing TE can unfold. But on the other hand, and equally important, TE-narratives also give us information about what we are *not* supposed to imagine (cf. Davies 2007, 35). In part this is done through the target-thesis that automatically narrows our focus to certain aspects and leaves out others. But it is also done through the determination of the quasi-world in which the TE must be embedded. By bracketing certain parts of our background knowledge in order to immerse ourselves in a particular quasi-world, we prevent certain rules of anticipation from even becoming operative. Our normal anticipations concerning objects sliding down inclined planes, for instance, are determined by the background knowledge according to which such objects are always subjected to kinetic friction and air drag. However, as reliable as this background knowledge may ever be under standard circumstances, the resulting rules of anticipation are entirely impractical for the performance of TEs like Stevin's chain. Hence, in order to let the imagination in such TEs be guided by the right kinds of anticipatory rules, parts of our common background knowledge must be bracketed already at the outset. Thus construed, bracketing is an essential tool for channelling our anticipations in desired ways.

The previous remarks on the role of bracketing also allow me to finally explicate my thesis according to which TEs should be understood as *anticipation pumps*.¹³ On the view proposed here, TE-narratives are well designed if they accomplish two objectives: First, they must trigger the right kind of background knowledge in order to motivate the right kinds of anticipatory rules. And secondly, they must ensure that these rules of anticipation aren't interfered by anticipations that are inessential or even an impediment for reaching the desired outcome of the imaginative process. Hence, to put it in a slogan: A TE is well-designed if the TE-narrative is successful in "pumping" the right rules of anticipations. The "right" rules are those that lead us from the initial TE-setup to the desired end point with a minimum of unnecessary distractions.

5.3 The Conclusion Stage

To perform a TE means to successfully "live through" the anticipatory horizons against the background of which fulfilled intentions towards particular phases of the imaginative process always stand out. The performance stage of a TE is completed when

¹³ Of course, in using this terminology, I take a cue from Daniel Dennett who famously dubbed TEs "intuition pumps". However, my agreement with Dennett's views is rather superficial. While I agree that the hallmark of a well-functioning TE is that it is "cunningly designed to focus the reader's attention on 'the important' features, and to deflect the reader from bogging down in hard-to-follow details" (Dennett 1984, 12), I do not share the dismissive attitude that is characteristic for Dennett's early takes on TEs.

we reach that state of the TE-scenario which the presenter deems relevant for the projected target-thesis. Like in the case of physical experiments, we can now ask whether the outcome of the TE supports the conclusion that its presenter draws. For instance: Do the changing states of a water-bucket-system in an otherwise empty universe really force us to accept the existence of absolute space? Does a light ray that enters an elevator horizontally really force us to accept that the effects of gravity and inertial acceleration are indistinguishable?

But scrutinizing the relation between the outcome of the imaginative process and the projected target thesis is not all that happens at the conclusion stage. In many cases it is also natural to reflect on whether the course of the imaginative process itself was inevitable, independently from its purported impact on the target thesis. Here is an example that might illustrate the point: Consider a modern version of Galileo's tower-TE in which we imagine a cannon ball that is tied to a musket ball. We are then asked to throw the combined system from a tower and to see what happens. If the speed of fall of bodies is really proportional to their weights, as the Aristotelian theory suggests, then the combined system will fall faster *and* slower. This outcome is usually taken to show that the Aristotelian law of falling bodies is false and must be replaced with Galileo's.

Now, critical reflection on this TE can occur on two different levels:¹⁴ On the one hand, it is natural to wonder whether both conclusions – the rejection of the Aristotelian law and the proclamation of the Galilean law – are equally supported by the outcome of the imaginative process. Yet, on the other hand, it is also possible to ask whether the outcome itself is inevitable. Is it really the case that the anticipations in this TE unavoidably lead us to two conflicting scenarios in which the combined system falls *both* faster *and* slower? Or do our anticipations vary depending on how tightly the musket ball and the cannon ball are connected? Do our anticipations vary depending on whether the two objects are connected with a rope or with a rubber band? Can a cannon ball that is connected with a musket ball really be treated as *one* object, as the TE seems to presuppose (cf., for versions of these and similar other objections, Koyré 1968, 51; Gendler 1998, 404-406)? Such questions obviously do not concern the relation between the outcome of TE-process and the projected target thesis. Rather, such questions concern the TE-process itself.

How can questions concerning the inevitability of the TE-process be resolved? Given what has been said so far, an answer to this question goes along the following lines: Since the evolution of a TE-setup is determined by the rules of anticipation that govern the TE, scrutinizing the outcome of the TE-process can only proceed by way of explicating the background knowledge that initially motivated the relevant rules of anticipation. During the performance stage, i.e. during the process of “living through” the horizons in which the phases of the TE are given, this background knowledge is tacit. However, if we wish to critically reflect on the strength with which the outcome of the TE-process imposes itself on us, the relevant parts of our background knowledge must

¹⁴ The two levels I am referring to here resemble Brown's distinction between experiment in the broad sense and experiment in the narrow sense: “In the narrow sense, an experiment includes the set up and the observation [...]. In the broad sense, the experiment includes background assumptions and initial theorizing, the setup, observation, additional theorizing, calculating, and drawing the final conclusion.” (Brown 2007, 157-158) What I call the level of the TE-process is roughly similar to Brown's experiment in the narrow sense.

be made explicit.¹⁵ The strength with which the outcome of a particular TE-process imposes itself on us will depend, among other factors, on the type of background knowledge that motivates the relevant rules of anticipation: If, for instance, a particular set of anticipations is motivated by a theory that operates in the background of a TE, then the degree to which we accept the outcome of the TE depends on the degree of belief in the underlying theory. If, on the other hand, a particular set of anticipations turns out to be determined by an essential law, then the outcome of the imaginative process will be regarded as necessary.

Let me conclude this section by briefly commenting on one last issue: On the view proposed here, to perform a TE is to “live through” the rules of anticipation that are co-given at each stage of the imaginative process. These rules are motivated by the background knowledge that is triggered during the preparatory stage, i.e. through the concepts of which the TE-narrative is composed. Hence, the course as well as the outcome of a TE is ultimately determined by chunks of pre-established background knowledge. But this raises an obvious question: How can we learn something *new* from TEs if their outcome is determined by background knowledge that must already have been acquired before the TE can even be set up? Isn't it a consequence of the view defended here that the performance of a TE is the mere recalling of what was already known?

Since scientific TEs are a very diverse lot, it is, I think, impossible to give a principled answer to how TEs manage to go beyond the knowledge one already has to possess in order to perform the TE. It is thus mainly for the sake of brevity that I will reduce myself to one single aspect that seems to be of particular importance in this context: As I have pointed out, TEs require us to immerse ourselves into quasi-worlds that differ from the actual world in more or less drastic ways. One of the reasons why immersion is epistemically significant is that it gives us the opportunity to test our background knowledge under conditions that could not be replicated otherwise. Of course, there is also a danger in projecting quasi-worlds that are far removed from the actual world: The more outlandish the quasi-world, the higher the risk that our anticipations eventually become ungrounded. But one of the things a well-designed TE accomplishes by way of projecting idealized quasi-worlds is to provoke collisions between parts of our background knowledge that would not even come close under normal circumstances.¹⁶ It is in this way that TEs allow us to detect and get rid of inconsistencies in our existing background knowledge, to explicate and scrutinize background assumptions that were previously left unquestioned and, in some particularly impressive cases, even to

¹⁵ It should be noted, however, that the explication of tacit background knowledge usually comes at the price of diminishing the cognitive efficacy and elegance of a given TE. This point has been emphasized by David Gooding (1992, 286).

¹⁶ An example might help make this point more vivid. As noted earlier (cf. footnote 6), our background image of the world is in part composed of pieces of non-propositional, practical and sensorimotor knowledge about how things are done and how we use our bodies to do them. At first blush, one might think that, since TEs are performed in the laboratory of the mind, implicit sensorimotor and kinaesthetic knowledge is irrelevant for the practice of TE-reasoning. However, as Gooding (1992) and Yiftach Fehige and I (2013) have argued, this is far from being the case. While Gooding employs a broadly naturalistic framework to make this point, Fehige and I have approached the issue from a phenomenological perspective. The aim of our paper was to a) show that implicit body knowledge is operative in many instances of TE-reasoning and to b) illustrate this claim by means of an analysis of Newton's bucket-TE.

reconfigure the conceptual apparatus with which we approach the world (Kuhn 1977; Gendler 1998).

6. Concluding Remarks

The goal of this paper was to set out some basic directions in which a phenomenological framework for the analysis of scientific TEs should be developed. On my view, phenomenology differs from the existing approaches, among other things, in its starting point: The main objective, at least initially, is not to construct a theory that seeks to bring TEs in line with certain pre-established epistemological, ontological and metaphysical views. The main objective of a phenomenological account is rather to give a faithful description of the *actual performance* of TEs from a first-person perspective and to go on from there. One of the aims in the preceding sections was to indicate how such a description might actually look like. But, of course, much work remains to be done in order to position phenomenology as a serious contender in the ongoing debate on TEs.¹⁷

7. Literature

- Arabatzis, Theodore (2012): "Hidden Entities and Experimental Practice: Renewing the Dialogue Between History and Philosophy of Science", in: Mauskopf, Seymour & Schmalz, Tad (eds.): *Integrating History and Philosophy of Science. Problems and Prospects*. Dordrecht: Springer, 125-140.
- Brendel, Elke (2004): "Intuition Pumps and the Proper Use of Thought Experiments", in: *Dialectica* 58/1, 89-108.
- Brown, James Robert (2004): "Why Thought Experiments Transcend Empiricism", in: Hitchcock, Christopher (ed.): *Contemporary Debates in Philosophy of Science*, Malden: Blackwell, 24-43.
- Brown, James Robert (2007): "Counter Thought Experiments", in: *Royal Institute of Philosophy Supplement* 61/82, 155-177.
- Brown, James Robert (2011): *The Laboratory of the Mind. Thought Experiments in the Natural Sciences*, London & New York: Routledge.
- Brown, James Robert (2013): "What Do We See in a Thought Experiment?", in: Frappier, Melanie, Meynell, Letitia & Brown, James Robert (eds.): *Thought Experiments in Science, Philosophy, and the Arts*, London & New York: Routledge, 53-68.
- Chudnoff, Elijah (2013): *Intuition*, Oxford: Oxford University Press.
- Crowell, Steven Galt (2006): "Husserlian Phenomenology", in: Dreyfus, Hubert & Wrathall, Mark (eds.): *A Companion to Phenomenology and Existentialism*, Malden: Blackwell, 9-30.
- Davies, David (2007): "Thought Experiments and Fictional Narratives", in: *Croatian Journal of Philosophy* 19, 29-45.
- Dennett, Daniel C. (1984): *The Elbow Room. The Varieties of Free Will Worth Wanting*, Oxford: Clarendon Press.
- Fehige, Yiftach & Wiltsche, Harald (2013): "The Body, Thought Experiments, and Phenomenology", in: Frappier, Melanie, Meynell, Letitia & Brown, James Robert (eds.): *Thought Experiments in Science, Philosophy, and the Arts*, London & New York: Routledge, 69-89.
- Faraoni, Valerio (2013): *Special Relativity*, Springer: Dordrecht.
- Foster, John (2000): *The Nature of Perception*, Oxford: Oxford University Press.

¹⁷ In developing my views over the last couple of years, I profited from the conversations I had with Jim Brown, Keizo Matsubara, Danny Goldstick, Sonja Rinofner-Kreidl, Walter Hopp, Mike Stuart, Yiftach Fehige, Geordie McComb, Hanne Jacobs, Sebastian Luft, Jeff Hilderley, George Heffernan, Michael Wallner and James McGuirk. I would also like to say "thanks" to the Austrian Science Fund for the generous support of my ongoing research on TEs (project number: J 3114).

- Froese, Tom & Gallagher, Shaun (2010): "Phenomenology and Artificial Life: Toward a Technological Supplementation of Phenomenological Methodology", in: *Husserl Studies* 26, 83-106.
- Gendler, Tamar Szabó (1998): "Galileo and the Indispensability of Scientific Thought Experiment", in: *British Journal for the Philosophy of Science* 49, 397-424.
- Gooding, David C. (1992): "What is Experimental about Thought Experiments?", in: *Proceedings of the Biennial Meeting of the Philosophy of Science Association. Volume 2*, Chicago: University of Chicago Press, 280-290.
- Gunstone, Richard & White, Richard (1981): "Understanding Gravity", in: *Science Education* 65/3, 291-299.
- Hopp, Walter (2011): *Perception and Knowledge. A Phenomenological Account*, Cambridge: Cambridge University Press.
- Hopp, Walter (2014): "Experiments in Thought", in: *Perspectives on Science* 22, 76-97.
- Horowitz, Tamara & Massey, Gerald J. (eds.): *Thought Experiments in Science and Philosophy*, Lanham: Rowman & Littlefield.
- Husserl, Edmund (1960): *Cartesian Meditations. An Introduction to Phenomenology*, translated by Dorion Cairns, The Hague: Martinus Nijhoff.
- Husserl, Edmund (1973): *Experience and Judgement. Investigations in a Genealogy of Logic*, translated by James S. Churchill and Karl Ameriks, Evanston: Northwestern University Press.
- Husserl, Edmund (1983), *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy. First Book*, translated by Fred Kersten, The Hague, Martinus Nijhoff.
- Husserl, Edmund (1997): *Thing and Space: Lectures of 1907*, translated by Richard Rojcewicz, Dordrecht et al.: Kluwer.
- Husserl, Edmund (2001a): *Logical Investigations. Volume 1*, translated by J. N. Findlay, London & New York: Routledge.
- Husserl, Edmund (2001b): *Logical Investigations. Volume 2*, translated by J. N. Findlay, London & New York: Routledge.
- Husserl, Edmund (2001c): *Analyses Concerning Active and Passive Synthesis. Lectures on Transcendental Logic*, translated by Anthony Steinbock, Dordrecht et al.: Kluwer.
- Husserl, Edmund (2005): *Phantasy, Image Consciousness and Memory (1898-1925)*, translated by John B. Brough, Dordrecht: Springer.
- Kind, Amy (2013): "The Heterogeneity of Imagination", in: *Erkenntnis* 78/1, 141-159.
- Koyré, Alexandre (1968): "Galileo's Treatise 'De Motu Gravium': The Use and Abuse of Imaginary Experiment", in: *Metaphysics and Measurement: Essays in the Scientific Revolution*, Cambridge: Harvard University Press, 44-88.
- Kuhn, Thomas S. (1977): "A Function for Thought Experiments", in: *The Essential Tension. Selected Studies in Scientific Tradition and Change*, Chicago: University of Chicago Press, 240-265.
- Kujundzic, Nebosja (1995): "Thought Experiments: Architecture and Economy of Thought", in: *Journal of the British Society of Phenomenology* 26/1, 86-93.
- Lampa, Anton (1924): "Wie erscheint nach der Relativitätstheorie ein bewegter Stab einem ruhenden Beobachter?", in: *Zeitschrift für Physik* 27/1, 138-148.
- McAllister, James W. (1996): "The Evidential Significance of Thought Experiment in Science", in: *Studies in History and Philosophy of Science* 27/2, 233-250.
- Miščević, Nenad (1992): "Mental models and thought experiments", in: *International Studies in the Philosophy of Science* 6/3, 215-226.
- Mohanty, Jitendra Nath (1991): "The Method of Imaginative Variation in Phenomenology", in: Horowitz, Tamara & Massey, Gerald, J. (Eds.): *Thought Experiments in Science and Philosophy*, Savage: Rowman and Littlefield, 261-272.
- Nersessian, Nancy (1993): "In the Theoretician's Laboratory", in: *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association, Vol. 1992, Volume Two*, 291-301.
- Noë, Alva (2004): *Action in Perception*, Cambridge et al.: MIT Press.
- Norton, John (1996): "Are Thought Experiments Just What You Thought?", in: *Canadian Journal of Philosophy* 26/3, 333-366.
- Norton, John (2004): "Why Thought Experiments do not Transcend Empiricism", in: Hitchcock, Christopher (ed.): *Contemporary Debates in Philosophy of Science*, Malden: Blackwell, 44-66.

- Palmieri, Paolo (2005): “‘Spuntar lo scoglio più duro’: did Galileo ever think the most beautiful thought experiment in the history of science?”, in: *Studies in History and Philosophy of Science* 36, 223-240.
- Penrose, Roger (1959): “The Apparent Shape of a Relativistically Moving Sphere”, in: *Mathematical Proceedings of the Cambridge Philosophical Society* 55/1, 137-139.
- Sartre, Jean-Paul (2004): *The Imaginary. A phenomenological psychology of the imagination*, London & New York: Routledge.
- Siewert, Charles (2007): “In favor of (plain) phenomenology”, in: *Phenomenology and the Cognitive Sciences* 6, 201-220.
- Smith, David Woodruff & McIntyre, Ronald (1982): *Husserl and Intentionality. A Study of Mind, Meaning, and Language*, Dordrecht, Boston & London: Reidel.
- Smith, David Woodruff (2004): *Mind World. Essays in Phenomenology and Ontology*, Cambridge: Cambridge University Press.
- Smith, David Woodruff (2007): *Husserl*, New York & London: Routledge.
- Sturgeon, Scott (2000): *Matters of Mind: Consciousness, reason and nature*, London: Routledge.
- Terrell, James (1959): “Invisibility of the Lorentz Contraction”, in: *Physical Review* 116/4, 1041-1045.
- White, Alan R. (1990): *The Language of Imagination*, Oxford: Blackwell.
- Wiltsche, Harald A. (2013): “How Essential are Essential Laws? A Thought Experiment on Physical Things and Their Givenness in Adumbrations”, in: Günzler, Ingo & Mertens, Karl (eds.): *Wahrnehmen, Fühlen, Handeln. Phänomenologie im Wettstreit der Methoden*, Münster: Mentis, 421-436.
- Wittgenstein, Ludwig (1980): *Remarks on the Philosophy of Psychology*, Oxford: Blackwell.