Generative and Demonstrative Experiments

Tommaso Bertolotti

Abstract Current scientific practice is often identified with the experimental framework. Yet, what "experimenting" means could be less than perfectly clear. Going beyond the common sense conception of experiment, two broad categories of experiments can be tentatively identified: the generative experiment and the demonstrative experiment. While the former aims at generating new knowledge, new corroborations of hypotheses etc., the latter—which is actually the kind of experiment most laypeople came to terms with in their lives—is designed so that, by being successful, it reverberates knowledge on the experimenters/witnesses, thus instructing them, albeit the experimental outcome was well known beforehand. *Prima facie* the uninformed observer may not always be able to tell whether an experiment is generative or demonstrative, therefore the existing distinction must rely on something else, namely the framework they are embedded into. The concept of epistemic warfare, recently introduced by Magnani, can be of help in investigating this distinction, also to the scope of showing that it is not a sterile dichotomy but rather a theoretically fruitful continuum, and can help the analysis of epistemically relevant issues such as the repetition/replication of experiments and their potential failure.

1 Introducing the Experiment

The idea of experiment is intuitively connected with the common conception of modern science. Yet, until the second half of the twentieth century, philosophy of science reenacted the ancient bias against craftsmanship and focused chiefly on the theoretical aspects of scientific endeavor. Breaking this tendency, philosophical

T. Bertolotti (✉)
Department of Humanities—Philosophy Section, and Computational Philosophy Laboratory, University of Pavia, Pavia, Italy
e-mail: bertolotti@unipv.it

L. Magnani (ed.), *Model-Based Reasoning in Science and Technology*, Studies in Applied Philosophy, Epistemology and Rational Ethics 8,
DOI: 10.1007/978-3-642-37428-9_27,

milestones such as Hacking [10] and Gooding [9] claimed the experimental question rightly back in the epistemological feud, but the topic was quickly seized by a different branch of studies, sometimes called social epistemology, sociology or anthropology of science, which focused more on the social dimension of experimental settings (consider [3, 14]).

The aim of this paper is to make the best of these two approaches (the epistemological care for experimentation, and the social-anthropological outlook), in order to set the framework for an updated and consistent consideration of experiment: that is, what I mean to provide is an analysis of experimentation able to comprehend both *crucial* experiments carried out in laboratories, and the more modest kind of experimentation we came across, for instance, during our high school years.

I believe that experimentation is a particularly pivotal topic for the understanding of science as a whole. Science is a many-headed deity: Hacking [10] claims in the Introduction that, albeit one decides to deal either with scientific rationality or scientific realism, one topic necessarily ends up defining the other. There are many roads leading to the same castle.

1.1 Defining the Experiment

I shall begin by sketching out an extensive definition of what can be considered an experiment. In order to be fruitful, the definition of what an experiment is must be neither too broad nor too narrow: I would rather enumerate a list of features that, in my opinion, make up the experiment rather than providing a full definition.

- *Any experiment is characterized by a manipulative dimension.*

Hacking [10] stressed the importance of *intervention*, and rooted his scientific realism not on our possibility to *know* but on our possibility to *intervene*. This intervention has to be understood in its fully dynamic display: an experiment is not the result of the experiment itself, but the whole process by which this result is achieved—or not. This can be said of thought experiments as well: as contended by Gendler [8], a thought experiment like the ones conducted by Galileo cannot be reduced to a more or less sound logical argument, because the manipulation and hence the manipulability by an agent are the pivotal feature.[1] This is all the truer if we think of real experiments: you can *tell* somebody how an experiment was conducted (the preparation, the procedure, and the outcome), but the narrative and communicative reconstruction of the experiment, and the description of the involved procedure, are not the experiment itself. Experiments are a manifest example of *manipulative abduction* [18].

[1] This view on thought experiments is not universal. Some scholars contend that they can in fact be reduced to straightforward arguments [28].

- *All experiments have a transformative nature.*

Experiments can be thought of as mechanical systems: to begin with, we have what is being experimented on; we then add what the experimenters bring to the experiment, that is models, heuristics, techniques, personal hunches and so on. Finally, manipulation, intervention, the *work* can take place. This work has an output, of any kind. It seems to me that without this *output* there would be no experiment to begin with: the result can be new knowledge, so to say "extracted" with the experimental manipulation. Of course one cannot know that she extracted all the available knowledge from the experiment: just like when you squeeze an orange to have the juice, someone might show you there was more juice left, or that there can be a better way to squeeze it. Yet, as I will explain along the next sections, an experiment does not only generate (absolutely) new knowledge: the output of the experimental transformation might also be a new affordance (of the experimented, for instance), or new commitments (e.g., toward the advance of science). A High Energy Particle Collider experiment will try to produce new knowledge about subatomic particles, whereas a high school physics experiment might help students acquire a new and better understanding of a certain phenomenon, or a renewed commitment towards scientific progress: that, too, is an effect brought about by the kind of transformations enacted by the experiment.

- *All experiments are "situated".*

Any experiment begs for a situation. It could be argued that experimentation itself projects its situation: experiments are often about the controlled manipulation of a number of variables in order to see "what happens when...", or "if it is the case that *x*...". The laboratory is the situation *par excellence*, and the next subsection will focus on it, but it is not the only one: even in a non-scientific setting, but when people wish to make use of a kind of rationality that can be called scientific, the suggestion "Let's make an experiment" entails the setting of a boundary: it is pragmatic (i.e. deciding what matters and what does not matter), but also regards the assigned social roles. Depending on the interactions, the peers of the person who calls for an experiment will be involved as active participants (as in, "Let's make an experiment: what would you do if...?"), or as onlookers/witnesses, expected to provide some opinion about the conclusion and the procedure. In any case, the experiment takes place within a well defined place, which can be more or less physically determined. Hacking refers to "mature" laboratory sciences those "in which we investigate nature by the use of apparatus in controlled environments, and are able to create phenomena that never or at best seldom occur in a pure state before people have physically excluded all 'irrelevant' factors" [11, p. 507]. Let us therefore take a closer look at what we called the boundaries of the experiment, which must be the boundaries of the laboratory.

1.2 Setting the Boundaries of the Laboratory

Let us accept our loose intuition about what a laboratory is: informally, we can think that a laboratory is the specific location where scientific experiments take place. But more can be said about the lab: first of all, what are the actual boundaries of the laboratory, understood as the *lieu* of the experiment? With this respect, I invite the reader to make a small recollection and consider what the word "laboratory" makes her think of: she could think of the instrumentation used for experiments, she could think about the instrumentation strategically laid out on workbenches, and perhaps of scientists carrying out experiments on these workbenches. Plus, thinking of the labs she might have attended, she could also think of all this *and* onlookers standing by and witnessing the experiment.

It is clear that the epistemological consequences of where we set the boundaries of the lab are quite significative. In the last item, the onlookers can be students, colleagues, sponsors, visitors at a science exhibition, and so on: I do not believe that including them among the possible target of the knowledge *transformation* enacted by the experiment necessarily means to shift the investigation from the epistemological plane to a social, anthropological one. Science as an *actual* human endeavor cannot be investigated excluding the human dimension it relies on. Knorr Cetina wrote that "the power of laboratories (but also their limitations) resides precisely in this 'enculturation' of natural objects. Laboratory sciences subject natural conditions to a 'social overhaul' and derive epistemic effects from the new situation" [3, p. 28].

I believe that the soundness of Knorr Cetina's statement does not exclusively follow from the adoption of an ethnographic outlook on science.[2] Conversely, it is easy to understand that the *enculturation* and the *social overhaul* are mutually implicating in our conception of science, because laboratories are more than a set of instruments, and even more than a set of affordances displayed by those instruments: labs prevent scientists from having to study a natural object "*as it is, [...] where it is [and] when it happens*" (p. 27). In this sense, labs allow the manipulation of the object far better than the natural context would (if it would at all). Even scientific models, playing a pivotal role in the economy of the lab, partake of a similar nature, being on the one hand highly manipulative human construals (therefore cultural too), while on the other they are necessarily bound to the natural object: in fact models are fundamental in binding the modeled natural object into a specific phenomenon [1, 19].[3]

The laboratory includes the experimenters as well, inasmuch as they are not separated from what is being experimented. Albeit Hacking is thankful toward the "large number of studies by philosophers, historians, and ethnographers of

---

[2] Also Nersessian's outlook on science is often characterized by a particular attention—called "ethnographic"—to the actual dynamics at play in a laboratory (cf. for instance [25–27]).

[3] Hacking [10] contends as well that many phenomena come to happen uniquely as they are *created* in laboratories.

experimental science," [11, p. 508] he seems less eager to concede a more significative role to human intervention per se, which is conversely mentioned by Heelan by means of the "instruments, standard procedures, experimental skills, laboratory traditions, and the social context of the research community" [7, p. 525]. Scientists are not simply interchangeable operators: two teams working on the same raw objects would not apply the same methodologies or necessarily obtain the same results. Indeed, "not only objects but also scientists are malleable with respect to a spectrum of behavioral possibilities. In the laboratory, scientists are methods of inquiry; they are a part of a field's research strategy and a technical device in the production of knowledge" [3, p. 29]. It seems to me that what could be at stake here is not the dispossession of epistemology by social approaches, but rather the opposite, that is the epistemological flooding of some aspects of scientific endeavor which—by default of better option—have so far been labeled as social but do rather concern an agent-based and factual approach to science, also leaning on an *actually happens rule*.[4]

Why then should we set the limits of the laboratory at the experimenters' level, and not admit the onlookers as well? Why should the "social overhaul" advocated by Knorr Cetina only involve the experimenters? As I will show in the following section, some experiments do not benefit the experimenters at all, in a strict sense, inasmuch as they produce knowledge that had already been acquired, and yet the same experiments cause indisputable epistemic effects on those who *observe* them. For them the experiment still produced a valuable transformation of knowledge, by which they gained a new *understanding* 1) of the phenomenon that was *explained* to them through the experiment, but also 2) of how scientific rationality works. It could be contended, in fact, that some experiments carried out at schools or science exhibits serve the chief purpose of exemplifying some tenets of scientific method.[5]

Concluding this section, we might say that there are indeed many kinds of laboratory, in which different kinds of experiments take place: there are labs for basic research, industrial labs, labs for medical research, and then there are laboratories in schools, science museums, and also the laboratories shown in

[4] This rule was introduced by Gabbay and Woods as a tenet of their new approach to logic, referring to the fact that logic should model how real agents think: one should try to correct the model so it fits the facts, and not try to amend or obliterate facts to make them fit the model [6, 35]. In this context, I use it to suggest that philosophy of science should match what science really is, and not arbitrarily cut out aspects of the problem by labeling them as external to the analysis (for instance, "social").

[5] The onlooker's gain of a renewed commitment towards science, be it specific for a particular research/discipline or to scientific endeavor in general, is just as vital for the development of science as the generation of new knowledge through experimentation. Contemporary *knowledge societies* massively rely on the development of science, which in turn relies on the will of citizens to care and spend for it [17]: funds are just as vital as genius and intelligence for the survival of science. This view is coherent with Magnani's conception of science as an *epistemic warfare* [19], which also includes non-epistemic strategies that are nevertheless crucial for science, such as those for the dissemination of knowledge, the acquisition of funding and so on.

educational TV programs: most of laypeople are acquainted with the latter kind of laboratory, that is the physics/chemistry/biology lab at high school, or those they see in science museums or on the Discovery Channel. Such acquaintance fuels our *thinking to know* what every lab should be like, which is in fact a *hasty generalization*.

In his classic book *Science in Action*, Latour enacts his anthropological approach to science narrating the epistemological adventure of an anthropologist taking a full immersion in the scientific endeavor. Interestingly, he makes the narrator say: "We came to the laboratory in order to settle our doubts about the paper, but we have been led into a labyrinth" [14, p. 67]. Specifically, the doubts referred to a reading of endorphin levels, which had to be interpreted through graphs and indicators, yet this bewilderment is common to many onlookers approaching a scientific setting: *we came, we saw, and yet we have not understood anything*. And yet, we saw experiments, at school, at the science museum, on TV, how comes?

In sum, experiments take place in laboratories, and laboratories may include onlookers. Yet not all experiments are geared towards onlookers the same way: to certain experiments anyone can be an onlooker and benefit of the epistemic effects, to others the onlooker is defined by very specific characteristics. In my opinion, this depends mostly on the kind of experiment at stake. If different kinds of experiment exist, it is legitimate to wonder how many kinds there are, and how we can tell them apart.

2 How Many Kinds of Experiment Are There?

A kind of *taxonomy* of experiments is not unusual among philosophers of science, and such differentiations sometimes merge into other connected ones. Gooding, for instance, links the concept of experiment to its *reconstruction*, obtaining six different kinds of reconstruction to be employed in different narratives: namely *cognitive, demonstrative, methodological, rhetorical, didactic, and philosophical* [9, p. 7], each with their peculiar scope. Notwithstanding the utility and soundness of this differentiation, I contend that its root lies at a lower level, and actually underdetermines it. The whole spectrum of experimental activity, as far as natural and model-based sciences are concerned, could in fact be reduced to two major forms of experimentations. One of the advantages of this proposal, which I dub a "plea for epistemological austerity", is that every distinction causes some unhappy left-outs: Steinle [33], for instance, lamented that the "standard view" in the Nineties of the past century would disregard as epistemologically irrelevant those experiments that were aimed at discovery—and not at the test of a clear hypothesis, or at the retrieval of a particular measurement. Such conception would in fact leave out a number of fundamental instances in the history of science: grouping the experiments into two sets, namely "generative" and "demonstrative" experiments could instead cause a lesser number of *homeless* instances.

2.1 The Generative Experiment

I could begin by suggesting that what I call the "generative" experiment is the kind of experiment that common sense has acquired, but this would be misleading. I contend, indeed, that the common-sense conception of experiment is somewhat blurred, so that the generative experiment, which is what we *should* think of when we think of a scientific experiment, does not coincide totally with our intuitive conception of experiment.

The *generative* experiment is the experiment whose outcome is *not known* beforehand, and its aim is to manipulate and transform the *experimentandum* (what is being experimented on) into knowledge that is new for everyone. To put it another way, it is the kind of experiment where the epistemic target,[6] that is what the experimenters want to obtain, is intrinsic to the experiment (this latter claim might seem a truism, but the next subsection should prove the opposite).

Most experiments in the history of science can be thought of as *generative* experiments. It is the kind of experiment where you *test* something (a hypothesis, a theory), and is usually comprised within a theoretical framework. It is also true that "one can conduct an experiment simply out of curiosity to see what will happen" [10, p. 154]: not only experiments that are well nested within a particular theoretical framework, for instance those aimed at testing a particular hypothesis, or at finding out a particular measurement (think of Millikan's experiment, projected to measure the elementary electric charge), but also entirely "exploratory" experiments are generative. According to Steinle, explanatory experiments do not rely on a "specific and well-defined procedure, but [include] a whole bundle of different experimental strategies", and their "central epistemic goal is the search for general empirical rules and for appropriate representations by means of which they can be formulated" [33, p. S73].[7]

But also in our everyday life, when we make use of scientific-experimental rationality to put some makeshift model to the test, we recur to generative experiments to gain some new knowledge. I can send myself an email to see if my IMAP server is really experiencing issues, and I can ask a friend to email me as well. I can put a five-dollar bill in a vending machine to test it before butting a twenty-dollar bill, to see if the machinery works properly. Generative experiments are often conducted as part of model-based activities: I can ask a relative to simulate a social situation to gain better knowledge about some possible

[6] I specify epistemic target, as the scope of the experiment, to differentiate it from Hacking's use of the word *target*, by which he refers to a part of the *"materiel"* of the experiment (cf. [11, p. 509]).

[7] Steinle's aim in describing exploratory experimentation is to allow the appreciation of the epistemological importance of this kind of experiment, while the "standard view" tended to disregard them as part of epistemically irrelevant *discovery* processes. Exploratory experiments are particularly relevant for entering new fields requiring new concepts and new general facts [33]. The explanatory experimentation can also be extremely tacit, and consist chiefly of "thinking through doing" [16].

consequences of an action of mine, or a man might cast small objects off a table to assess the likelihood of himself surviving after jumping from a cliff with his car. In those cases, what I gain from the manipulative intervention of the target (that is from the experimentation) is some knowledge I did not possess before.

In sum, the focus in generative experiments should be put on their ability to intrinsically produce new knowledge. This is the kind of experimentation that engages theory (and theories): as suggested by Steinle [33] and Hacking [11] among many others, some experiments—which I label as generative—can precede theory inasmuch as they can illuminate new fields of scientific research and provide it with new concepts.[8]

With respect to this kind of experiment, even scientific common sense knows that theories should behave according to the already mentioned "actually happens rule": experimental observations affect theories. Experiments are where theories can be falsified [29], and experiments that do not go as expected can affect the scientific paradigm, taking it to an eventual crisis [12], or causing scientists to fix the protective belt of the program to keep it progressive [13]. In the next two subsections I will show how only one kind of experiments indeed affects theories, and then move to analyze a wide and yet peculiar class of experiments, that—even though they can be called experiments to their full right—are not expected to affect theories at all.

2.2 The Demonstrative Experiment

It is now time to deal with an apparent contradiction: we know that experiments are, so to say, the field artillery of scientific progress, and it is on experimental grounds that new knowledge is either discovered or validated. On the other hand, we also know very well that most experiments we—as laypeople—witnessed (even in decent laboratories) did not add anything to scientific knowledge. It would not be right to arbitrarily exclude them from the category of experiments, because they display all the traits I pointed out in Sect. 1.1, and they also fit with the more demanding description proposed by Hacking [11].

I am referring to most experiments carried out in schools, exhibitions, museums, and so on. For instance, they can be experiments aimed at demonstrating or

[8] Hacking suggests several examples from the actual history of natural science that refute Popper's claims according to which "theory dominates the experimental work from its initial planning up to the finishing touches in the laboratory" [10, p. 155]. The debate on the theory-ladenness of experimental facts is often brought to quasi-metaphysical issues: one way to tackle it is to appeal to the intuitive notion of theory (as folk theory). Experiments may precede particular theories, and yet rely on past sub-theories about substances, agency, causation etc. Thus, to say that an experiment precedes theory—and so does the experimental observation that follows such experiment—does not indeed equal saying that the experiment generates new coherent knowledge *ex nihilo*. After all, we could claim that intuitive, hard-wired theory precedes even out every-day observation, even at the lowest levels of the perception of images, sounds etc. [30].

illustrating a law or a theory, fostering a better understanding of it. With this respect, at least in the Italian school system, theory overwhelmingly precedes experimentation: in chemistry or physics courses, experiments are not even used to stimulate theorization upon the students' minds, but rather as a persuasive proof to show that what was explained in theory *actually happens*.[9]

This kind of experiment could be thought of as *deduced* from theory in a strong sense, opposed to the *weak* Popperian sense of experiments *informed* by theory: The procedure of the experiment is vouched for by the theory it means to put in display. Is it a *paetitio principii*? Not really. Consider this example:

1. *Experiment E* (for instance Maxwell's or Faraday's experiments on electromagnetism) is crucial for the establishment of a *Theory T*;
2. *Theory T* is established;
3. *Experiment E** is used at school to prove the adequacy[10] of *Theory T*.

*Experiment E** is a (usually easier) version of *E*, updated according to the theory it means to demonstrate. If its real aim was to *test* the theory, then of course it would be begging the question. But who would expect high-school students to be *actually testing* a theory? Everybody knows that high-school level optics, or electromagnetic physics and so on *do work*. Proving it *n* more times every day, in *n* school laboratories, does not add one bit to the robustness of those theories. *Experiment E** aims at providing students with an actual proof that what they studied (or they are going to study) is really so.

Even if you think about experiments that do not aim at demonstrating a law, but rather at isolating a phenomenon so that it can be shown for some theoretical scope, the defining element is that the experimental *outcome is known beforehand*.[11] Contrarily to the *generative* experiment, in this case the epistemic goal is extrinsic to the experiment itself: it means little to say that the experiment *in se* was successful, because it was planned to be successful. The experiment is successful in its *actual* scope if it operates any *change* within the epistemic configuration of the observer, after she witnessed the positive (staged) outcome of the experiment. That is to say, the experiment is successful if it triggered a new awareness in the observer, for instance a student might be further persuaded about the empirical adequacy of a theory, or a citizen might reconsider the importance of electing a prime minister advocating more funding for scientific research. Or, simply, their aim could be to convey indeed a bit of *local* knowledge about some

[9] This concept is well exemplified by a sign hung in my chemistry laboratory at high school, which would read something along the lines of "If I listen I will forget; if I see I will remember; if I do I will understand". The experimental dimension is taught as completely subsidiary to abstract theory.

[10] Please understand this word in an intuitive sense, as in "What they taught me about the *Theory T* does indeed happen in real life", and not as laden with implications about the epistemological debate about the truthfulness or acceptability of a scientific theory.

[11] This claim clearly begs for some considerations about the *failure* of an experiment: I will address this issue in Sect. 3.2.

phenomena, but on the overall, to infuse the belief that science is "interesting", or just "cool".

This class of experiment could be defined as *demonstrative* or *explanatory*, contrasting it with *generative* experiment. Interestingly, one could say that in their scope of disseminating scientific knowledge (for various purposes), demonstrative experiments have become more and more widespread together with the growing impact of science on society. Living science shows in the eighteenth and nineteenth century, analyzed by Raichvarg [31], provide a clear example of a *demonstrative* experimental framework, which could be seen as the ancestor of modern science exhibitions or scientific shows for general audience on TV. One of the scientists/showmen mentioned by Raichvarg would start his experiments with the following call:

> And if I am here among you, it is because all of you must draw from my demonstrations, the true and natural principles of the forces which are above us, these forces which frighten the ignorant but supply the educated with all the moral pleasures of intelligence (p. 3).

Raichvarg draws from his analysis a list of characteristics that were typically common to science shows, and still apply to scientific dissemination aimed at general public:

> – They reach a wide audience, an audience which could be defined as a public with no scientific training...They come to the fairground for anything but science, but then they meet science face to face...
>
> – The importance of the current events of science, mostly because on a fairground one must astonish everybody to attract everyone and get your pennies back...
>
> – A continuous desire for good pedagogy, together with a continuous desire for wonder, if not for the supernatural! (p. 4).

These *experimental shows*[12] did not contribute to form scientists, just as contemporary science classes at high school do not mean to train scientists, and science museums do not either [15]. On the other hand, these forms of dissemination do play a pivotal role in educating people that might undertake or value a scientific career. A living science show, just as a school experiment, may indeed induce in the observer a taste for scientific methodology, or just make her aware of its existence.

A final question concerning the demonstrative experiment might arise: since we are accustomed with experiments carried out at school, often with obsolete equipment, it seems that the difference between a generative and a demonstrative experiment should be most easily noticed. In my opinion, from a phenomeno-

[12] The expression is a bit of an oxymoron, but it means to stress the *staged* dimension of many demonstrative experiments. Concerns about the esthetic dimension of their replication will be addressed in Sect. 3.1.

logical point of view, it is not so. We should not be fooled by the *time lag*: if most of our school labs look like museums of past century science it is just because those instruments were once upon a time the cutting edge of generative experimentation (think of Volta's battery and most electromagnetism-relate devices). If we removed this time lag, which is merely contingent, we would be phenomenologically unable to tell one kind of experiment from another, if not by considering the cultural and social framework an experiment is nested in.

Let us make a quick thought experiment: imagine in the near future a highly-funded high school in some advanced country, whose politicians place a great emphasis on education. Just as our high schools have a physics lab, that high school as a High Energy Particle Collider in its basement, and teachers use it to instruct pupils about quantum physics. If the same-old-friendly alien landed on Earth and could not understand human language, and witnessed the experiments carried out in that school, and those carried out at CERN (for instance), it could not be able to tell any difference: what goes on, apparently, is the same. Yet, once our alien managed to set its intergalactic translator to understand our language, it would see at once the difference, since the HEPC at school would be embedded in a pedagogical framework of demonstrative experiments, whose outcome are already known by the teacher who can therefore lead the pupils along the right path. Time lag, and thus the obsolescence of experimental materials, are not a necessary criterion to tell a generative experiment from a demonstrative one, since *prima facie* they cannot be told one from another, unless considering—as I said—the setting they partake of.

3 Consequences of the Distinction

Now that the distinction between generative and demonstrative experiment is in place, I will use it to tackle two epistemological problems, namely the repetition/replication of experiments and their failure. Once again, I will try to match common-sense expectations with the actual scientific practice, past and current.

3.1 Differentiating Repetition and Replication of Experiments

In an interesting paper about the conception of experiment repetition in the past centuries, Shickore [32] sets out stating that "[t]oday it is generally assumed that isolated experimental outcomes—'one-offs'—are insignificant. Twentieth-century philosophers of science, most notably Popper, made the reproducibility of experimental results the basic methodological requirement for successful experimentation: if an experiment cannot be re-done, it is invalid" (p. 327–328). Indeed, the possibility of re-doing an experiment became one of the first tenets of contemporary scientific rationality.

Before applying my distinction (between kinds of experiments) to the problem of redoing experiments, a brief semantic interlude is required, which—I think—might let the reader foresee my claim before I make it clear. It is sometimes said that experiments are "repeated", while sometimes they are "replicated". I believe that the two terms can be sensibly separated, each with its own proper meaning.

- **Repeating** an experiment exemplifies the epistemological tenet towards the re-doing of experiments. You *repeat* an experiment when you put the known outcome between brackets and proceed entirely as if it was unknown. The focus of repetition is on *what* outcome will be obtained, and *whether*—changing certain factors—the same outcome will be obtained again.[13]
- **Replicating** an experiment focuses on the replication of the procedure and not only on obtaining the same outcome. You *replicate* an experiment without necessarily putting the outcome between brackets, because what matters is observing *why* a particular procedure yielded such an outcome. Once the reason is found out, it is possible to replicate the experiment with the pragmatic certainty that if the outcome differs from what expected, then a mistake was committed in the procedure.

I suppose that such semantic characterization foreshadowed quite clearly the rest of my argument. As for generative experiments, I think that *repetition* is the case. Repetition engages the intrinsic epistemic goal of the experiment. Repeating the experiment does not mean necessarily to redo the same experiment over and over. This is what happens every day in schools worldwide, and we know that it has little epistemic value for the progress of science.

> A number of scholars have stressed that scientists rarely try to copy the exact same experiment. Rather, experimenters seek to obtain similar results in different experimental settings, and experimental results are considered valid if multiple determinations of the evidence are possible. The crucial notion here is reproduction by *doing something different* [32, p. 328].

The repetition of an experiment in a generative epistemic context is valuable because it may challenge the previous outcomes of the same experiment, for instance it can interfere with claims of universality (by "doing something different"). Repetition has therefore chiefly epistemic concerns.

Repetition can indeed be about the *same* experiment, but in this case it is about looking for *freak factors* of the experimental procedure, and make sure that the result is accurate. Even if scientific truths are notably *provisionally* true, the search

[13] To make students assimilate this concept, physics teachers often deploy plethoric lists of settings (e.g. here, at the Equator, on mount Everest, on the Moon, on Mars, in a billion years, and so on) where a law (such as "All metal bars expand when heated") must apply for it to be universal. The different settings correspond to a series of real or potential repetitions of one or more experiment concerning the law in question.

for freak factors will end at some point. No branch of science still heats metal bars every day on normal conditions to see if they expand and by what coefficient.[14]

Whereas such use of repetition was already in vogue in Early Modern science, its role was chiefly to corroborate (and make appear *as reliable*) one's own experimental results by the method of the *slight modifications*: Schickore, building his case study on an Italian eighteenth-century microscopist and physiologist, states that "Fontana's methodological thought is particularly interesting because he stressed the importance of *repetition* of his own experiments. The text is packed with claims that experimental trials were repeated 'a hundred times' or even 'a thousand times,' and that thousands of animals were used. Also, the experiments were varied 'in a thousand ways"' [32, p. 328]. Only subsequently the stress was placed on the assessment (via repetition) of experimental results obtained by other scientists.[15]

If the repetition is meant to engage the outcome of some other scientist's experiment, then again it can partake of a generative nature. Assessing someone else's experiments is, as a matter of fact, one of the pillars of contemporary scientific practice: a purpose of publishing experimental procedures in peer-reviewed journals is to offer the experiment to the assessment of peers, so that other scientists can repeat it and see *if* they obtain the same results. With this respect, an experiment is scientific if it is available for repetition, so that somebody else can repeat it and—perhaps—falsify its previous outcome: it is not necessary, for an experiment to be deemed *scientific*, to obtain necessarily the same outcome upon every different repetition.[16] Also thought experiments, inasmuch their repetition does not lead *necessarily* to one indisputable result, can be seen as generative in their repetition [2].

As far as demonstrative, or explanatory experiments are concerned, it follows from the initial argument that we should be mostly dealing with *replication*, for a number of reasons. First of all, whereas the redoing of generative experiments has epistemic concerns (since the previous outcome is what has to be challenged), the redoing of demonstrative experiment must face different constrains: indeed, their outcome is already known and their scope is to disseminate knowledge for the

14 "Scientists do not repeat the same experiment *ad nauseam*. They perform an experiment a 'sufficient' number of times (whatever that might be), and then perform it no more. The experiment becomes a part of history, to be performed again, if at all, only by science students as an exercise" [23, p. 248].

15 This conception was rather absent in early modernity: "Recent methodological frameworks highlight robustness, the importance of multiple determinations of experimental outcomes through a variety of independent procedures. While some parts of Fontana's project could perhaps be reconstructed in hindsight as multiple determinations of experimental results, neither he not Redi [a physician and naturalist at the court of the Grand Duke of Tuscany] explicitly called for independent determinations by different means to make an experimental result more reliable" [32, p. 344].

16 Of course, in the latter case, something must be wrong either in one of the procedures, or in the theorization on which the experiment relied. About this issue, see Sect. 3.2.

benefit of the observers, therefore their peculiar constraints are chiefly *esthetically oriented*.

I do not mean this in a strong sense, *à la* Feyerabend: it is not that experiments carried out in contexts of dissemination are a work of rhetorics. My contention is that the will to reproduce a successful experiment may focus the attention on the reproduction of the *same* procedure, which therefore acquires a ritualized dimension that laminates the epistemic concern.

As a matter of fact, being certain about the outcome (be it an experimental result or an experimentally-confirmed theory) causes a shift in the perspective: the objective is not to redo the experiment to see what happens anymore, but to replicate it in the most convincing and understandable way. This can also be said of actual scientific experiments: sometimes, in the reconstruction of a discovery, when things seem to go too smoothly, it may be the case that a more pleasing *demonstrative* experiment was smuggled in place of the original *generative* one.

Interestingly, Hacking reports an annotation of Maxwell's about the work of Ampère which sums up quite neatly the essence of the replicated demonstrative experiment:

> We can scarcely believe that Ampère really discovered the law of action by means of the experiments which he describes. We are led to suspect what, indeed, he tells us himself that he discovered the law by some process he has not shewn us, and *when he had afterwards built up a perfect demonstration he removed all traces of the scaffolding by which he had raised it* [10, p. 162, added emphasis].

This methodological reconstruction is akin to the one I put forward in Sect. 2.2, by which the demonstrative experiment is somehow deduced from a theory already confirmed as adequate. Consequently, this kind of experimentation (already drawn from a successful experimental confirmation) is ready for replication without excessive worries about the outcome, but rather about its development: if the experiment is carried out correctly, it will be successful and prove our initial hypothesis.[17]

[17] Furthermore, Schickore seems to connect the early-modern care for repetition *in se* with a chiefly demonstrative dimension: "References to multiple repetitions have been interpreted as an echo of an Aristotelian conception of experience; as a literary device to bolster an experimental report; as a literary tool to highlight the wealth of the experimenters' patrons; or as an expression of a general commitment to experience that marked the beginning of modern experimental science" [32, p. 329]. Such an understanding of repetition clearly embeds it in a demonstrative framework akin to the non-epistemic strategies advocated by Magnani's *epistemic warfare* (see footnote 5). Schickore also hints at how repetition, in Galileo, served as a conceptual wrapper to *run* experimental observations as general facts: "Claiming results that accrued from trials repeated 'a full hundred times' was a way of saying 'things *always* behave this way,' and hoping that the reader would believe it" [4, p. 134].

3.2 The Meaning of "Failure"

Repetition and failure are strictly interconnected. As I suggested in the previous subsection, repetitions of *generative* experiments are aimed at testing the outcome of the experiment (and so at testing the hypothesis, theory or measurement that had been carried out during the experiment):

> Our ability to recognize when data fail to match anticipations is what affords us the opportunity to systematically improve our orientation in direct response to such disharmony. Failing to falsify hypotheses, while rarely allowing their acceptance as true, warrants the exclusion of various discrepancies, errors, or rivals, provided the test had a high probability of uncovering such flaws, if they were present. In those cases, we may infer that the discrepancies, rivals, or errors are ruled out with severity [21, p. 18].

If "[a] test 'uncovers' or signals the falsity of *H* by producing outcomes that are discordant with or that fail to 'fit' what is expected were *H* correct" [22, p. 352], then it sparks a procedural loop involving a careful check of the experimental conditions (looking for freak factors), a revision of the hypothesis or—ultimately—a revision of the model [1, 19]. Therefore, in case of experimental failure, the existing tension between the experiment*er* (and her background knowledge) and the experiment*ed* is resolved in favor of the latter, and thus the dignity of the falsifying failure is *respected*. Failure becomes yet another manipulative factor at play in a subsequent experiment. Failures are able to climb back over the experimental framework and crawl inside of general theories from one minimal experimental discrepancy.

> When we falsify a prediction, however "local" it is, we falsify whatever entails that prediction, however general or large-scale. There is, in this respect, no localization of the refuting process. The fact that we may try to find out which part of the refuted whole is to blame is another question–the Duhem question [24, p. 105].

Think of how the inaccurate predictions fostered by Newtonian mechanics about the orbit of Uranus jeopardized the adequacy of Newton's theory in toto: this failure was accepted by Le Verrier, and transformed into new knowledge that managed not only to preserve the adequacy of the theory but also discover a new planet, Neptune.

I suggest, though, that in particular (yet scientific) settings, namely in demonstrative experimental frameworks, Musgrave's claim [24] is wishful, or at best it is the object of a mere lip-service. That is, sometimes a "local" falsification does not affect what entailed the falsified prediction at all. Experiments carried out in schools, for instance, can "not work out" for a number of reasons, in a more or less meaningful way (the phenomenon may not occur, or some measurements might be different). What happens in this cases? Nothing at all.

When a demonstrative experiment fails, the general/expected outcome of the experiment is not questioned. This peculiar "experiment *token*" may have failed, but not the "experiment *type*" it stood for [23, p. 252]. Failure is made into something relative to this peculiar occurrence: it is a matter of *here an now*—this particular experiment failed, but by no means it falsified the theory it was meant to prove. This can be supported by a dialectical interplay with the observers, aimed at illuminating and then filling ignorance bubbles with demonstrative *emergency knowledge*: this process is usually introduced by rhetorical questions along the lines of “Okay, you know why the experiment didn’t work out?”, followed by information—often in-between *ad verecundiam* arguments and plain magical thinking—about the involved instrumentation, secondary phenomena affecting the materiel involved and so on. I label this filling as *magical* because the leading experimenter is saving the expectations of the others by strategically deploying information that was only in her background knowledge: sub-experimenters (for instance pupils, or laypeople visiting a science exhibition) lack the necessary background to make sense of this information, which is therefore offered as self-justifying, or rather justified by the authority of the leading experimenter. There is a significant *appeal to authority* at work in the dissemination of scientific knowledge, even if the latter is presented as immune to authority constraints. Furthermore, it could be said that this *authority overhaul* is *necessary* if only to convey and evoke commitment towards scientific method and its *unconstrained nature*.

One last epistemological effect of this mechanism is worth noting: constructing his argument against the fictionality of models, Magnani [19] contrasts a static understanding of science—for instance the one conveyed by textbooks—with the actual understanding of the dynamic nature of scientific endeavor, and states that if they are seen statically then of course models appear as fiction. The demonstrative experimental framework I described raises the stakes. Demonstrative experiments seem to entail the kind of *fictionalism* that sees models as fictions depicting missing systems [20, 34]. Why? Consider failure in a demonstrative experiment: the unexpected wrong outcome is injected with emergency knowledge (“I am telling you why the experiment did not work out”), and so the model indeed appears as an awkward fiction (the phenomenon that the model should actualize does not happen). Furthermore, a demonstrative failure turns the observed reality into a fiction as well (a missing system, “which you should have seen in the experiment but you didn’t...”), in order to support the cost of the what was to be demonstrated (be it a model, a law, etc.). In case of failure, the tension is resolved in favor of the experimenter and her background knowledge.

What is the final result? Once the observer is faced with a model which underwent a neglected experimental failure (that is, solved through authority-based emergency procedures), she will understand that “there are no actual, concrete systems in the world around us fitting the description it contains” [34, p. 283]. The experimental learning achieves the result of teaching scientific theories as something *necessarily* abstract and incoherent with everyday perceived reality: such a configuration of the experiment awkwardly clashes with Hacking’s breakthrough intuition, according to which experiments (and the models they embed) *create* phenomena that might very well not give themselves in everyday reality [10]. The constructed/modeled nature of phenomena is a consequence of the experimental framework, and not something that the experiment must cope with as the byproduct of the clash between theory and actual reality.

4 Conclusion

The aim of this paper was to provide a sensible analysis of the veritable experimental framework in science. As noted in footnotes 5 and 17, this study is coherent with—and was partially inspired by—Magnani's conception of "epistemic warfare", which sees

> [...] scientific enterprise as a complicated struggle for rational knowledge in which it is crucial to distinguish epistemic (for example models) from non epistemic (for example fictions, falsities, propaganda, etc.) weapons. I certainly consider scientific enterprise a complicated epistemic warfare, so that we could plausibly expect to find fictions in this struggle for rational knowledge. Are not fictions typical of any struggle which characterizes the conflict of human coalitions of any kind? During the Seventies of the last century Feyerabend [5] clearly stressed how, despite their eventual success, the scientist's claims are often far from being evenly proved, and accompanied by "propaganda [and] psychological tricks in addition to whatever intellectual reasons he has to offer" (p. 65), like in the case of Galileo. These tricks are very useful and efficient, but one count is the *epistemic* role of reasons scientist takes advantage of, for example scientific models, which directly govern the path to provide a new intelligibility of the target systems at hand, another count is the *extra-epistemic* role of propaganda and rhetoric, which only plays a mere ancillary role in the epistemic warfare. So to say, these last aspects support scientific reasoning providing non-epistemic weapons able for example to persuade others scientists belonging to a rival coalition or to build and strengthen the coalition in question, which supports a specific research program, for example to get funds [19, p. 3].

Magnani's concept was devised arguing about the use and nature of models in science, but it can be applied fruitfully to the understanding of other aspects of scientific endeavor. Thinking of generative and demonstrative experiments, it can be said that the former reflect epistemic weaponry, while the latter partake of a non-epistemic nature. Yet both kinds of experiment are crucial and unremovable for a correct functioning of science: while generative experiment engage the natural framework, and are thus the first-line of scientific and technological progress, demonstrative experiments engage the human framework. Science is a human activity, therefore a fittingly shaped human framework (eager to invest funds, commitments, priorities etc.) is just as essential as the correct exercise of method and rationality.

The distinction I proposed should not be considered a dichotomy, but rather consists in the two poles of a continuum specter covering the experimental dimension. Even if it is possible to find some experiments (as in Newtonian mechanics) that are carried out only in patently demonstrative settings, there is not a fixed number of repetitions after which an experiment switches from being generative to demonstrative: Popper had already faced this problem, when dealing with the *diminishing returns* from repeated experiments [23].[18] On the other hand, the distinction between the two kinds of experiment is sometimes blurred in the actual scientific practice (not in the dissemination to a lay public): as shown by Ampère's

[18] See also footnote 14.

example in Sect. 3.1 (and other ones in [10]), what I called generative experiment has often had a scaffolding role, and once its outcome is assessed, the scaffolding is replaced by a more straightforward and *nicer* experiment informed by the already confirmed theory. Lastly, demonstrative experiments have a minor (if only nominally) role to play as *watchdogs* of the adequacy of well-assessed theories. That is to say—in Lakatosian terms?—they provide a further protective layer to the protective belt of a research programme: repeating *ad nauseam* experiments about basic chemical reaction, light properties, metal bars that expand when heated and so on, we keep assessing the adequacy of fundamental predictions.

It should be noted that even to consider the distinction as two poles of a continuum is slightly problematic because of some anomalies posed by contemporary sciences: in robotics, computer sciences or for instance genetics most experiments can be generative and demonstrative at the same time. A robot, for instance, is at once the product of the manipulative transformation generating new knowledge, and the mediator of dissemination of that same knowledge. This aspect is worth further studying, as is the relationship between my distinction and thought experiments: thought experiments can be seen at the same time as both generative and demonstrative experiments, depending on the conception of thought experiment rooted in one's background [2, 8]. If one considers thought experiments as reducible to arguments, then she might think of them as *demonstrative*; conversely if thought experiments are seen as rightful experiments, then—no matter how many times a thought experiment is repeated—it could remain *perennially generative*.

Acknowledgments I would like to thank the participants of the MBR'012_Italy conference for the valuable comments on my presentation: many of them contributed to the final form of this paper. In particular, I would like to thank Viola Schiaffonati for the insight about the anomaly of robotics with respect to my distinction.

References

1. Bertolotti, T.: From mindless modeling to scientific models: the case of emerging models. In: Magnani, L., Li, P. (eds.) Philosophy and Cognitive Science: Western & Eastern Studies, pp. 75–104. Springer, Heidelberg (2012)
2. Bishop, M.A.: Why thought experiments are not arguments. Philos. Sci. **66**(4), 534–541 (1999)
3. Cetina, K.K.: Epistemic Cultures. How Sciences Make Knowledge. Harvard University Press, Cambridge (1999)
4. Dear, P.: Revolutionizing the Sciences: European Knowledge and its Ambitions, 1500–1700. Princeton University Press, Princeton (2001)
5. Feyerabend, P.: Against Method. Verso, London-New York (1975)
6. Gabbay, D.M., Woods, J.: The new logic. Log. J. IGPL **9**(2), 141–174 (2001)
7. Galison, P.: Philosophy in the laboratory. J. Philos. **85**(10), 525–527 (1988)
8. Gendler, T.S.: Galileo and the indispensability of scientific thought experiment. Br. J. Philos. Sci. **49**(9), 397–424 (1998)
9. Gooding, D.: Experiment and the Making of Meaning. Kluwer, Dordrecht (1990)

10. Hacking, I.: Representing and Intervening. Introductory Topics in the Philosophy of Natural Science. Cambridge University Press, Cambridge (1983)
11. Hacking, I.: On the stability of the laboratory sciences. J. Philos. **85**(10), 507–514 (1988)
12. Kuhn, T.S.: The Structure of Scientific Revolutions. University of Chicago Press, Chicago, 1962. Second expanded edition, 1970
13. Lakatos, I.: Proofs and Refutations. The Logic of Mathematical Discovery. Cambridge University Press, Cambridge (1976)
14. Latour, J.: Science in Action: How to Follow Scientists and Engineers through Society. Harvard University Press, Cambridge (1987)
15. Macdoland, S., Basu, P. (eds.): Exhibition Experiments. Blackwell, Malden (2007)
16. Magnani, L.: Thinking through doing, external representations in abductive reasoning. In: AISB 2002 Symposium on AI and Creativity in Arts and Science, London, Imperial College (2002)
17. Magnani, L.: Morality in a Technological World. Knowledge as Duty. Cambridge University Press, Cambridge (2007)
18. Magnani, L.: Abductive Cognition: The Epistemological and Eco-Cognitive Dimensions of Hypothetical Reasoning. Springer, Berlin (2009)
19. Magnani, L.: Scientific models are not fictions: model-based science as epistemic warfare. In: Magnani, L., Li, P. (eds.) Philosophy and Cognitive Science: Western & Eastern Studies, pp. 1–38. Springer, Heidelberg (2012)
20. Mäki, U.: MISSing the world. Models as isolations and credible surrogate systems. Erkenntnis **70**, 29–43 (2009)
21. Mayo, D., Spanos, A.: Introduction and background. In: Mayo, D., Spanos, A. (eds.) Error and Inference: Recent Exchanges on Experimental Reasoning, Reliability, and the Objectivity and Rationality of Science, pp. 1–27. Cambridge University Press, Cambridge (2010)
22. Mayo, D.: Explanation and testing exchanges with clark glymour. In: Mayo, D., Spanos, A. (eds.) Error and Inference: Recent Exchanges on Experimental Reasoning, Reliability, and the Objectivity and Rationality of Science, pp. 351–363. Cambridge University Press, Cambridge (2010)
23. Musgrave, A.: Popper and 'diminiscing returns from repeated tests'. Australas. J. Philos. **53**(3), 248–253 (1975)
24. Musgrave, A.: Critical rationalism, explanation, and severe tests. In: Mayo, D., Spanos, A. (eds.) Error and Inference: Recent Exchanges on Experimental Reasoning, Reliability, and the Objectivity and Rationality of Science, pp. 88–112. Cambridge University Press, Cambridge (2010)
25. Nersessian, N.J., Patton, C.: Model-based reasoning in interdisciplinary engineering: cases from biomedical engineering research laboratories. In: Meijers, A.W.M. (ed.) The Handbook of the Philosophy of Technology and Engineering Sciences. Springer, Berlin/New York, Forthcoming (2009)
26. Nersessian, N.J.: How do scientists think? Capturing the dynamics of conceptual change in science. In: Giere, R.N. (ed.) Cognitive Models of Science, Minnesota Studies in the Philosophy of Science, pp. 3–44. University of Minnesota Press, Minneapolis (1992)
27. Nersessian, N.J.: Should physicists preach what they practice? Constructive modeling in doing and learning physics. Sci. Educ. **4**, 203–226 (1995)
28. Picha, M.: How to reconstruct a thought experiment. Organon F **18**(2), 154–188 (2011)
29. Popper, K.R.: The Logic of Scientific Discovery. Hutchinson, London, New York (1959)
30. Raftopoulos, A.: Is perception informationally encapsulated? The issue of theory-ladenness of perception. Cogn. Sci. **25**, 423–451 (2001)
31. Raichvarg, D.: Science on the fairgrounds: from black to white magic. doi: 10.1007/s11191-006-9011-4 (2006)
32. Schickore, J.: The significance of re-doing experiments: a contribution to historically informed methodology. Erkenntnis **75**, 325–347 (2011)

33. Steinle, F.: Entering new fields: exploratory uses of experimentation. Philos. Sci. **64**, S65–S74 (1996)
34. Thomson-Jones, M.: Missing systems and the face value practice. Synthese **172**, 283–299 (2010)
35. Woods, J.: Epistemic bubbles. In: Artemov, S., Barringer, H., Garcez, A., Lamb, L., Woods, J. (eds.) We Will Show Them: Essays in Honour of Dov Gabbay (Volume II), pp. 731–774. College Publications, London (2005)