



Classifying exploratory experimentation – three case studies of exploratory experimentation at the LHC

Peter Mättig¹ 

Received: 21 May 2022 / Accepted: 19 October 2022 / Published online: 15 November 2022
© The Author(s) 2022

Abstract

Along three measurements at the Large Hadron Collider (LHC), a high energy particle accelerator, we analyze procedures and consequences of exploratory experimentation (EE). While all of these measurements fulfill the requirements of EE: probing new parameter spaces, being void of a target theory and applying a broad range of experimental methods, we identify epistemic differences and suggest a classification of EE. We distinguish classes of EE according to their respective goals: the exploration where an established global theory cannot provide the details of a local phenomenon, exploration of an astonishing discovery and exploration to find a new entity. We find that these classes also differ with respect to the existence of an identifiable target and their impact on the background theory. The characteristics distinguish EE from other kinds of experimentation, even though these different kinds have not yet been systematically studied. The formal rigor and precision of LHC physics facilitates to analyze concept formation in its early state. In particular we emphasize the importance for nil–results for conceptualization and argue that conceptualization can also be achieved from nil–results only.

Keywords Exploratory experimentation · Classification of experiments · Particle physics · Conceptualization · Effective field theories

1 Introduction

Exploration is a key scientific practice and percolates all scientific fields and methods. It is the motor of scientific dynamics and as such, understanding the procedures and conditions of exploration is crucial for understanding science. Without exploration, science is reduced to a sterile and static endeavour, even if it may be simpler to understand and formalize. Exploration can be seen everywhere in science. Lenhard (2007) (182) diagnoses an ‘explorative cooperation’ between

✉ Peter Mättig
maettig@physik.uni-bonn.de

¹ Department of Physics, University of Bonn, Bonn, Germany

simulation and data in the sense that the models used in simulation have to be iterated to reproduce the data. Morgan (2003, 218) connects ‘exploring a mathematical relation’ inside a model as tool for theory development and understanding of the world. Sargent (1995) and Elliott (2007) acknowledge explorations in the development of instruments and experimental techniques, while theorists may explore which fundamental concepts (Schnitzer, 2020) can solve a certain problem. As pointed out, for example by Massimi (2019) and Gelfert (2016), models can serve as tools to explore theoretical ideas. Currie (2018) (302f) describes exploratory observation in paleontology and O’Malley (2007) considers the connection of EE to natural history/experimentation (352). Common to these is the characterization of exploration as a systematic study of an object, which is not, at least not well, known. Exploration involves many different paths and attempts and exceeds a single study.

In this paper we will focus on exploratory experimentation (EE). In their seminal papers (Steinle, 1997) and (Burian, 1997) showed that a class of experiments exists that is not theory guided, in contrast to the dominant philosophical school at the time of their publications. Steinle developed this concept along the 19th century experiments of Ampere and Faraday on the back of the then just recently discovered phenomenon of electro–magnetic interactions, for which no conceptual foundation, not to speak of a theory, existed. Burian addressed the contributions of the bio–chemist Brachet to understand the structures of RNA and DNA. Both emphasize the systematic and open procedures in these examples, the development of new concepts, providing the groundwork for Maxwell’s theory of electromagnetism, respectively Watson’s and Crick’s deciphering of the DNA structure some decades later.

We will study EE in the context of particle physics and the Large Hadron Collider (‘LHC’) at the European Center for Particle Physics in Geneva (Switzerland). The LHC is a facility of detectors, notably ATLAS and CMS (Aad et al., 2008; Chatrchyan et al., 2008), and an accelerator (Evans & Bryant, 2008). The LHC is interesting to analyze EE since it is comprehensive and allows physicists to perform many kinds of experimental studies, in particular different classes of EE. Indeed, a large fraction of the by now 1000 publications of each experiment, can be categorized as ‘exploratory’. Furthermore, as the LHC is a huge facility with some 10000 scientists working on a broad range of topics and as it collects some 10 billion of collision events each year, leading to some 10 PBytes of stored data, it is an outstanding instance of Big Science both in data–rate and variety of its research.

Discussing three case–studies representing different EE classes, we will highlight the conditions of EE. While most of our results are quite general for EE, they should also be seen in the context of Big Science.

The main results of our analysis are

- Even though flexibility in instrumentation and operating conditions is reduced, EE is possible and even facilitated by the comprehensiveness of LHC experiments.
- EE can be meaningfully classified along different goals: filling gaps that are inaccessible to a global theory, searching for new, theoretically not defined phenomena, and responding to an astonishing signal.

- To characterize the role of theory in EE we distinguish between background theory, motivations for a measurement and theory of the target system. All classes of EE work within a background theory, they all lack a target theory. However, they are distinct in their theoretical motivations, have a different kind of target, and affect the background theory differently.
- Hypothetico-deductive inferences are not how scientists advance conceptualization at least in its early phases. We will show the importance of nil–results for conceptualization and that it is possible to conceptualize even if only nil–results exist.
- We distinguish EE from other types of experimentation and argue for a more systematic typology.

The paper will proceed as follows. In Section 2 we will summarize the literature on EE and the ensuing questions. After a broad overview of the LHC physics in Section 3, we will discuss three different classes of EE at the LHC. Firstly, at the LHC a phenomenon is explored that is known to exist in a theory but cannot be solved within the theory: the detailed structure of hadron jets (Section 4). Secondly we discuss how the LHC responded to an indication of an astonishing signal (Section 5). Thirdly, in Section 6 we discuss EE to make a discovery by systematically searching for deviations from the predicted kinematic properties of the Higgs boson. These case–studies will finally be used to address questions from the literature survey (Section 7). We conclude on the main lessons in Section 8.

2 Summarizing literature on exploratory experimentation

After being ignored for some time during the 20th century, EE has aroused significant attention during the past decades because of its contrast to the presumed theory guidance of experiments. Hacking (1983), and then Steinle (1997) and Burian (1997), who coined the term ‘exploratory experimentation’, discussed EE as a scientific procedure along historical examples, in which experiments act to a large part autonomously from theory. Hacking (1983) used Herschel’s exploration of radiant heat (177ff), Steinle the explorations of Ampere and Faraday on electro–magnetic interaction, Burian contributions of Brachet to the understanding of RNA and DNA.

Not surprisingly, by addressing different scientific fields and problems, some differing emphases exist. For example, Ampere and Faraday solved their specific problem largely using existing detection technologies, Brachet had to develop and to validate new detection methods. However, there is at least one common principle, emphasized by Elliott (2007) (322): “the most fundamental characteristic of EE seems to be” that “in contrast to other types of experimentation, [it] does *not* serve the aim of testing theories or hypothesis” (highlighted in original).

After Steinle and Burian, other instances of EE have been studied such that a more comprehensive picture about its range and features emerges. EE is mostly analyzed in life–science (Franklin, 2005; Burian, 2007; Elliott, 2007; O’Malley, 2007; Colaço, 2018) and in physics, Cobb (2009); Karaca (2013, 2017); Panoutsopoulos (2019). Further examples can be found in Steinle (2016) (326ff).

2.1 Classifying EE

The case studies show a diversity of EE and suggest a classification. Elliott (2007) (313) identified three properties that give rise to different “types” of EE: “(1) the aim of the experimental activity, (2) the role of theory in the activity, and (3) the methods or strategies employed for varying experimental parameter”. While these properties indeed affect the roles of EE, they are not independent of each other and obey a hierarchy. Instead, we will argue that the role of theory follows from the aim of EE. Furthermore, methods and strategies of parameter variation are not specific to EE but also adopted for other kinds of experimentation. Thus, we consider the aim of EE as the overarching criterion along which, we suggest, EE should be classified.

To do so we will use the term ‘parameter region’ to describe the experimental parameters addressed in experimentation. These parameters characterize the reach and sensitivity of a measurement. In general they may refer to resolution and precision of a measurement, the kinds of objects to be analyzed, special locations or times within a complex system etc.. In particle physics, the field from which we will take our case–studies, they are strongly related to energies, angles, precisions, etc. Progress in experimental science is intimately connected with extending these parameters and particularly, EE always probes extension of known parameter regions by addressing new entities and properties or addressing them with higher precision and variety. There is a relation between the aims of EE and how the parameter regions are extended.

The evaluation of examples of EE in the literature leads us to identify three classes:

[a.] EE for theory deficiency in a parameter region: ‘gap–filling’

Here EE fills gaps in theoretical understanding: the target is part of a complex global system described by a ‘global theory’. However, the global theory by itself cannot describe the target in sufficient depth, no ‘local theory’ exists. Let us list a few examples: Burian (1997) (42) writes “Brachet and his colleagues sought improved understanding of phenomena and sequences of change already partially understood” and later, in his analysis of microRNA he summarizes “current fundamental theories do not contain enough information to provide a functional analysis of RNA regulatory networks” (Burian, 2007) (305). Franklin (2005) addresses high–throughput experimentation where a global, but no local theory existed (893). Colaço (2018) includes experiments, where no “interesting hypotheses” can be derived “from the theory that [researchers] do have” (18).

As will be discussed in more detail later, the reason why the global theory is not sufficient to address the local problem, depends on the problem under consideration. It may be of technical nature or because the substructure of the local target is not known.

[b.] EE extending parameter regions: ‘extension–mode’

In contrast to filling gaps within a chartered parameter region, exploration also transcends such regions. *Per se* knowledge of those is insecure, modelling often uncertain and thus explorations of these new regions are open to discoveries, which add significantly to theory or even change theoretical understanding in a fundamental way. Karaca (2013) discusses how, by increasing the initial energy of electrons scattered on nucleons,¹ the substructure of nucleons was discovered. Variations of experimental parameters in those Deep Inelastic Scattering (DIS) experiments revealed the existence of pointlike partons, which turned out to be identifiable with quarks and gluons. This discovery was astonishing and unintended.

Discoveries also occur by exploring a wide range of processes within the same facility, as many modern instruments are set-up to do (see Section 3.2). Karaca (2017), evaluated how data selection at the LHC prepares for discoveries. Panoutsopoulos (2019) considers how exploration at significantly higher energies to be reached with a planned new particle physics accelerator may foster discoveries. O'Malley (2007) discusses how discoveries emerged from metagenomics studies, where large amounts of DNA from a wide range of environments are analyzed. Today, such broad studies of large data are frequently using artificial intelligence. Examples have been addressed by Leonelli (2014) and Pietsch (2015). In all these cases, the parameter regions are extended by an increased sensitivity to unknown processes.

[c.] Exploring new parameters: 'systemizing astonishing discoveries'

An astonishing experimental discovery calls for systematic measurements of a broad number of properties of such a phenomenon to embed it into a theory. Thus EE is also instrumental after an astonishing discovery has been made. This is the case of Ampere's and Faraday's reaction to Oersted's discovery of electromagnetic interaction as discussed by Steinle (1997, 2016) and Cobb (2009). Also O'Malley (2007) finds that beyond a mere discovery, broad studies lead to "novel theoretical frameworks as well as challenges to old" ones (347). She connects such exploration with 'natural experiments' without human intervention, which can be explored by Big Data studies (352).

In all cases the goals of EE are to gauge theoretically unknown parameter spaces: either theory alone is not sufficient to solve a problem, or there is no place for an observed phenomenon in the existing theory, or one searches for a new phenomenon with unknown properties.²

¹ Nucleons are protons and neutrons, the building blocks of the atomic nucleus. Nucleons are part of the large family of hadrons, which in turn are made up of quarks and gluons.

² This classification of EE's goals differs from the one of Elliott (2007) (324): (1) "identifying regularities and developing new concepts", (2) "isolating or manipulating particular entities or phenomena", (3) "developing experimental techniques, instrumentation or simulations" and (4) "resolving anomalies". Elliott's aims mix methods with aims or outcomes of EE. Instead, we will argue that (1) is the outcome of all kinds of EE, (2) is not special to EE but applies to all kinds of experimentation. Furthermore, his aim (3) is only one of the many other kinds of exploration beyond EE, listed in Section 1. We agree with Elliott's fourth aim, which corresponds to our goal [c.].

2.2 Conceptualization

For Steinle (1997) a “key role” and a “central epistemic goal” of EE is the “the search for general empirical rules and for appropriate representations by means of which they can be formulated” (73), respectively, the “formation and stabilisation of concepts and classification schemes” (72). We will refer to this as ‘conceptualization’. Conceptualization is thus based on results of measurements, which are ordered, classified and correlated. Conceptualization leads at best to phenomenological laws (see also Hacking (1983) (165)) by finding empirically justified correlations between phenomena and their properties. Being pre-theoretical, conceptualization falls short of explaining the correlation, but by constraining the empirical range of correlations, they guide the development towards concepts and theories. Arabatzis (2012) (149) distinguishes two levels of concepts: ‘phenomenological concepts’ that “impose order in a domain of natural or experimentally produced phenomena” (158) and those concepts “that emerge in ... more mature stages of the investigative process”. Although these two stages may not be strictly separable but are part of the iterative process of concept formation, pointed out by Nersessian (2008), we consider them as a valuable, although approximate way of organizing conceptualization. Interestingly, conceptualization is mostly discussed as a one-way street from experiment towards building a theory.

Most discussions of conceptualization deal with Arabatzis’ second stage. Hanson (1960) and Lugg (1985), for example, start from Tycho Brahe’s measurements and discuss how Kepler ordered these into his’ laws. Nersessian (2008) discusses Maxwell’s conceptualization of fields allowing him to develop the fundamental equations of electromagnetism. Less frequently addressed is the role of experiments, which are essential at least for the first step. Exceptions are Steinle (1997) and Cobb (2009). Several studies of conceptualization show that it does not follow a hypothetico–deductive procedure.

The role of conceptualization in EE has different connotations. While Steinle (2016) (334) sees EE to play “a key role in the process of forming and stabilizing concepts and conceptual schemes”, O’Malley (2007) sees conceptualization as a ‘frequent’ part of EE (338) and Elliott (2007), as discussed, sees conceptualization as one of several aims of EE.

2.3 Methods of EE

As mentioned before, the very notion of exploration requires broad studies using different roads into the unknown. Steinle (2016) (331) notes that “individual experiments carry little weight in exploratory experimentation, it is chains, series, or networks of experiments that lead to conclusions”. Burian (1997) argues in the same vein by characterizing Brachet’s research as depending on the “immensely elaborate series of interconnected experiments” using “multiple ways” (43–44). Thus, in EE multiple means and procedures address the same problem to identify

different properties. The experimental practices must be “open for a large variety of outcomes, even unexpected ones” (Steinle, 2002) (307).

While Steinle concluded from his 19th century example that such broadness of EE cannot be performed in fixed facilities, modern experimentation and data analysis has moved beyond this limitation. Multi-purpose facilities of Big Science and high-throughput have implications on how experimental parameters are varied and the bundle of experimental strategies are realized. The advantage of high-throughput instruments – “which allow the simultaneous measure of many features of an experimental system” (Franklin, 2005) (888) – is obviously not only a faster, but importantly a more comprehensive observation of patterns and potential discoveries. In particular sensitivity to rare processes is increased, which are normally not considered. Such broadness and huge data rates are also significant for LHC experimentation that will be outlined in Section 3.2.

Elliott (2007) (324) suggested to classify EE also along a list of methods and strategies for varying parameters. However, the individual methods and strategies are not specific to EE. For example, Big Data, Artificial Intelligence or high-throughput measurements are used for both theory confirming experimentation and EE. However, what is characteristic for EE is the broad application of different methods.

2.4 The role of theory

The discussion and development of the EE concept at the turn of the century was significantly motivated by the rejection of the long-held view of experiments as mere ‘handmaiden’ Schickore (2016) (23) to theories. There is a general consensus that (even) EE is not theory free (Steinle (2016), 330).³ However, “[t]he distinction between exploratory and theory-driven experiments centers not on whether an experiment depends on theory, but on the way(s) in which it depends on theory” (Waters, 2007) (277). The role of theory is thus a central issue for EE and their relation requires a more sophisticated view of the term ‘theory’.

One such qualification is needed for the relation of a ‘global’ to a ‘local’ theory, which is frequently used in the discussion of EE and particularly relevant for our class [a.]. The primary meaning here is the range of applicability: the global theory encompasses more phenomena than the local one. One way to look at this is that a global system A consists of several known local subsystems a_i , which themselves are complex, each having a specific structure. Although the existence of a_i is anticipated in A , its local internal structure cannot be derived from the global theory of A . Thus the structure and mechanisms of a_i and how these affect A may not be understood. In fact, depending on the autonomy of a_i , its detailed substructure may even be of only marginal relevance to understand the whole system A . Such relations between A and a_i are prominently discussed in the EE case studies of life science and cell structures. Another reading is that the elements are the same for local and global theories, but it is impossible to derive and calculate the local phenomenon using the global

³ Steinle clarified in later publications his account of Steinle (1997), which was frequently read as a claim of EE being theory free.

theory, for example, due to the complexity of the local phenomenon. This requires to conceptualize the local phenomenon within the constraints of the global theory. This is sometimes characterized as a relation between a theory and a model.⁴ Indeed, in this sense, sometimes models can be obtained by straight-forward approximations of a theory. However, such an approximation does not always lead to a meaningful model, but has to be significantly complemented by EE. We will discuss such an example in Section 4.1

The other qualification of interest for our study is about the difference between ‘theory’ and ‘thinking’ (Steinle, 2016) (316). Discussing the presumed theory-ladenness of experiments, Hacking (1983) distinguished a weak and a strong notion of theory. The weak one states that one “must have some ideas about nature and [an] apparatus before conducting experiment” (153), a statement “no one would dispute”. He concludes that if “one wants to call every belief etc. a theory, the claim about theory-ladenness is trifling” (175f), see also Steinle (2016) (316). More substantial is Hacking’s strong notion: the experiment is significant only if one “is testing a theory about the phenomenon under scrutiny” (154). Here theory is a worked-out scientific system of a few principles and entities by which multiple different phenomena and dynamical processes can be derived and causally explained. This is the notion we will adopt in the following. We will distinguish three kinds of possible theory impact on experiments.

First: an experiment always builds upon a ‘theoretical background’. Franklin (2005) takes this in her case-study as the systematic knowledge of molecular biology (a ‘global theory’). In his study of electron–proton scattering, Karaca (2013) analyses experiments where “basic instrumental and conceptual requirements to perform experimentation” (126) were used. Even though the existence of a theoretical background is consensus, the terminology is different: Karaca denotes this as ‘theory ladenness in a weak sense’, Franklin speaks of ‘theoretical background’, Waters of ‘theory informedness’ and Colaço (2018) of ‘auxiliary hypotheses’.⁵ We will talk of background theory referring to both the established experimental apparatus and methods and the known theoretical environment on top of which a target phenomenon appears.

Second: a ‘target theory’ denotes the role of the target and predictions for it within a more encompassing theory. Such a theory can be the well established background theory or a perspectival one (Massimi, 2018). This is related to what Waters (2007) denotes ‘theory directed in a strong sense’: “a theory generates expectations about what will be observed when the experiment is conducted” (277). The absence

⁴ We will refrain from discussing differences between theory and model since it has no bearing for our discussion. Therefore, even if the terminology becomes less precise, we will maintain frequently used expressions. For example, we will talk of ‘theory guidance’, although our examples refer to guidance by a model and we will retain the notion of Standard Model of particle physics, although it is frequently considered a theory.

⁵ Colaço (2018) separates experimental methods from background theory and combines them with the motivation into the category of auxiliary hypothesis. We do not adopt this, since using experimental methods for EE requires them to be validated as much as the proper theoretical background.

of a target theory is a defining moment for EE and separates it from experiments geared to theory testing.

Third: Waters (2007) addresses the question, why “experimenters .. search for facts about some things and not about other things” (277) and calls this ‘theory-directed in a weak sense’. Franklin (2005) (891) calls this just ‘theory-directed’ and remarks “[t]he theoretical background serves to guide the explorer to look for certain classes of objects whose activities are known, as a class, to relate to one another, but it need not direct the explorer to one group of those objects over another” (894). Thus theory may play into the motivation of scientists to measure a certain target phenomenon. While in some cases a background theory may motivate, in general background and motivating theory need not be identical. Even more, if a theory may suggest to measure a process, experimentalists may be motivated by different ideas. In fact, there are plenty examples where the motivation to measure something has little to do with theory.⁶ For simplicity we retain the term ‘theoretical’ motivation and consider it as a separate category of theory impact.

Schickore (2016) (23) correctly criticizes that “no terminological agreement has been reached on the question of how to characterize the role of theory for exploratory experimentation”. But even though the terminology is different, there seems to be a consensus about categories of theory impact in EE (and experimental studies in general). For clarity we will adopt the following classification and terminology:

- ‘Background theory’: defines the experimental practice and current theoretical understanding (in literature corresponding to ‘theory-informed’, weak sense of ‘theory ladenness’),
- ‘Motivating theory’: theory that motivates to perform a particular experiment (in literature corresponding to ‘theory-directed’ – in a ‘weak sense’)
- ‘Target theory’: theory about the target phenomenon (in literature also named ‘theory-directed in a strong sense’, ‘strong theory ladenness’, ‘local theory’).

Taking up Waters (2007) notation of the various theory inputs as dimensions, we will talk of ‘theory space’ with the axes given by the three different kinds of theory. The values assumed for these are either discrete (‘existent / non-existent’) but can be continuous (as Waters argues) for the motivating theory. We will return to this in Section 7.2.

2.5 Lines of debates on exploratory experimentation

There has been some debate about what the concept of EE actually encompasses, what, if at all, is special and how it relates to experimentation in general.

⁶ See Mättig and Stöltzner (2020) and Section 6, also O’Malley (2007) (350) states that “[m]arine DNA was sequenced because it was there.... with nothing more theoretical in mind than the general idea it could be similar to bacteriorhodopsin”. Furthermore, although not completely independent of theory, searches for patterns using machine learning for high-throughput data depend less on theory (Pietsch, 2015).

One direction of criticism claims a missing clarity of what does or does not belong to EE. Guralp (2019) misses a “fully worked out account of what might be the central distinguishing aspect of an exploratory experiment” (76), asking if conceptualization, in his words ‘systematic phenomenology’, is a “necessary ingredient or simply one possible outcome”. His concern thus reflects the different views on conceptualization mentioned above. The scope of EE is also criticized by Schickore (2016), who considers it as being “too broad for its intended purpose” (24).

Missing clarity about the definition of EE calls into question its relation to other kinds of experimentation. There is a general agreement that EE brings to the surface questions that have been neglected in philosophy. Guralp (2019) sees EE as one step in the broader task to delineate “how experimental activity relates to theory in general”. In this vein (Waters, 2007) (275) emphasizes the multiple ways of how theory relates to experimentation. He denies a sharp distinction between theory driven and EE but suggests that “the difference between exploratory and theory-driven experimentation might represent continuums” in some dimensions. Schickore (2016) acknowledges the concept of EE serving as a “heuristic tool” that has aided historians and philosophers of science to better characterize diverse experimental practices (23f).

Another question concerns the dynamics and rules of knowledge generation. Schickore (2016) claims that Steinle assigns a ‘new kind of knowledge’ to EE in so far it “generates concepts and classifications rather than confirmations or refutations of hypotheses and theories”. Schickore, recurring to the ‘knowing how’ and ‘knowing what’ categories denies that they are “categorically different from theories or hypotheses; all of these are forms of propositional knowledge” (22). Steinle (2016) on the other hand asks for a more differentiated view to understand in detail “the process by which concepts are stabilized in the course of experimental activity” (321).

We will address these points in this paper along three case studies of EE at the LHC and its abilities to cover a broad range of very different processes and provide our conclusions in Section 7.

3 Physics and experimental conditions at the LHC

To set the scene, we summarize the environment of LHC experimentation, insofar as it is relevant for our three case–studies of EE. Our main observations are:

1. Particle physics is conducted in the context of a highly confirmed background theory, the Standard Model (SM), yet aims to embed the SM in a more encompassing theory, beyond the SM, denoted as ‘BSM’.
2. The sheer size of the detectors and the huge amount and diversity of data result in significantly different epistemic conditions from previous experiments. The LHC is a prominent example for Big Science and high–throughput experimentation.

3.1 The highly confirmed SM and the search beyond

The theoretical concept of the SM of particle physics was developed during the 1960's and 70's (Glashow, 1961; Higgs, 1964; Englert & Brout, 1964; Salam, 1968; Weinberg, 1967; 't Hooft & Veltman, 1972; Fritzsche et al., 1973). It consists of three pillars: exactly twelve elementary matter particles (prominantly electrons and quarks), exactly three interactions (the electromagnetic, weak and strong interactions, all with identical structures), and a sector to generate masses of elementary particles with the higgs h as the one observable particle. During the past decades the main aim of particle physics was to test the SM by finding and precisely measure all its elements. After 50 years of continuously better confirmation, in 2012 the final missing particle, the higgs, was found at the LHC. As of today, all measurements agree with the SM expectation to an astounding precision of typically 10^{-4} up to even 10^{-9} . The SM turned out to be so precise that masses of not-yet observed particles could be inferred from tiny quantum fluctuations, inferences that were eventually confirmed by experiments. Thus particle physics happens in the framework of a very precise background theory. The past successes of the SM do not mean that theory testing has terminated. Indeed both higher energies or higher precision at the LHC and future accelerators, but also at smaller dedicated experiments, probe the SM in new parameter spaces. These will also allow physicists to address the few pieces that have still evaded experimental studies, notably the higgs self-coupling ($h \rightarrow hh$).

And yet, because of internal deficiencies and astrophysical indications, there is a general belief among particle physicists that the SM is just part of a larger, more encompassing theory. For example, the SM can neither explain the number and apparent relations between the matter particles, nor the ones between the interactions. Furthermore it does not include gravitation and has no room for dark matter or dark energy, which are suggested by astrophysical observations. To address these deficiencies, hundreds of BSM models were developed that hypothesized BSM phenomena.⁷

While these models were intensively targeted at the LHC and other experiments, no evidence for BSM was found. Furthermore, principles that guided theoretical development in the past have (as yet) turned out to be futile: hopes to unify all interactions were frustrated by the stability of the proton, cures to avoid the perceived unnaturalness⁸ have not borne out because of the absence of BSM observations at the TeV scale. It is thus not only that nothing has been found, but basic concepts for predicting BSM physics are perceived to have been turned down. Particle physicists

⁷ One of these, the 'pMSSM' model, was addressed by Massimi (2019) as an exploratory model. Beyond a single family of models, we consider the multitude of BSM models as a phase of exploratory modelling: they are devised to try out concepts to see what problems they can solve and if they can meet the observational tests.

⁸ 'Naturalness' was a highly considered theoretical guideline for BSM models. The naturalness criterion avoids corrections through renormalization procedures that by far exceed the value to be corrected. It was high-lighted for the Higgs mass of 125 GeV, which required radiative corrections of the order 10^{18} GeV.

are moving more and more to searches that are ‘model independent’⁹ and expectations that experiments will guide theory (Bechtle et al., 2022). This is underlined in a survey among particle physicists.¹⁰ Asked, for example, about conditions when to give up a BSM model, 35% of particle physicists replied that they have a ‘low commitment’ to any currently discussed model. A similar fraction saw model-independent searches as the ‘most promising way to find signs of New Physics’.

Today particle physicists face the epistemic challenge to find some new effect that is not accounted for by the SM. Discovering BSM is a primary goal of current research possibly transforming the field. Steinle (2016) claims that “in historical case studies performed to date, we do not see researchers starting out with the goal of creating new concepts.” (332) (cp. Kuhn (1996) (59ff)). Currently particle physicists aim to do just this.

3.2 How LHC experiments are performed

At the LHC, high energy protons are made to collide inside detectors, producing a spray of particles. The construction principle of the detectors is to measure at least all SM particles comprehensively and with very high precision. Comprehensiveness of LHC measurements means that a vast range of processes can be measured under different conditions like energies, directions, kinds of particles etc. This allows physicists also to scan spaces of free parameters for many theoretical BSM ideas. The strategy of measuring all kinds of SM particles optimizes sensitivity to BSM effects in proton collisions: time-reversal of the physics laws implies that if BSM particles are produced by the SM particles of proton collisions, they should also decay in SM particles, to be seen as final state products of the interaction. It also takes up the long tradition in experimental physics of observing new effects by signatures of known phenomena.

The details of the two largest LHC detectors are described in Aad et al. (2008) for ATLAS and Chatrchyan et al. (2008) for CMS. In a nut-shell: to measure and identify the SM particles, LHC detectors are structured in different layers, each of which is sensitive to a special property of particles. In addition, LHC detectors cover all directions around the interaction point with just tiny exceptions. Combining the signals from the different layers provides complementary signatures¹¹ for each of the SM particles. While the detectors are primarily optimized to precisely measure SM particles and effects and are huge (and costly),¹² they have a limited flexibility

⁹ This terminology by physicists denotes analyses without a target theory – although, of course, they use background theories.

¹⁰ This is an outcome of a poll in 2018 with replies by some 1500 particle physicists. The unpublished poll was conducted by C.Chall, M.King, P.Mättig and M.Stöltzner within the project ‘Model Dynamics’ of the DFG funded ‘Research Unit Epistemology of the LHC’.

¹¹ As discussed in Mättig and Stöltzner (2019), signatures are stable classes of signals in the detector that can be defined operationally, black-box theory and are related to the concept of phenomena by Bogen and Woodward (1988).

¹² E.g. the ATLAS detector is 40m in length, and 20 m wide. Almost the full volume is covered by materials and electronics.

to also accommodate signatures that were not anticipated in the original design and allow small additions (see e.g. Ariga et al. (2018)).

At the LHC, per year, some 10 billion events are stored, complemented by about 20 billion simulated events. However, it is not only the sheer number of events that is important - the comprehensiveness of the detectors together with the large number of events allow a vast range of processes to be recorded. In terms of the classification by Elliott (2007), LHC experiments are, within just one facility, (a) high-throughput instruments and (b) use multiple experimental techniques to characterize a phenomenon.

The LHC experiments are run by thousands of physicists with a strong division of labour and a distribution of special expertise. Roughly speaking, the LHC community combines theorists, simulationists, scientists for data analysis, for developing and maintaining the software to store and access the data, and finally those that build and operate the detector.¹³

The LHC allows physicists to measure a broad range of different processes, as needed for exploratory experimentation. In the following we will cover three distinct processes, related to the classification suggested in Section 2.1.

4 EE as gap-filling: Hadronization of quarks

Studies of EE in life-science frequently addressed cases, where a system is understood in terms of a global theory anticipating the existence of a subsystem. The global theory by itself, however, is unable to predict the structure of the subsystem, a 'local theory' is missing (see Section 2.1). We find similar situations in case of strong interactions of partons (i.e. quarks and gluons) in particle physics. There exists the highly confirmed and precise 'global' theory of Quantum Chromo Dynamics (QCD), which however, cannot describe features for certain 'local' situations. QCD processes are parametrized in terms of Q^2 , a measure of the hardness of an interaction, and thus, roughly speaking, of the energy of particles emitted from an interaction.¹⁴ Most of the LHC processes of interest are at high Q^2 and can be precisely calculated - in contrast to low Q^2 strong interactions.¹⁵ However, these low Q^2 properties affect how accurately partons can be measured and are thus of relevance for LHC experimentation.

We will focus on one of the strong interaction effects at low Q^2 , deemed 'hadronization', in which the partons at high Q^2 turn into narrow bundles of hadrons,¹⁶ 'jets', at distances of $\mathcal{O}(10^{-14}\text{m})$. QCD can predict the properties of the original partons,

¹³ While these social aspects may be of interest, it is beyond this paper to address these.

¹⁴ More precisely Q^2 parametrizes the difference between the momenta of incoming and outgoing particles. Quantum mechanical arguments relate Q^2 with the inverse of time: $Q^2 \propto 1/t^2$, such that heuristically the evolution of a process from high Q^2 to low Q^2 corresponds to an evolution in time.

¹⁵ At low Q^2 the coupling between partons becomes so large that a perturbative expansion is prohibited.

¹⁶ QCD assigns a 'color charge' to its entities, quarks and gluons, similar to the electric charge of electrons and protons. Hadrons are built from quarks and gluons such that their total color charge adds up to zero (cp. footnote 1).

but not the detailed distribution of hadrons in the emerging jets. Conceptualizing and modelling the structure of jets has been developed by explorative experiments and adopting constraints from QCD (e.g. Mättig (1989)).¹⁷

4.1 Understanding hadronization

The existence of jets was inferred from the global theory of QCD. However, as Field and Feynman (1978), noted, “[t]here is no comprehensive theory of the details of the jet structure”. Instead Field and Feynman attempted to describe jet structure with a so-called ‘hadronization model’. While today there is agreement that the transformation of partons into hadrons in jets is a low Q^2 QCD phenomenon, it is impossible to arrive at a consistent model of such a transformation from QCD. The Field–Feynman model assumed a recursive radiation of hadrons from quarks, parametrizing each radiation by the energy distribution along the initial partons. The model essentially contained three free parameters and could be solved by Monte–Carlo simulation. While basic ideas of the Field–Feynman approach were retained, with emerging measurements and the solidification of QCD, subsequent models were developed incorporating more details and concepts that were better motivated by theory (for example, Andersson et al. (1983)).

Such models provided guardrails for measurements. Within these, experiments embarked fairly autonomously on a broad range of measurements to explore the jet structure. For each of the tens of different hadrons, experiments determined functional dependences of the fragmentation, their different production rates and other properties and measured correlations between those. They found effects that had not been included before – for example, rather high baryon production, emission of photons or production of high spin particles.

These experimental results, combined with QCD motivated ideas and constraints, are today part of computer simulations of proton–proton interactions at the LHC. But even after 40 years of research on jet properties, physicists have not arrived at a single model but instead utilize several hadronization models based on different concepts like the one of Webber (1984). After adjusting tens of model parameters, several models and conceptual ideas agree fairly well with the data.¹⁸

4.2 Theory and experiments on hadronization

Hadronization is understood within the background theory of QCD that in parallel acts as a motivating theory. However, while QCD identifies a target, a target theory of hadronization was missing, making it necessary to measure how hadrons are organized within jets.

¹⁷ Most of what will be discussed in this section has been achieved in experiments before the LHC. However, the results are iterated and controlled at the LHC.

¹⁸ For reviews see Seymour and Marx (2013) or Boge and Zeitnitz (2021a, 2021b).

These measurements do not test a model. Indeed, different models accounting for QCD constraints describe the measurements. Although based on quite different ideas,¹⁹ it is impossible to (dis)confirm them. The main reason is that they are flexible enough to accommodate measurements by adjusting parameters and adding auxiliary hypotheses.²⁰ Instead of aiming to confirm a certain hadronization model, the diversity of adjusted models is used to estimate uncertainties in the analyses. Also, measurements do not simply provide values for free model parameters. Instead by finding new features they play an active role in shaping the understanding of hadronization and the development of models.

In virtue of such broad range of measurements, since these address a physics range that has not been probed in detail before, and since they did not aim to confirm a target theory, they agree with what is considered exploratory experimentation of jets, even though the existence of jets themselves follows from theory.

4.3 Conceptualizing hadronization

In Section 2.4 we discussed the relation of global and local theory in the case where exploration is needed to fill theoretical gaps. Different to cases in life–science, hadronization is an instance where the global theory A (QCD) provides all ingredients for the target phenomenon a_i (jets): ‘jets’ consist of hadrons, which are made out of the partons of A , but how these partons turn into hadrons cannot (as of today) be calculated within the global theory. Instead, measurements are required to allow ordering the production yields and properties of the multitude of hadrons of different flavours, different spins etc. Determining such ordering required to disentangle different effects, which in itself required modelling. However, as discussed in Section 2.2, to arrive at a model or theory, one first needs to conceptualize measurement results. Thus conceptualization of hadronization is an iterative process, similar to what Nersessian (2008) (184) has diagnosed for model–building.

Although the global theory does not permit calculating the local phenomenon of hadronization, by virtue of its power for high Q^2 phenomena, it impacts the conceptualization of hadronization. The measures and ordering principles of hadronization are strongly suggested by the global theory: the very properties of QCD predicted jets as a narrow bundle of hadrons to emerge from partons, which in general terms means high hadron momenta $p_{||}$ along the direction of the bundle, but low momenta p_T transverse to it. Thus, suggestive measures to parametrize hadrons in jet are both $p_{||}$ and p_T .

¹⁹ For example, the model of Andersson et al. (1983) orders recurrent hadron emission along colour–neutral strings of adjustable length, the model of Webber (1984) combines partons into least massive colourless entities and lets them decay according to a statistical procedure. Note that the limited predictive power of these models becomes apparent in the number of parameters needed to describe jet structure within the concept of Andersson et al. (1983), which over the years increased from just a few to over a hundred.

²⁰ There are exceptions: for example, the simple Field–Feynman model was shown to be inferior to the string model of Andersson et al. (1983) and Webber (1984) with regard hadron distributions in more complex topologies.

Conceptualization thus was an interplay between theory–motivated properties and experimental exploration. In contrast to other examples, conceptualization of hadronization did not lead towards a theory, but started from a theory. Acting as constraint, theory “limits the number of possible ways of proceeding, without rigidly specifying the moves one can make within the space of possibilities” as Nersessian (2008) (184) notes in discussing model development. We can therefore talk of a (theory–) ‘constrained conceptualization’ of hadronization.

When no exact and unambiguous theoretical understanding of a phenomenon exists, it is up to experiments to systematically explore it. Only by such fact–finding can a phenomenon be conceptualized. Thus, the example of hadronization meets up with those of Section 2.1. However, having a different relation between A and a_i than those, also the reason for the non–existence of a target theory of a_i is different. It is not that the ingredients of a_i have to be explored, it is just that the mathematical abilities are missing to infer the structure of jets from A .

5 EE for exploring a discovery: The 750 GeV story

If the goal of EE is different, for example, understanding a discovery that had not been anticipated in any theory, its relation to theory and conceptualization is rather different from the previous example. The exploration of the interaction between electricity and magnetism of the early 19th century is the prototypical example that lead (Steinle, 1997) to introduce the term ‘exploratory experimentation’. As yet no such unexpected discovery has been confirmed at the LHC, however, its potential and interesting lessons can be inferred from a brief period when a small indication of a new particle was observed (for a discussion about how the indication developed and was received by physicists, see Ritson (2020)). In the end, it turned out to be a statistical fluke and disappeared after more data was collected. For simplicity, in the following we will talk about the indication as if it were a signal \mathcal{I} (for imposter).

5.1 The rise and fall of \mathcal{I}

At the end of 2015 the ATLAS and CMS experiments presented their preliminary physics results from the first data collected at a significantly increased energy of 13 TeV. Physicists were excited to learn the results since the higher energy allowed for a significantly higher sensitivity to new phenomena, especially at high mass. Both experiments independently showed small indications of an enhancement \mathcal{I} of about 25 events in the distribution of the mass of a pair of two photons at 750 GeV (The ATLAS Collaboration, 2016; The CMS Collaboration, 2016).²¹ Although they

²¹ Such an enhancement, also denoted by physicists as ‘resonance’, signals the existence of a particle.

This episode has to be seen on the backdrop of the previous LHC running at lesser energies culminating in the discovery of the Higgs boson h (see e.g. Chall et al. (2021)). Also the signature of the Higgs was a narrow excess in the spectrum of the invariant mass of two photons at 125 GeV. While first indications for the Higgs were reported with a statistical significance similar to what was seen for the \mathcal{I} , with increased statistics, the Higgs was confirmed with a significance of more than five standard deviations elevating the peak to a ‘discovery’. With the signal at 750 GeV there was a sense of history repeating

emphasized that the significance was too small to claim a discovery and could well be a statistical fluke, in the weeks and months that followed, hundreds of papers were submitted debating how \mathcal{I} could be embedded in a theoretical framework. All theoretical analyses showed that, if the signal would turn out to be significant, it could neither be accommodated by the SM, nor by any previously developed BSM model. Strumia (2016) summarized the situation as “[t]he $\gamma\gamma$ excess is either the biggest statistical fluctuation since decades, or the main discovery”.

In contrast to the frantic thinking by theorists about the meaning of \mathcal{I} , experimentalists were sitting on a fence, knowing that a decisive word on its survival would be spoken with more data to be collected in the next few months. And indeed, in mid 2016, when ATLAS and CMS released new data based on four times the previous statistics, the \mathcal{I} saga came to an end. The excess at 750 GeV was not confirmed, it turned out to be a statistical fluctuation (The ATLAS Collaboration, 2017; The CMS Collaboration, 2018).

In the following we want to address, how \mathcal{I} was explored: how and which data were collected and how conceptualization was approached.

5.2 \mathcal{I} and its theory relation

\mathcal{I} was observed using established experimental strategies and contrasting it to the background theory of the SM. But the Standard Model (SM) has no place for an additional particle. Furthermore, nothing like \mathcal{I} was expected in any of the many BSM models: no target theory of \mathcal{I} existed.

Without theory expectation, why did physicists measure the di-photon mass spectrum at all? In fact, it is a standard analysis at the LHC, because, as Aad et al. (2015) summarizes: “new high-mass states decaying to two photons are predicted in many extensions of the SM, and since searches in the diphoton channel benefit from a clean experimental signature: excellent mass resolution and modest backgrounds”. Thus, searching for a resonance in the di-photon channel is motivated by both (BSM) model testing and, independent of any BSM model, by its experimental virtues of simplicity and cleanliness. While several BSM models suggested the existence of a diphoton resonance, they predicted substantially different properties than were observed for \mathcal{I} .²² Physicists, however, neither needed model predictions nor wanted to be constrained by such. Instead, the diphoton channel was scrutinized for a resonance with a range of masses and properties, only limited by the statistical and systematic constraints of the experiment itself, but independent of any model expectations. To observe \mathcal{I} , experimentalists had to search broadly and to decouple themselves from theoretical motivations.

Footnote 21 (continued)

itself, although at different levels. The Higgs boson was expected, the 750 GeV indication came as a complete surprise.

²² As an example, the popular model suggested by Randall and Sundrum (1999) expected di-photon decays of gravitons, the hypothetical mediators of gravitation. As all models, it predicted a higher production rate and smaller width than observed for \mathcal{I} .

5.3 Conceptualizing \mathcal{I}

While, obviously no concept of \mathcal{I} was finally developed, the physicists' approach of trying to understand \mathcal{I} , allows us to address what Arabatzis (2012) calls first level of conceptualization (see Section 2.2).

5.3.1 Step 1: Starting from data

The first step was to take stock of what is known about \mathcal{I} from measurements. Since measurements at the LHC are comprehensive, facts about properties of \mathcal{I} were immediately obtained even without dedicated analyses. Together with data from previous running ('Run 1') at lower energies, Franceschini et al. (2016) could list: "the anomalous events are not accompanied by significant missing energy,²³ nor leptons [l] or jets [j]. No resonances at invariant mass 750 GeV are seen in the new data in ZZ, l^+l^- , or jj events. No resonances were seen in Run 1 data at $\sqrt{s} = 8$ TeV, although both CMS and ATLAS data showed a mild upward fluctuation at [a mass of] 750 GeV".²⁴

Considering the list, the immediate question is, why did physicists consider these properties important given a vast range of other ones to characterize \mathcal{I} . Why did physicists not choose to mention, say, correlations between \mathcal{I} production and the tides of the moon. As has been frequently diagnosed, the hypothetico–deductive approach falls short in describing (early) conceptualization. What data physicists considered relevant for \mathcal{I} did not emerge from an even tentative conjecture or hypothesis. Take the observation 'no additional leptons' exist. Physicists did not hypothesize ' \mathcal{I} is (not) accompanied by leptons'. They had no reason to even *assume* that \mathcal{I} is (not) connected to leptons. Void of any model to predict \mathcal{I} and a possible associated production with leptons, their procedure cannot be translated in an 'if-then' relation (beyond the trivial tautology 'if the hypothesis is true, \mathcal{I} is (not) accompanied by leptons'). They collected facts without commitment to any of them as a solution for their problems.²⁵ How physicists approached \mathcal{I} is more in line with the observation of Hanson (1960) that in conceptualizing an astonishing discovery, "[n]atural scientists do not 'start from' hypotheses. They start from data" (100). However, they selected the data according to the problem, based on very general principles and restricting them to the scientific field of concern. For example, the time–invariance of physics processes mentioned before, implies that since \mathcal{I} was

²³ 'Missing energy' is a term used by physicists to denote situations, where an energetic particle did not deposit any signal in the detectors, although it entered the detector. Some types of particles like neutrinos interact extremely weakly with matter such that they only very rarely leave a trace in a detector. This can also be true for other (hypothetical) particles. Note that physicists maintain that energy and momentum is conserved.

²⁴ To be precise, claiming 'not seen' means upper limits on its production rates could be derived.

²⁵ Notably there may be one exception: looking for associated 'missing energy' is motivated by one of the primary goals of current physics, to understand Dark Matter, for which 'missing energy' is supposed to be a signature. But even in this case there is no explanatory model to lead to a non-trivial 'if-then' relation.

produced via SM particles, it should also decay into SM particles. Thus a signal of \mathcal{I} was searched for in all kinds of SM particles.

For the first step of conceptualization, physicists extrapolated successful strategies and arguments of the past and applied established general concepts. It is because of such an approach, that gravitationally caused moon–tides were not considered – gravitation is too weak to be relevant at the small scales of LHC physics.²⁶ Adopting a picture of the many–dimensional space of possible correlations, this can be characterized as a ‘nearest–neighbour’ approach – nearest to the background theory. Because no hypothesis or tentative model for the problem solution was at hand, and *a priori* it was not known, which data would be relevant to devise a model, searching for relevant information implied an element of chance. However, the physicists’ strategy was significantly different from random ‘trial and error’, but they systematically adopted established procedures.

5.3.2 Step 2: Embedding in a global framework

Let us now turn to the second step, which addresses what Lugg (1985), in analyzing Kepler’s laws, considers the important question: “how a particular conceptualization was discovered, not whether novel hypotheses can be ‘derived’ from suitably conceptualized data” (209). Indeed all facts on \mathcal{I} listed in the above quote were important for attempting to embed \mathcal{I} in a larger framework (for a summary see Strumia (2016)). Broadly speaking, the observation was approached from two directions. Top–down methods assumed a worked–out model with explanatory power at an energy scale higher than the SM. Alternatively, bottom–up approaches start from the SM and the very general framework of Quantum Field Theory (QFT) to find constraints on \mathcal{I} . The top–down approach, which accords with hypothetico–deductive argumentation and reaches beyond conceptualization, simply did not work for \mathcal{I} , although a plethora of hypotheses were studied. Thus we will focus on the bottom–up approach, which serves to classify the observed properties by supplementing the SM with minimal assumptions.

Instead of speculating about properties of \mathcal{I} , constraints on \mathcal{I} were obtained from fairly general principles of QFT and symmetries of the SM. Here we follow Franceschini et al. (2016). Assuming some new physics at an unknown energy scale Λ , the SM Lagrangian was supplemented by an operator representing \mathcal{I} with an unknown effective coupling. Such a procedure can be systematically developed in the ‘Effective Field Theory’ (EFT) approach.²⁷ Applying this approach to the available experimental information allows one to, in the words of Nersessian “specify moves ... within the space of possibilities”. For example, the low production rate suggests that direct couplings of \mathcal{I} to gluons and light quarks to \mathcal{I} are disfavoured and a quantum

²⁶ It may be indicative of what may come out of explorations and long detailed studies without theory guidance, that moon tides, surprisingly and significantly, affected the energy calibration at LEP, the predecessor of LHC at CERN. However, the accelerator is a large object with a diameter of about 8.5 km, such that gravitational interactions played a role.

²⁷ For the idea and formalism of Effective Field Theories see e.g. Manohar (2020).

loop of heavy new particles should be introduced. The large width, on the other hand, called for additional possible decays of \mathcal{I} , which, in the absence of decays into gluons or light quarks, could be accommodated by significant decays into heavy quarks (bottom or top), with a rate that was allowed by the experimental upper limit.

This analysis does not invoke a concrete target model, does not hypothesize and has *a priori* no immediate bearing on much else than \mathcal{I} . The only assumption is that general principles apply. The use of an EFT is one example of how physicists conceptualize. In this case they start with tools that have been developed and justified from the background theory.

5.3.3 Step 3: From constraints to model building

Only in a third step the general constraints of the EFT approach were translated into possible physical mechanisms and particles to explain the observed \mathcal{I} production and decay. But even for this step, physicists did not start from a coherent model or a physics motivation, but in a kind of bricolage, they merged entities from different perspectival models, none of which anticipated \mathcal{I} .²⁸ Such attempts start providing ideas of further experimental studies, like, for example searching for an entity at higher masses, and they would be an important step towards a model with explanatory power. As it stands, however, while such bricolages can accommodate \mathcal{I} and some properties, model parameter have to be ‘stretched’ making them unlikely (Franceschini et al., 2016) (39). This phase is a kind of exploratory modelling: assuming certain particles - could the observation be accommodated? And does one observe the predicted implications? Only this step starts resembling hypothetico-deductive argumentation with first, rather general and testable hypotheses.

5.3.4 Summarizing conceptualization of \mathcal{I}

The approach to \mathcal{I} reflects the very early steps in conceptualization, much earlier than what is mostly discussed in literature. Also here we diagnose the iterative and reasoned procedure that Nersessian (2008) claims for later stages. Evidently, at each step a wide range of possible ways to conceptualize \mathcal{I} is open, however, physicists follow a strategy starting from established research traditions to eventually narrowing down the range of possibilities by accounting for additional experimental information.

During these first steps of conceptualization neither a theory nor a hypothesis in the sense of “tentative answers to a problem” (Hempel (1966), 17) or being “universal” (Popper (2002), 43) is generated, not to speak of an explanatory model. These would only become relevant with progress in conceptualization.

²⁸ Such tools included prominently heavy vector-like quarks (Aguilar-Saavedra et al., 2013), which have different helicity properties than ordinary quarks, do not couple to the Higgs boson and are frequently used in concrete BSM models. While limits on the number of ordinary quarks can be set from the production rate of the Higgs boson, no such constraints exist for vector-like quarks. This allows one to introduce the required quantum loop from step 2 without violating any experimental and theoretical constraint.

Let us point to an additional element in conceptualizing \mathcal{I} , which becomes more relevant for the case–study in Section 6: the crucial role of negative results. In case of conceptualizing \mathcal{I} itself, the non–observation of decay modes or associated particles contributed significantly to conclusions about the possible character of \mathcal{I} . Here the ‘nearest neighbor’ approach by using EFTs makes this rather transparent.

6 EE to discover: Exploring higgs properties

The previous case study considered the exploration of an astonishing phenomenon that has been observed. In this section we will discuss exploration to find an astonishing phenomenon, one that is not theoretically predicted. As discussed in Section 3.1, particle physics is in a state, where physicists turn towards searching for signs of BSM physics with ‘model–independent’ analyses. We follow this turn and do not consider searches along concrete BSM models.

Without a target model, in principle all observables are candidates for a discovery, a virtually infinite amount of possibilities. Naturally this requires measurements in a broad range of parameter space that has as yet not been probed before, and thus experiments of an exploratory type. In this section we will address how physicists try to organize these searches by strategically targeting analyses that appear most promising to find deviations from the SM. In a first step they divide the SM into different sectors and focus on what they consider a good portal for new physics. Particularly promising appears the higgs sector, which we will take as an example and for which, as characterized in de Florian et al. (2016)(1) “experiments will be able to measure more precisely the kinematic properties of the 125 GeV Higgs and use these measurements to probe for possible deviations induced by new phenomena”. As yet, no significant deviation from the SM has been observed in any SM sector.

6.1 The role of theory

In Section 2.4 we distinguished three kinds of theory impact on experimental analyses. How do these play into the search for deviations in the higgs sector?

In view of the first of our classes of theory impact: evidently, the search for deviations from the SM requires the SM as a background theory. The high precision of both the SM and the understanding of LHC measurements translates into a high sensitivity to potential deviations. The second impact are theoretical motivations. For studies of the top quark at the LHC, Mättig and Stöltzner (2020) found a wide range of motivations from purely pragmatic and factual ones, e.g. from simply that one has the data, to highly theoretical motivations, e.g. the perceived role of the top quark for the stability of the universe. A similar broad range can be identified for studying the higgs particle. Importantly, none of these motivations *per se* necessarily translates into a target model yielding predictions, of where, what kind of, and how strong a deviation should appear.

Even limiting the search for deviations to the higgs sector, a wide range of measurements is possible. Restricting this range further, physicists perform a network

of measurements on higgs production at high Q^2 , motivated by discoveries at high Q^2 in the past. Such measurements have been performed by the ATLAS and CMS experiments using a large number of observables²⁹ and comparing them to the SM prediction in search for deviations (Aad et al., 2020; Aaboud et al., 2018).

6.2 Conceptualizing negative search results

In contrast to the previous examples, here not only a target model is missing, but there is not even a well-defined target. Both jets of Section 4 and the \mathcal{I} of Section 5 had been observed, but in our current example, one does not know if and where deviations show up in the high Q^2 higgs distributions. Thus exploration has to start with nil-results. However, we will argue that even nil-results allow physicists to ‘impose order’, which, as discussed in Section 2.2 is an important element of conceptualization.

To understand conceptualization of the nil-results in Higgs deviations one first has to realize that non-observation can only be claimed within the experimental sensitivity ϵ . That is, to say ‘we have not observed phenomenon P ’, actually means ‘the phenomenon P is produced with a rate that is at most ϵ ’. Thus instead of excluding an observed phenomenon one can only put a constraint on its rate.³⁰

Secondly, physicists are able to organize all SM measurements with operators, where operators that describe BSM have more than four dimensions. This provides an accounting scheme for constraints on deviations from the SM, the most popular one being the SM Effective Field Theory (SM-EFT) (e.g. Manohar (2020) and de Florian et al. (2016)(279ff)).³¹ Each of the (at least) 2499 operators that are relevant for the LHC has an unknown ‘Wilson coefficient’.³² The important point here is that these 2499 operators are sensitive to any possible deviation from SM processes. Measurements, for example the high Q^2 higgs production, provide constraints on the values for some Wilson coefficients, or – in case a significant deviation – their values. As discussed in Bechtle et al. (2022), the SM-EFT has no representational meaning but, in the words of Contino et al. (2013), “[parametrizes] our ignorance”.³³

²⁹ E.g. jet multiplicities, transverse momenta of the Higgs for different jet multiplicities, angles between the produced SM particles etc. All of these reflect high mass scales.

³⁰ Note that this is true for an analysis without a target model. It does not apply if P is predicted in a model to have a certain rate, where one can definitely exclude P , if the experimental sensitivity is sufficient.

³¹ EFTs were already discussed in Section 5.3. However, different to the \mathcal{I} saga, in SM-EFT no new field \mathcal{I} is introduced, synonymous with not identifying a target. Instead, operators of higher dimension are constructed that only contain SM fields. For LHC physics one obtains 2499 relevant operators at the lowest relevant dimension six. Restricting the operators relevant to the Higgs boson, a subset of 66 operators is retained.

³² The Wilson coefficients are parametrized as g_i/Λ , where g_i is the strength of the operator i and Λ the energy scale of BSM.

³³ Although we agree with Koberinski (forthcoming) that measurements interpreted within SM-EFT provide constraints for concrete models, we are more skeptical as regards the representational status of its operators and instead emphasize the advantage of SM-EFT to organize measurements.

Knowing these constraints has important implications for the strategy of further measurements – comparable to conceptualization of an observed signal. They give physicists guidance for measurements and model building: the magnitude of the coefficients, together with the potential experimental sensitivity, suggests which distribution is of potential interest to find deviations. Furthermore, they provide guard-rails for building models: constraints on the Wilson coefficients should not be violated. Thus, the framework of SM-EFT provides physicists with an “orientation and organization” (Steinle, 2016) as required for conceptualization, even if no target exists. Negative results are not only a necessary condition for conceptualization, but a sufficient one, if they are based on a network of measurements and a systematically evaluated for positive results.³⁴

There is another line of argument for the role of nil-results for conceptualization. If, as is commonly anticipated, positive evidence will eventually be found, the current situation is transitory. Thus, compared to what had been discussed for \mathcal{I} , only the sequence of positive and negative evidence changes: instead of conceptualizing by using positive evidence and negative results simultaneously, one now starts with just negative results, waiting for the positive evidence. The major difference between conceptualization of BSM using only negative results and a positive signal like \mathcal{I} is that the former still allows a potentially infinite amount of possible solution, whereas the former singles out one (or a few) solutions.

6.3 Theory testing vs. EE

The broad search for deviations as an instance of EE, brings to the fore the observation of Waters (2007) that “the same experimental procedure might be employed for the purposes of exploratory research in some situations and theory-driven research in others” (280) – see also O’Malley (2007) (351). Indeed, operationally the search for deviations from the SM is identical to its testing. By trying to confirm a theory there is always the potential of observing a deviation. A stark operational demarcation between methods of theory testing and search for deviations from that theory has probably never existed in history.

Instead different emphases between testing and search for deviations apply to the same kind of measurements at different times. Take the SM: roughly speaking, the four decades between the break-through for the quark hypothesis by finding the charm quark in 1973 (Augustin et al. (1974); Aubert et al. (1974), see also Masimi (2007)) and the Higgs discovery in 2012 can be considered as a theory-driven period aiming to confirm the SM. During this time, whole accelerators were built to find particular SM particles, test if their properties agree with the SM and measure

³⁴ The importance of negative results on scientific strategies can also be seen in the failure to find signals of hypothesized concrete BSM models over the past decades. As a result of these nil-results, not only whole classes of models, like supersymmetry, became disfavoured, but they had important repercussions on guiding principles for model development. By far the most prominent one during the past decades was naturalness (see, e.g. Giudice (2018)). By non-observing BSM effects at the TeV scale, naturalness turned from being the holy grail of model-building to an almost neglected principle.

the basic parameters of the SM. In 1978, while the SM was still in the infancy of confirmation, a study characterized the projected LEP accelerator at CERN as “the ideal instrument for a detailed study of the [then hypothetical] Z^0 and more generally for the analysis of the expected merging of weak and electromagnetic interactions” – as conceived in the SM (Jacob, 1978)(7). But Jacob added “[a]n attavistic [sic!] attitude does however call for surprises and still more exotic things to happen.” (ibid.). Today, with all elements of the SM being observed, the proposed new several ten billion euros CERN machine, aims to “[i]mprove by close to an order of magnitude the discovery reach for new particles at the highest masses” (Abada et al., 2019)(1). Testing SM processes is still a key item, but now with the aim to “[p]robe energy scales beyond the direct kinematic reach, via an extensive campaign of precision measurements sensitive to tiny deviations from the Standard Model (SM) behaviour.” (ibid).

But even if theory testing and the search for deviations are operationally identical, they can be distinguished in theory space. As the quotes show, different weights were assigned to the two goals depending on the status of the field. If many features of a theory are not yet experimentally established, the dominant goal of experimentalists (and theorists) is to confirm the missing pieces. With more and more features of a theory being experimentally confirmed, experiments have increasingly the goal to address its inherent short-comings and to explore ways out. True, as Waters and O’Malley argue, the motivation is fluent, but this does not mean theory–testing and exploration are not identifiable as distinct kinds of experimentation. Instead, they differ by the (non)–existence of a target theory. Theory testing is “designed to settle a very precisely specified question” (Elliott, 2007) (332): a target theory is known. In case of searching for deviations in high Q^2 Higgs production, there is no precisely specified question to be settled – the target theory is unknown. Thus, if a target theory is known, experimentation would be testing, if not, it is exploratory (if other requirements are met).

7 Lessons on exploratory experimentation from the LHC case studies

The three case–studies serve us to address questions on the concept of exploratory experimentation outlined in Section 2, particularly in Section 2.5.

7.1 Requirements of EE

Guralp (2019) criticizes the absence of a worked-out account of the distinguishing aspects of EE (76). His criticism has two aspects, firstly there is no clear definition, second, such definition should distinguish EE from others kinds of experimentation.

Let us start with defining EE by setting its requirements, which are together necessary. They are largely agreed upon in literature.

Almost trivially, for some parameter region to be explored it has to be unknown, at least not well known. Thus

- 1 exploratory experimentation addresses a parameter space and properties of physical objects that have not yet been probed by experiment or that have not yet been addressed in sufficient depth.

In addition, for this parameter space,

- 2 EE is performed without a target model.

This is rather unanimously identified as the central aspect of EE. This requirement does not exclude that vague ideas about the new parameter region exist, but these are not the reasons to probe it. The absence of a target model can have several reasons and depends on the scientific problem. For our three case-studies these were inability to solve mathematical equations, an astonishing discovery beyond the current theory and the search for new, yet unknown entities. Importantly, this requirement lends a certain autonomy to experimentation in its methods and goals. The absence of a target theory does not mean EE is theory-free. EE always builds on a background theory, i.e. it is performed on the back of a theoretical framework and established instrumental and experimental methods. Even more, the better the background theory is known, the more meaningful can EE be performed, since the target stands out more clearly.

Furthermore, the notion of exploration implies that EE extends beyond a single measurement, such that

- 3 EE requires a broad and systematic experimental study of a target or a search for such.

How these ‘networks of measurements’ are realized, has changed from the early EE analyses of Steinle. He assumed, in the words of Schickore (2016) (24) that “only the instruments that are not fixed lend themselves to exploratory experimentation”. We have shown that the rather fixed experimental conditions at the LHC provide the required broadness and openness of EE, typical for Big Science and high-throughput measurements.

While requirements 2 and 3 are commonly accepted, Guralp (2019) worries if the role of conceptualization is a “necessary ingredient” for EE, or “simply one possible outcome of it” (76). We will return to this later.

7.2 The distinctiveness of EE

Guralp’s criticism further calls to distinguish EE from other kinds of experimentation. This is seconded by Schickore (2016) (23f), who sees EE as a “heuristic tool” to “better characterize diverse experimental practices”, but is skeptical that it is too broad. Thus, do the requirements meaningfully distinguish EE from other types of experimentation?

Requirement 1 makes EE different from experiments repeating previous measurements either to confirm or improve those or to apply a different method. Examples in

particle physics are measurements of masses or coupling constants, which are often used to reduce experimental uncertainties.

Requirement 2 distinguishes EE from theory-testing experiments, which indeed are abound, like the classical example of the solar eclipse measurement of Eddington and Dyson to test General Relativity, or the tests of all predictions of the SM during the past five decades. These theory-confirming experiments have distinctively different goals than EE. Requirement 2 rejects a purely theory-centered view, which still dominates explicitly or implicitly philosophy of science today.³⁵

The third requirement separates EE from singular experiments that may probe new parameter regions, and are not guided by a theory. An example is the high precision $g - 2$ measurement of the anomalous moment of the muon (see Abi et al. (2021), for a philosophical discussion see Koberinski and Smeenk (2020)). Koberinski (forthcoming) calls the $g - 2$ experiment an “exploratory test” and relates it to EE. While the experiment can be part of the systematic searches for deviations using SM-EFT as discussed in Section 6, by itself, the $g - 2$ should not count as EE.

This brief discussion shows that the requirements of EE are powerful enough to distinguish it from other types of experimentation, which makes Guralp’s criticism appear unjustified. Instead it supports Schickore’s call for a ‘better characterization’ of the diverse ways of experimentation. While to introduce a classification of experimentation is much beyond the paper, we agree that it is underexposed in the philosophical debate: contrast the wealth of model types on the theoretical side listed, for example, in Frigg and Hartmann (2020) with those for experiments, say in Franklin and Perovic (2021). The classification criteria that Elliott (2007) listed for EE may be relevant here, and special purpose type experimentation like the ‘method-driven’ experimentation recently suggested by De Baerdemaeker (2020) also may point to such classification.

7.3 Classes of EE

Since EE can be clearly distinguished from other kinds of experimentation, the broadness of EE that Schickore (2016) criticizes, is not a negative property *per se*. Instead, it suggests to separate EE into different classes. Analyzing EE case-studies in the literature, we suggested in Section 2.1 three classes of EE which we successfully adopted for our case-studies. While the requirements of Section 7.1 as well as the existence of a background theory, are common to all classes, their differences are notable in the role of a target, the theoretical motivation and the impact on the background theory. Let us characterize how classes [a.] to [c.] differ in these respects.

[a.] Targets are identified within a global theory, which, however, cannot provide a local theory of the target. The motivation to explore the local phenomenon

³⁵ Its extreme version may be seen in Popper (2002) who claims the experimentalist “to elicit a decisive answer to these questions [put to him by theoreticians], and to no other. All other questions he tries hard to exclude.” (89, original emphasis).

thus is provided by the global theory. The aim of EE is a conceptualization and parametrization of the local phenomenon, which may lead to a local target model. In most cases this local target model does not change the foundations of the global theory.

- [b.] Targets are confounding experimental discoveries. Thus the targets are known, but a theoretical understanding is missing. As such the background theory does not motivate exploration, apart from the trivial motivation to cure its failure. The outcome of EE and the subsequent embedding of the discovery into a theoretical framework requires fundamental changes to the background model, for example, new structures or ontologies.
- [c.] A target is missing if a new parameter region is systematically probed without a theoretical expectation, possibly leading to an astonishing discovery. The motivations for such measurements could be scientific curiosity or the explicit aim to discover some new fundamental entity due to global deficiencies of the background theory. If indeed something astonishing is discovered, it affects the foundations of the background model.

Surveying the literature, these three classes cover at least a large fraction of EE, although we do not exclude that additional classes may exist. These three types exhibit well identified features and are clearly separable, however, in the process of exploration transitions from one class into another may occur. For example, a successful EE search for astonishing discoveries ([c.]) may lead to EE to understand those ([b.]).

In passing, let us comment on observational surveys that are prominent in astrophysics. We find that the three types of EE (or better exploratory observation) can also be identified. They may explore some target which is known to exist, but at this stage needs more input to develop a model (for example the Sloan Digital Sky Survey on galaxy formation), they may explore instances and conditions of an astonishing discovery, say collecting measurements on dark energy (Dark Energy Survey), or they are equivalent to [c.]: astonishing discoveries, like those of the accelerated expansion of the universe obtained by systemically surveying high redshift galaxies. These examples also obey the requirements we set on EE before.

7.4 Conceptualization in EE

In our discussions we studied the process of concept formation. We saw that because particle physics is highly formalized and because of its very high precision of both theory and experimentation, the steps in concept formation are particularly transparent. Let us discuss some take-aways.

7.4.1 Role of conceptualization for EE types

While in [a.] a global theory is not able to understand a phenomenon in detail, it constrains conceptualization. Here a concept does not preexist a theory, as in most other case–studies, but vice versa: conceptualization starts from a global theory. However, the global theory just provides boundaries, which have to be filled by experimental measurements. As mentioned before, sometimes such a situation may also be solved by mere model–building without further exploration, e.g. by setting a theory parameter to some limiting value. But often this does not lead to a viable concept and model. Instead the phenomenon may be too complex to be conceptualized without experimental exploration. Here the role of EE is to collect facts that need to be ordered. The empirical rules to be developed connect to, without being derivable from, the global theory. In fact attempting such derivation may even not be pursued if no fundamental new insights are expected but high operational investment is required.

In contrast, for our case–studies [b.] and [c.], conceptualization reaches beyond the foundations of the background theory. They are the classes, for which Carrier (1998) (184) stated, they would typically open up new areas of research. Here conceptualization is a step towards extending the background theory or embedding it into a larger framework. This makes EE an important part of the scientifically most interesting periods, where the dynamics of science becomes manifest. Such cases have been frequently discussed in, for example, Hanson (1960); Lugg (1985) or Nersessian (2008). However, their discussions take up mature states, in which all relevant measurements of the phenomenon are available such that conceptualization is the immediate preparation of an encompassing theory. In contrast, we discuss conceptualization at a more infant state, where significant experimental information on phenomena is missing but has to be collected and eventually ordered.

All our cases underline that also in the absence of a theory, physicists approach conceptualization in a very organized and systematic way instead of a random trial–and–error procedure. Within such a systematic approach, certain ideas may fail, but these play a significant role to develop a viable concept.

7.4.2 Conceptualization without hypotheses

The absence of a target theory in EE makes it at least difficult to develop a testable hypothesis, contrasting it to the commonly assumed hypothetico–deductive argumentation. In our case–studies of the early state of conceptualization, physicists did not hypothesize certain properties of a phenomenon, they did not propose specific new measurements or develop a prediction that could be tested. Instead, taking the case–study of Section 5, they embedded the observation of \mathcal{I} and a collection of measurements into the framework of Effective Field Theory, allowing physicists to derive constraints on properties, \mathcal{I} could or could not have. Doing so, underlines that before developing concepts of phenomena that may allow one to express hypotheses, one has to collect facts about it. This reflects the observation of Hanson (1960) that scientists start from facts. But then, without a concept, it is not clear *a priori*, which facts should be collected. We saw in our examples that scientists started out from

facts close to the established background theory, thereby implicitly acknowledging and building upon the past progress of their field. ‘Nearest neighbor’ properties of the phenomena were collected to embed them into a framework, but not following a testable hypothesis.

Such limitation of a hypothetico–deductive procedure towards a discovery has been frequently shown, for example, by Hanson (1960) and Lugg (1985) for Kepler’s derivation of ‘his’ laws. Also O’Malley (2007) in her analysis of proteorhodopsin finds that “the less controllable nature of the phenomena means that strict hypothesis testing is practically impossible” (352).

7.4.3 Conceptualizing Nil–results

The absence of a hypothetico–deductive procedure for conceptualization becomes obvious in case–study [*c.*]. Here the particular challenge is to find an order in nil–results, at least for a transition period. As discussed in Section 6.2, this is possible by building upon very general principles of the background theory. Conceptualization in this context allows physicists to identify research paths, both in experiments and model–building, which appear less promising to reach the goal, or those which amplify possible observations.

The importance of nil–results has also been discussed in other examples of conceptualization, however, with a different flavour. For example, in the context of Kepler’s attempts to find the order in Tycho Brahe’s measurements of positions of planets with time, Hanson (1960) (105) concludes “ways in which scientists sometimes reason their way towards hypotheses, by eliminating those which are certifiably of the wrong type, may be as legitimate an area for conceptual inquiry as are the ways in which they reason their way from hypotheses”. Here a nil–result is a failed hypothesis, not a non–observation as in our case. It is also different from nil–results as obtained by the Michelson–Morley experiment, which refuted a concrete model prediction (of the existence of an ether) and could be considered a classical experiment for testing a hypothesis. Our example differs from Kepler and Michelson–Morley. Firstly, we do not consider single nil–results, but a broad and different variety. Secondly, no model prediction exists for a phenomenon, instead the nil–results are hoped to help to eventually lead to a model.

7.4.4 The need of conceptualization for EE

After analyzing the relation of conceptualization to EE, we now return to the question of Guralp (2019) if conceptualization should be seen as a defining requirement for EE. For [*a.*] and [*b.*] exploratory experimentation has the explicit aim of conceptualization. Even though this is different for [*c.*], also in this case it results in an organization of the measurements and a quantitative assessment. Thus, even if conceptualization may not be a defining requirement for EE, it is integral part and natural outcome and necessarily connected to EE. It is unreasonable to assume that scientists would develop a network of experiments to just collect facts and not try to order their findings.

This is underlined by examples outside of LHC physics: as discussed by Franklin (2005), the microarray survey is followed by “looking for and finding similarities between them that they later hypothesized explained their similar behaviors” (894). Also the discovery of partons, discussed by Karaca (2013), was systemized and organized by subsequent measurements.

One may even argue that conceptualization builds upon the requirements for EE. Scientists will conceptualize predominantly new measurements – for which no established target theory exists. Instead, conceptualization is a first step to devise such a target theory. Finally, conceptualization can only work if a broad number of facts are known. The three requirements for EE are thus preconditions for conceptualization. Since conceptualization follows from these requirements of EE, it should not count as a requirement of EE itself.

8 Conclusion

Exploratory experimentation at the LHC highlights the conditions and methods of EE in the framework of Big Science and high-throughput experiment. Early accounts of EE emphasized the need for measurement facilities to be flexible, to be able to devise new ones etc.. In contrast, we observe that the broadness of experimentation and the high-throughput at the fixed LHC facilities is beneficial for EE.

In general agreement with discussions on exploratory experimentation in the literature we defined EE by three requirements: probing a new region of parameter space, the absence of a theory of the target and a broad range of measurements. We find that these requirements distinguish EE from other kinds of experimentation, which may be interesting for an eventual comprehensive analysis of different kinds of experimentation and their respective roles for scientific progress.

While these requirements make EE well defined and distinct, they are also broad and encompass experimentation with different qualities and thus suggest to classify EE. We suggest three classes that align with the case-studies in the literature and EE at the LHC. The overarching criterion to differentiate between the classes is their respective (theoretical) motivation to explore: either they fill a local gap in an otherwise well understood global theory, or they explore properties of an astonishing discovery in conflict with the global theory, or they aim for a discovery that is not predicted by theory. Beyond their (theoretical) motivation, this classification implies other differences: some classes have a known target, some do not, also they have different repercussions on the background theory. Thus, the three classes are clearly identifiable and play different roles in the scientific inquiry.

Our case-studies also reveal how concepts of a target system are formed in a very early state. Indeed, in almost all examples of EE discussed in literature, EE triggered conceptualization. This may not be surprising in view of the large overlap between the requirements of EE with those needed for concept formation, at least in its early stage. Conceptualization thus has to be seen as a consequence from EE but, in contrast to some opinions, it does not seem meaningful to count it as a requirement for EE. We also saw in our examples that to a large extent, concept formation did not start by hypothesizing. The absence of a target model, or even a general idea

about a concept of the target, makes it is difficult to formulate hypotheses. Thus, our case–studies support observations that a hypothetico–deductive approach is not followed, although it certainly will become relevant at more mature stages of concept formation. Instead of trying to hypothesize, what is seen as center-stage in our examples, is the collection of relevant measurements, which eventually should be ordered. These relevant measurements were collected based on established research practices, acknowledging the scientific progress reached up to now. Our case–studies further show the high importance of nil–results. Concept formation does not only work with positive evidence, but of similar value are constraints that identify possible concepts to be dead–ends. We discussed case–studies where nil–results worked together with positive evidence to lead to concept formation. However, we also showed that conceptualization in EE in the absence of positive evidence, but with a broad range only nil–results is possible, meaningful, and an important means to define scientific strategies.

Acknowledgements This paper was written with the support of the German Research Foundation (DFG) and as part of the Research Unit ‘The Epistemology of the Large Hadron Collider’ (FOR 2063). The author is grateful for comments by many members of the Research Unit, notably those by H.Sorgner, F.Steinle and M.Stöltzner. He profited from comments by two anonymous reviewers. The author acknowledges in particular many discussions with S.Ritson in the early phase of the project. Her contributions to the literature review and her many constructive remarks helped to clarify several important ideas.

Funding Open Access funding enabled and organized by Projekt DEAL. Research Group ‘The Epistemology of the Large Hadron Collider’, funded by the German Science Foundation DFG (grant FOR 2063)

Data availability Not applicable

Code availability Not applicable

Declarations

Conflicts of interest/Competing interests Not applicable

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Aaboud, M., et al. (2018). Measurements of Higgs boson properties in the diphoton decay channel with 36 fb^{-1} of pp collision data at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS detector. *Physical Review D*, 98, 052005.
- Aad, G., et al. (2008). The ATLAS experiment at the CERN Large hadron collider. *JINST*, 3, S08003.
- Aad, G., et al. (2015). Search for high-mass diphoton resonances in pp collisions at $\sqrt{s} = 8 \text{ TeV}$ with the ATLAS detector. *Physical Review D*, 92(3), 032004.

- Aad, G., et al. (2020). Higgs boson production cross-section measurements and their EFT interpretation in the 4ℓ decay channel at $\sqrt{s} = 13$ TeV with the ATLAS detector. *The European Physical Journal C*, 80(10), 957.
- Abada, A., et al. (2019). FCC physics opportunities: Future circular collider conceptual design report volume 1. *The European Physical Journal C*, 79(6), 474.
- Abi, B., et al. (2021). Measurement of the positive muon anomalous magnetic moment to 0.46 ppm. *Physical Review Letters*, 126(14), 141801.
- Aguilar-Saavedra, J., Benbrik, R., Heinemeyer, S., & Pérez-Victoria, M. (2013). Handbook of vectorlike quarks: Mixing and single production. *Physical Review D*, 88(9), 094010.
- Andersson, B., Gustafson, G., Ingelman, G., & Sjostrand, T. (1983). Parton fragmentation and string dynamics. *Physics Reports*, 97, 31–145.
- Arabatzis, T. (2012). Experimentation and the meaning of scientific concepts. In U. Feest & F. Steinle (Eds.), *Scientific concepts and investigative practice* (pp. 149–166). De Gruyter.
- Ariga, A. et al. (2018). Technical proposal for FASER: Forward search Experiment at the LHC. arXiv:1812.09139.
- Aubert, J., et al. (1974). Experimental observation of a heavy particle. *J. Physical Review Letters*, 33, 1404–1406.
- Augustin, J., et al. (1974). Discovery of a narrow resonance in e^+e^- annihilation. *Physical Review Letters*, 33, 1406–1408.
- Bechtel, P., Chall, C., King, M., Krämer, M., Mättig, P., & Stöltzner, M. (2022). Bottoms Up: The standard model effective field theory from a model perspective. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*. <https://doi.org/10.1086/710618>.
- Boge, F. J., & Zeitnitz, C. (2021). Polycratic hierarchies and networks: what simulation-modeling at the LHC can teach us about the epistemology of simulation. *Synthese*, 199(1), 445–480.
- Boge, F. J., & Zeitnitz, C. (2021). Correction to: Polycratic hierarchies and networks: what simulation-modeling at the LHC can teach us about the epistemology of simulation. *Synthese*, 199(3), 11767–11768.
- Bogen, J., & Woodward, J. (1988). Saving the phenomena. *Philosophical Review*, 97(3), 303–352.
- Burian, R. M. (1997). Exploratory experimentation and the role of histochemical techniques in the work of Jean Brachet, 1938–1952. *History and Philosophy of the Life Sciences*, 19(1), 27–45.
- Burian, R. M. (2007). On MicroRNA and the need for exploratory experimentation in post-genomic molecular biology. *History and Philosophy of the Life Sciences*, 29(3), 285–311.
- Carrier, M. (1998). New experimentalism and the changing significance of experiments: On the shortcomings of an equipment-centered guide to history. In M. Heidelberger (Ed.), *Experimental essays: Versuche zum Experiment* (pp. 175–191). Nomos Verlagsgesellschaft.
- Chall, C., King, M., Mättig, P., & Stöltzner, M. (2021). From a boson to the standard model Higgs: a case study in confirmation and model dynamics. *Synthese*, 198(16), 3779–3811.
- Chatrchyan, S., et al. (2008). The CMS Experiment at the CERN LHC. *JINST*, 3, S08004.
- Cobb, A. D. (2009). Michael Faraday's "historical sketch of electro-magnetism" and the theory-dependence of experimentation. *Philosophy of Science*, 76(5), 624–636.
- Colaço, D. (2018). Rethinking the role of theory in exploratory experimentation. *Biology & Philosophy*, 33(5), 38.
- Contino, R., Ghezzi, M., Grojean, C., Muhlleitner, M., & Spira, M. (2013). Effective Lagrangian for a light Higgs-like scalar. *JHEP*, 07, 035.
- Currie, A. (2018). *Rock, bone, and ruin: An optimist's guide to the historical sciences*. Cambridge: The MIT Press.
- De Baerdemaeker, S. (2020). Method-driven experiments and the search for dark matter. *Philosophy of Science*, 88(1), 124–144.
- de Florian, D. et al. (2016). Handbook of LHC higgs cross sections: 4. Deciphering the nature of the Higgs sector. arXiv:1610.07922.
- Elliott, K. (2007). Varieties of exploratory experimentation in Nanotoxicology. *History and Philosophy of the Life Sciences*, 29(3), 313–336.
- Englert, F., & Brout, R. (1964). Broken symmetry and the mass of gauge vector mesons. *Physical Review Letters*, 13, 321–323.
- Evans, L., & Bryant, P. (2008). LHC machine. *Journal of Instrumentation*, 3(08), S08001–S08001.
- Field, R., & Feynman, R. (1978). A parametrization of the properties of Quark jets. *Nuclear Physics B*, 136, 1.

- Franceschini, R., Giudice, G. F., Kamenik, J. F., McCullough, M., Pomarol, A., Rattazzi, R., et al. (2016). What is the $\gamma\gamma$ resonance at 750 GeV? *Journal of High Energy Physics*, 2016(3), 144.
- Franklin, A., & Perovic, S. (2021). Experiment in physics. In E.N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/sum2021/entries/physics-experiment/>.
- Franklin, L. R. (2005). Exploratory experiments. *Philosophy of Science*, 72(5), 888–899.
- Frigg, R., & Hartmann, S. (2020). Models in science. In E.N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/spr2020/entries/models-science/>.
- Fritzsche, H., Gell-Mann, M., & Leutwyler, H. (1973). Advantages of the color octet gluon picture. *Physics Letters B*, 47, 365–368.
- Gelfert, A. (2016). *How to do science with models. A philosophical primer*. Berlin: Springer.
- Giudice, G.F. (2018). The dawn of the post-naturalness era. In A. Levy, S. Forte, & G. Ridolfi (Eds.) *“From my vast repertoire ...: Guido Altarelli’s Legacy”* (pp. 267–292). World Scientific.
- Glashow, S. L. (1961). Partial symmetries of weak interactions. *Nuclear Physics*, 22, 579–588.
- Guralp, G. (2019). Essay review of Exploratory experiments: Ampere, Faraday, and the origins of electrodynamics, by Friedrich Steinle. *Studies in History and Philosophy of Science Part A*, 73, 72–76.
- Hacking, I. (1983). *Representing and intervening: Introductory topics in the philosophy of natural science*. Cambridge: Cambridge University Press.
- Hanson, N. R. (1960). Is there a logic of scientific discovery? *Australasian Journal of Philosophy*, 38(2), 91–106.
- Hempel, C. G. (1966). *Philosophy of natural science*. Englewood Cliffs, N.J.: Prentice-Hall.
- Higgs, P. W. (1964). Broken symmetries and the masses of gauge bosons. *Physical Review Letters*, 13, 508–509.
- ’t Hooft, G., & Veltman, M. J. G. (1972). Regularization and renormalization of gauge fields. *Nuclear Physics B*, 44, 189–213.
- Jacob, M. (1978). LEP summer study: Summary report. In *LEP summer study* (pp. 1–34). CERN Yellow Report 79-01
- Karaca, K. (2013). The strong and weak senses of theory-ladenness of experimentation: Theory-driven versus exploratory experiments in the history of high-energy particle physics. *Science in Context*, 26(1), 93–136.
- Karaca, K. (2017). A case study in experimental exploration: Exploratory data selection at the large hadron collider. *Synthese*, 194, 333–354.
- Koberinski, A. & Smeenk, C. Q.E.D., QED (2020). *Studies in history and philosophy of science part B: Studies in history and philosophy of modern physics*, 71: 1–13
- Koberinski, A. (forthcoming). “Fundamental” “constants” and precision tests of the standard model. *Philosophy of Science*. <https://doi.org/10.1017/psa.2022.41>.
- Kuhn, T. (1996) [1962]. *The Structure of Scientific Revolutions*. ISSR collection. University of Chicago Press
- Lenhard, J. (2007). Computer simulation: The cooperation between experimenting and modeling. *Philosophy of Science*, 74(2), 176–194.
- Leonelli, S. (2014). What difference does quantity make? On the epistemology of Big Data in biology. *Big Data & Society*, 1(1), 2053951714534395.
- Lugg, A. (1985). The process of discovery. *Philosophy of Science*, 52(2), 207–220.
- Manohar, A. V. (2020). Introduction to effective field theories. *Les Houches Lecture Notes*, 108. <https://doi.org/10.1093/oso/9780198855743.003.0002>.
- Massimi, M. (2007). Saving unobservable phenomena. *The British Journal for the Philosophy of Science*, 58(2), 235–262.
- Massimi, M. (2018). Perspectival modeling. *Philosophy of Science*, 85(3), 335–359.
- Massimi, M. (2019). Two kinds of exploratory models. *Philosophy of Science*, 86(5), 869–881.
- Mättig, P., & Stöltzner, M. (2019). Model landscapes and event signatures in elementary particle physics. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*.
- Mättig, P. (1989). The structure of jets in e^+e^- collisions. *Physics Reports*, 177, 141.
- Mättig, P., & Stöltzner, M. (2020). Searching for signatures. *Philosophy of Science*, 87(5), 1246–1256.
- Morgan, M. S. (2003). Experiments without material intervention. In H. Radder (Ed.), *The philosophy of scientific experimentation* (pp. 216–235). University of Pittsburgh Press.
- Nersessian, N. J. (2008). *Creating scientific concepts*. Cambridge: MIT Press.

- O'Malley, M. A. (2007). Exploratory experimentation and scientific practice: Metagenomics and the prothorhodopsin case. *History and Philosophy of the Life Sciences*, 29(3), 337–360.
- Panoutsopoulos, G. (2019). The Culture of exploratory experimentation at CERN and the future colliders. arXiv:1907.12946.
- Pietsch, W. (2015). Aspects of theory-ladenness in data-intensive science. *Philosophy of Science*, 82(5), 905–916.
- Popper, K. (2002) [1959]. *The logic of scientific discovery*. Routledge.
- Randall, L., & Sundrum, R. (1999). A Large mass hierarchy from a small extra dimension. *Physical Review Letters*, 83, 3370–3373.
- Ritson, S. (2020). Probing novelty at the LHC: Heuristic appraisal of disruptive experimentation. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 69, 1–11.
- Salam, A. (1968). Weak and electromagnetic interactions. *Proceedings of the Nobel Symposium held 1968 at Lerum, Sweden*, pp 367
- Sargent, R.-M. (1995). *The diffident naturalist*. Chicago: University of Chicago Press.
- Schickore, J. (2016). 'Exploratory experimentation' as a probe into the relation between historiography and philosophy of science. *Studies in History and Philosophy of Science Part A*, 55, 20–26.
- Schnitzer, H.J. (2020). The crucial calculation as a motivating force in particle physics. arXiv:2012.08887.
- Seymour, M.H., Marx, M. (2013). Monte carlo event generators. In *69th Scottish universities summer school in physics: LHC Physics* (pp. 287–319). https://doi.org/10.1007/978-3-319-05362-2_8.
- Steinle, F. (2002). Challenging established concepts: Ampere and experimentation. *Theoria: An International Journal for Theory, History and Foundations of Science*, 17(2), 291–316.
- Steinle, F. (1997). Entering new fields: Exploratory uses of experimentation. *Philosophy of Science*, 64, S65–S74.
- Steinle, F. (2016). *Exploratory experiments: Ampère, faraday, and the origins of electrodynamics*. Pittsburgh: University of Pittsburgh Press.
- Strumia, A. (2016). Interpreting the 750 GeV digamma excess: a review. In *51st Rencontres de Moriond on EW interactions and unified theories* (pp 407–426). ARISF
- The ATLAS Collaboration (2016). Search for resonances in diphoton events at $\sqrt{s} = 13$ TeV The ATLAS Collaboration (2016). Search for resonances in diphoton events at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Journal of High Energy Physics*, 2016(9)
- The ATLAS Collaboration. (2017). Search for new phenomena in high-mass diphoton final states using 37 fb of proton-proton collisions collected at $\sqrt{s}=13$ tev with the atlas detector. *Physics Letters B*, 775, 105–125.
- The CMS Collaboration. (2016). Search for resonant production of high-mass photon pairs in proton-proton collisions at $\sqrt{s} = 8$ and 13 TeV. *Physical Review Letters*, 117(5), 051802.
- The CMS Collaboration. (2018). Search for physics beyond the standard model in high-mass diphoton events from proton-proton collisions at $\sqrt{s} = 13$ TeV. *Physical Review D*, 98, 092001.
- Waters, C. K. (2007). The nature and context of exploratory experimentation: An introduction to three case studies of exploratory research. *History and Philosophy of the Life Sciences*, 29(3), 275–284.
- Webber, B. (1984). A QCD model for jet fragmentation including soft gluon interference. *Nuclear Physics B*, 238, 492–528.
- Weinberg, S. (1967). A model of Leptons. *Physical Review Letters*, 19, 1264–1266.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.