Higgs Discovery and the Look Elsewhere Effect

Richard Dawid[1]

The discovery of the Higgs particle required a signal of five sigma significance. The rigid application of that condition is a convention that disregards more specific aspects of the given experiment. In particular, it does not account for the characteristics of the look elsewhere effect in the individual experimental context. The paper relates this aspect of data analysis to the question as to what extent theoretical reasoning should be admitted to play a role in the assessment of the significance of empirical data.

1: Introduction

In July 2012, CERN announced the discovery of a scalar particle that was likely to be the Higgs particle, the last by then unconfirmed prediction of the standard model of particle physics (ATLAS 2012, CMS 2012). Based on further improvements of the data, this discovery has by now been acknowledged as the discovery of a Higgs particle.[2]

Already in December 2011, CERN had first announced indications for a Higgs particle at both LHC experiments, ATLAS and CMS, with a combined significance of nearly 4σ. Since the agreed upon statistical limit for acknowledging a discovery in particle physics was 5σ, that data amounted to significant evidence for a new scalar particle but did not constitute a discovery. The present paper focuses on the epistemic status of the Higgs particle between December 2011 and July 2012. Assessments of the December 2011 data could be divided into two clearly discernible 'camps'. Those who adhered to the letter of the definition of a discovery in particle physics recommended caution and warned against being overly confident based on insufficient data. Others, however, emphasized the striking coherence of the data with theoretical knowledge about the Higgs particle and, on that basis, argued that the evidence for the Higgs was stronger than suggested by the formal statistical analysis. Motivations for leaning towards one or the other understanding were diverse and arguably included subjective elements, ranging from personal dispositions to the specific requirements of the individual agents' professional roles within the research process. However, the disagreement did reflect a substantial conceptual difference of opinion regarding the overall take on empirical confirmation: it hinged on the question how to evaluate the 'look elsewhere effect' in the given case. The present article analyses this conceptual disagreement, which points towards a more general question about the epistemological status of measurement in high energy physics and may be expected to be of increasing importance in the future.

After a brief characterization of the Higgs particle and the look elsewhere effect in Sections 2 and 3, the core problem associated with interpreting the data is presented in Section 4 by defining two different perspectives on the status of measurements in high

---

[1] Univ. of Vienna, Dep. of Philosophy and Institut Vienna Circle. email: richard.dawid@univie.ac.at

[2] It remains to be seen whether the discovered particle has the properties predicted by the standard model of particle physics or must be understood in terms of an extension of the standard model, such as supersymmetry.

energy physics. This discussion is followed by a brief assessment of the significance of the ensuing analysis in Section 5. Section 6 construes the two perspectives in terms of the distinction between a frequentist and a Bayesian understanding of statistical analysis. Section 7 then proceeds to discuss potential problems of both perspectives and builds a connection to the novel confirmation debate in philosophy of science.

2: The Higgs Particle and its Discovery

The standard model of particle physics introduces the Higgs mechanism in order to explain the observed mass spectrum of fermionic and bosonic particles. Nuclear interactions are conceptualized as gauge interactions, that is in terms of a Boson exchange that satisfies a local gauge symmetry. A gauge symmetry, if not broken, would enforce massless gauge bosons and equal masses for the fermions that are connected by the gauge symmetry. The fact that W- and Z-bosons, the gauge Bosons of the weak interaction, are massive and the fermions connected by the electroweak gauge symmetry have different masses therefore implies that the electroweak gauge symmetry must be broken. In order to retain the advantages of gauge theory, specifically its capability to guarantee the renormalizability of the theory and therefore to provide the basis for consistent quantitative calculations of scattering processes, it was suggested in the 1960s that the electroweak gauge symmetry should be broken at the theory's ground state while the theory's Lagrangian remained gauge symmetric. This structural feature, which is called 'spontaneous symmetry breaking', can be obtained by introducing a scalar field, the Higgs field, with a specific potential. After the Higgs mechanism had emerged as the only promising mechanism capable of reconciling gauge field theory with the observed mass spectrum, the striking conceptual and empirical success of the standard model established increasing trust in the viability of the Higgs mechanism. Once all other predictions of the standard model had been empirically confirmed by the mid 1990s, the Higgs particle remained the only standard model particle to be searched for. The LHC at CERN was build with the aim to test the entire parameter space where the Higgs particle could be expected to be found (apart from a few patches that would constitute particularly unfortunate finetuning).

Higgs particles can be produced in deep inelastic scattering processes: particles are smashed together with very high kinetic energy so that, according the laws of relativity and quantum physics, new kinds of particles are generated in the collision process. If the Higgs particle exists and particle collisions set free energies above the rest mass energy of the Higgs particle, Higgs particles are generated in accordance with the physical conservation laws. The rate of the Higgs production in collisions at a specific energy depends on the involved coupling constants and the phase space (the space of locations and momenta) of the possible experimental outcomes which involve a Higgs particle. The empirical confirmation of the Higgs particle in collider experiments is a particularly difficult enterprise due to the specific properties of the particle. Because of its high mass, the Higgs particle decays into energetically favoured lighter particles after a very short period of time. This time period is too short for generating an observable trace or gap (the standard model Higgs particle is electrically neutral and therefore could not generate a particle trace) in a detector. Generation and decay of the Higgs particle occur, for all practical purposes, at the same spot in the detector and must be attributed to one and the same vertex in a picture of particle traces extracted from the detector.

This fact makes it very difficult to identify Higgs particles on pictures of scattering events. We can understand this by comparing the situation of the Higgs particle with a scenario where a generated new particle lives long enough for leaving a trace in the detector before decay. A particle trace contains a lot of information that can be extracted from a picture of an individual scattering process. To begin with, a trace univocally indicates a charged particle. Furthermore, a relation between the particle's charge, mass and kinetic energy can be extracted from the way the trace is curved by the magnetic field in the detector. Finally, some information on the particle's lifetime can be gained from observing whether, and if so where, the trace ends in the detector. All that information in conjunction with the theoretical knowledge about the observed scattering process can lead to a univocal identification of an individual trace with a specific type of elementary particle. The observation of a mere vertex on a scattering picture does not provide a comparable amount of information about the particles which are generated and decay within it. Thus it is often possible to interpret an individual vertex on a scattering picture in several different ways, each of them involving the generations and decays of different particles. The higher the scattering energy is, the more types of particles can be produced that are capable of generating the observed effective vertex, which makes the situation increasingly complex. In the case of the Higgs particle, which is very heavy and therefore can only be produced with very high scattering energies, a scattering picture that can be interpreted as containing a Higgs particle always allows for a number of other interpretations which do not involve a Higgs particle. The generation and decay of a Higgs particle therefore can never be univocally attributed to an individual vertex in a scattering picture.

For that reason, the existence of the Higgs particle must be demonstrated on a statistical basis. The rates of specific types of events that contain vertices which might arise due to Higgs production must first be calculated based on the known and well established theories of particle physics without assuming the existence of the Higgs particle. Then, it must be checked whether the measured rate of such events lies significantly above the calculated event rate. If that is the case, one finally has to check whether the observed excess of the given events is consistent with the predictions of one of the known Higgs models (that is, a specific realisation of spontaneous symmetry breaking based on a Higgs sector). The empirical analysis thus consists of two separate parts. First, it is checked whether the data is compatible with the dynamics predicted by standard model physics without the inclusion of a Higgs particle. This corresponds to testing the null hypothesis in the given case. If the null hypothesis is excluded, physicists are justified to claim that they have observed new physics. Second, it is analysed whether this new physics, based on our theoretical knowledge, is univocally consistent with the Higgs hypothesis. If that is established, the collected data can be acknowledged as conclusive empirical evidence for a Higgs particle.

3: The Look Elsewhere Effect

We now want to take a closer look at the first step of the data analysis described above. It deals with the question whether or not the collected data implies new physics. Let us assume that the detector measures a certain number of events which could stem from the generation and decay of Higgs particles. As discussed above, physicists always face a background of events which can explain the observed signature but are caused by other kinds of processes that do not involve a Higgs particle. Since physicists know the standard model well and can calculate its scattering amplitudes, they can calculate the expected background. However, since they are dealing with a stochastic quantum process, statistical fluctuations can produce events in excess of the numbers to be expected. Faced with a specific measured number of events of a given type, experimentalists thus must calculate the probability that a number of events at least as high as the measured number was produced based on the null hypothesis that amounts to assuming standard model physics without a Higgs particle. That probability is called the p-value of the data with respect to the given null-hypothesis. If the p-value lies below a certain limit, the null-hypothesis can be taken to be refuted and experimentalists are justified to speak of the discovery of new physics. Particle physicists have set limit for announcing a discovery at a 5σ confidence level, which – expressed in terms of the Higgs search - corresponds to the condition that the probability of finding an excess rate of at least the measured size that is caused by standard model particles without a Higgs particle is lower than $3x10^{-7}$.

Setting a discovery limit is, of course, a matter of convention. It is based on a trade-off between the advantage of calling a viable scientific claim empirically well-established and the potential damage of endorsing a false scientific claim. In many scientific fields, a 3σ effect, that is a probability of less that 0,15% that the measured excess rate of events could be reached if the null hypothesis were true, is considered sufficient for establishing a phenomenon. In particle physics, the discovery of a new particle is taken to be of very high importance and is used in the analysis of the background in all future high energy scattering experiments. Therefore the risk of erroneously acknowledging a discovery of a particle should be kept particularly low and a stronger criterion seems advisable.

Still, a 5σ limit for discovery might seem surprisingly high at first glance. One important reason for this remarkably high limit has to do with a characteristic aspect of experimentation in high energy physics that is called the look elsewhere effect (LEE). Normally, experiments in high energy physics don't just search for new phenomena at one specific energy scale. Rather, they test a wide energy spectrum. The chances that an experiment shows a certain deviation from the predicted event rate at some energy scale within the tested energy range thus must be calculated by summing up the chances of finding it at each specific energy scale. The number of 'places' where one can find a signal is roughly given by the tested energy range over the width of the signal.

To give a specific example, let us imagine that energy levels of specific events can be specified with the accuracy of 1 GeV and a range of 100 GeV is tested. The probability of finding a 5σ deviation from the predicted event rates somewhere within the tested energy range then is roughly 100 times the probability of finding it at one specific energy level. That is, the probability of getting a 5σ effect due to oscillations anywhere within the measured energy range is not $3x10^{-7}$ but rather $3x10^{-5}$. The latter value may be taken to be a reasonable limit for acknowledging a discovery of a new particle once one takes into consideration that 1) in the face of a considerable number of experimental tests of various kinds of new physics one wants to have a small probability that any announcement of a discovery is spurious and 2) it seems wise to introduce some extra error margin in order to account for unknown systematic errors which might distort the empirical data.

The look elsewhere effect thus is one main reason for setting the limit for acknowledging a discovery as high as 5σ. Historically, the 5σ limit was established based on largely pragmatic considerations. While statistical fluctuations at a 4σ level did and do occur from time to time in high energy physics experiments, no 5σ signal in a particle experiment has up to this point ever turned out to be a fluctuation. A 5σ limit therefore seemed plausible simply based on the historical record.[3] The fact that 4σ fluctuations do occur can be statistically explained based on the number of experiments that are carried out in conjunction with the size of the look elsewhere effects which usually applies in those contexts.

Let us now look specifically at the Higgs detection at the LHC. Higgs particles can be found at the LHC in a number of different scattering processes. Two types of process are of particular importance because they allow for a rather precise specification of the Higgs mass and have a comparably small background in conjunction with a sufficiently high Higgs production rate. In the first type of process, a Higgs is produced and decays, via a few intermediate steps, into a photon-anti-photon pair. In the second type, the Higgs ends up decaying into four leptons. Both kinds of Higgs candidate events, the photon-anti-photon event as well as the four lepton event, can be detected at the ATLAS as well as at the CMS detector. When CERN announced its results in December 2011, excess rates of both events had been measured by both detectors. The significance of the entire excess rate of both event types over both detectors was assessed to be somewhere close to 4σ.[4] This corresponded to a probability close to $3x10^{-5}$ that the observed rates would arise as a fluctuation of standard model physics without a Higgs particle. Following a conservative assessment of the energy range where a Higgs-like particle could be detected in the experiment, a Higgs-like particle is looked for in about 80 energy bins at the LHC. Taking the resulting LEE into account, one gets a probability of about $2{,}5x10^{-3}$ that an oscillation of the size of the December data or higher could occur at the LHC experiment within the range of sensitivity for the discovery of Higgs-like particles. This probability does not even amount to a 3σ effect and therefore was clearly insufficient for establishing the existence of a new particle. The data of July 2012 then had a significance above the 5σ level. This meant that the probability of reaching the measured excess rate without the exchange of new particles was less than $3x10^{-5}$ even after taking LEE into account, which was sufficient for declaring the data a discovery of a new particle.

4: Two Ways of Interpreting the Data

The problem to be discussed in this paper can be seen clearly when looking at the situation between the first CERN announcement of a potential Higgs finding in December 2011, and the announcement of a discovery of a scalar particle in July 2012. The official CERN announcement of December 2011 strictly adhered to the 5σ rule and called the data an indication of a possible Higgs particle that did not constitute a discovery. Some

---

[3] The problem of the assessment of the statistical significance of data in a field where lots experiments were searching for new particles in lots of places was appreciated by physicists since the early days of particle physics in the 1940s. In the 1960s and 1970s, more specific formulations of LEE were given (see Rosenfeld 1974) and attempts were made to tackle the problem by looking at computer simulations of background fluctuations. (See e.g. Lederman et. al. 1976 and a comment on page 267 of Alvarez 1968.) Only later it became customary to use a rigid 5σ discovery criterion instead of simulating LEE in the individual cases. For a recent take on the 5σ level, see Feldman 2006. (I am very grateful to Robert Cousins for providing most of the information contained in this footnote.)

[4] The calculation of the overall significance of data collected in several different experiments is difficult and suffers from ambiguities. (See Cousins 2007 for an analysis and further references.) It was particularly tricky in the given case due to the fact that the characteristic energies of the measured excess rates did not precisely coincide. We need not go into the details of this analysis, however. For the present purposes it is sufficient to have an approximate value for the overall significance.

commentators emphasised the preliminary character of the measured effect by pointing out that 4σ effects had turned out to be statistical fluctuations on several occasions in the past. The understanding that lay behind the described kind of reasoning shall be called the 'experimentalist position' for reasons which will become clear later on.

Despite the official caveat, however, many particle physicists felt pretty sure already after the announcement of the December data that actual Higgs events had been detected. Some of them did present an argument for their enhanced trust in the significance of the data. They argued that the Higgs case differed from those contexts where 4σ effects had vanished after further experimental testing in the past. In the latter cases, a strong LEE had to be taken into account. In the Higgs case, to the contrary, one could considerably reduce the LEE because i) one could be quite confident for theoretical reasons that the Higgs existed and ii) empirically based arguments had already predicted within which energy range the Higgs particle should be found. Physicists who took up that position implicitly suggested that one should look carefully at the theoretical arguments which could constrain the LEE and, on that basis, determine the trustworthiness of the data beyond the rigid application of the 5σ rule. I want to call this position the 'theoretician's position' as opposed to the 'experimentalist position' presented before.

In order to fully understand the difference between the 'experimentalist' and the 'theoretician's' position, we have to remember the two steps of the data analysis mentioned in Section 2. The first step establishes that new physics has been found by demonstrating that the collected data cannot be accounted for by the set of empirically confirmed particles alone. The second step then must demonstrate that the observed new physics can indeed be identified with the Higgs particle. The close to 4σ effect measured up to December 2011 characterises the analysis at step one. It can be specified without any knowledge about the Higgs particle. Keeping this part of the analysis free from reasoning based on the Higgs hypothesis in fact is based on a core principle of the experimental method: experimentalists want to keep the data analysis as independent as possible from the scientific concepts the data is supposed to confirm. If one wants to establish that some data is incompatible with the well-established physics WITHOUT the Higgs particle, it would feel like begging the question to demonstrate that fact by relying on theoretical reasoning that assumed the Higgs particle's existence.[5] Still, that is what the 'theoretician's position' proposes. According to the 'theoretician's position' it would be misleading to pretend that we did not have theoretical knowledge that makes us expect the existence of the Higgs particle. By discarding theory-based indications to that end, we may be experimentally flawless but do not provide a realistic picture of the way the collected data actually influences our trust in the Higgs particle.

The designations 'experimentalist' and 'theoretician's position' are not meant to imply that all or most experimentalists adopt the former and most theoreticians the latter position. Rather, it points at the positions' core concerns. The experimentalist position has

---

[5] The accusation of ‚begging the question‘ must be made more precise in the given context. Formally, using an assumption in a line of reasoning that aims at supporting that assumption is only invalid if one takes that assumption to be certainly true in advance. If one just attributes a certain probability to its truth in advance, an argument on its basis can be perfectly valid, as formally demonstrated in Bayesian analysis. Since no-one pretended to be absolutely sure about the existence of the Higgs particle before its discovery, the claim that the theoretician's perspective amounts to begging the question clearly cannot imply that its line of reasoning is formally invalid. Rather, it suggests that its line of reasoning is "dialectically ineffective" (I am grateful to Luca Moretti for emphasising this point and providing the term.): to those who do not share the trust in the Higgs particle, the reduction of the look elsewhere effect is not plausible and the argument based on it cannot be convincing.

the priority to defend the purity of the process of data analysis by keeping it free of theoretical reasoning that is about to be tested by that very data. The theoretician's position, to the contrary, has the priority to be frank about our actual beliefs with respect to the hypothesis in question. That belief, however, clearly does rely to a considerable extent on the theoretical knowledge we have about the overall situation.

Before assessing the strengths and weaknesses of the two positions, we have to specify a little more clearly what the theoretical support for the Higgs data amounted to. This support was based on the understanding that a Higgs particle was likely to exist. When the LHC search for the Higgs particle began, the standard model was an empirically well confirmed theory except for the Higgs particle. Based on that fact, a seemingly cogent line of reasoning generated trust in the existence of a Higgs particle.

> 1: It was considered highly unlikely that an empirically equally successful description of the physics described by the standard model could be formulated that was entirely independent from the principles of gauge field theory that stood behind the standard model.
>
> 2: In order to make gauge field theory compatible with the data, it seemed inevitable to introduce a concept of spontaneous symmetry breaking.
>
> 3: The only satisfactory way of introducing spontaneous symmetry breaking (disregarding the far more difficult and conceptually in some respects doubtful approach of 'Higgsless models' which were discussed in recent years) seemed to be the introduction of a Higgs scalar – which could be an elementary particle or a boundstate.

On the described basis, high energy physicists felt highly confident about the existence of a Higgs scalar even in the absence of direct empirical evidence for it.

This trust in the existence of the Higgs particle provided the basis for a second step of reasoning. The Higgs field can have empirical effects in two ways. On the one hand, a Higgs particle can be produced in collisions as a real particle if the collision energy is high enough for generating a particle of its mass. This is the effect searched for at the LHC experiments. In addition, however, the Higgs field also has an effect on perturbative calculations of processes which do not generate real Higgs particles. Due to the uncertainty principle of quantum physics, calculations of perturbative corrections in quantum field theory must take into account contributions of 'virtual particles', i.e. particles that could not be produced in accordance with the laws of energy and momentum conservation. These virtual Higgs contributions to perturbative corrections play a role already at energy levels that are too low for generating real Higgs particles. The size of those effects of the Higgs field on scattering cross sections can be calculated. Calculations of scattering amplitudes which take into account virtual contributions of a Higgs particle of a given mass then can be compared with actual measurements of the corresponding scattering processes at the given energies.

No high energy experiment before the start of the LHC provided any data that could be understood as an effect of virtual Higgs-contributions. This observed lack of Higgs-induced effects constrained the Higgs mass quite strongly already at the time the LHC experiments started. On that basis, one could set an upper bound of 141 GeV for the Higgs mass. In conjunction with the fact that earlier experiments had already directly excluded a Higgs mass of less than 115 GeV, there remained a rather small window of possible Higgs mass values. Therefore, physicists had strong reasons to expect that, if a Higgs particle would be found at all, it should be found between 115 and 141 GeV.

The described theoretical status quo had the potential to influence the evaluation of the empirical data on the Higgs from December 2011. The theoretician's position suggested taking seriously our theoretical knowledge about the Higgs particle. First, it seemed justified to presume on theoretical grounds that a scalar Higgs field was likely to exist. Second, one knew from previous data that, if it existed, it was most likely to have a mass between 115 and 141 GeV. A significant excess of Higgs-like events therefore could only count as a serious candidate for a Higgs discovery if it corresponded to a Higgs mass between 115 and 141 GeV. Frome this perspective, a signal between 115 and 141 GeV had to be related to a very different look elsewhere effect than an imagined signal outside that mass window. The latter had to be treated as a potential signal for unexpected and not yet understood new physics whose appraisal had to account for a strong look elsewhere effect (let us say, in the given case, 80 bins or even more). When looking at a signal between 115 and 141 GeV, however, which was consistent with the entire available knowledge about the Higgs particle, that knowledge justified accounting for a smaller look elsewhere effect that was reduced to the described window for the possible Higgs mass. Given a width of the signal of 1 GeV, this corresponded to a factor of 26, which increased the significance of the data from December 2011 by a factor 3.

Strictly speaking, the argument did not depend on attributing a high absolute probability to the existence of the Higgs particle. It was sufficient to believe that the existence of the Higgs particle was far more likely than new physics that could explain the observed excess rate but was not a Higgs particle. To understand this point, let us assume, for a moment, that we could strictly rule out the existence of any other new physics that could explain an excess rate of the observed kind. In that case, we clearly could reduce the look elsewhere effect to 26 even if we had no confidence in the existence of the Higgs particle whatsoever.

Contrary to the theoretician's position, the experimentalist position suggests that we must not take into account any information about the probability of the Higgs mechanism or any other kind of new physics when assessing the data which is collected to confirm it. On that perspective, theoretical 'prejudices' regarding the existence of an object must not be responsible for announcing the discovery of that object. When disregarding all knowledge about the probability of new physics, however, no justification remained to take the Higgs particle to be more probable than so far unknown kinds of new physics that were not constrained to the mass window between 115 and 141 GeV. In other words, there was no justification to reduce the look elsewhere effect to the Higgs-mass window. The rigid experimentalist position therefore had to account for the full look elsewhere effect implied by the experimental setup. It thus did not generate an argument for moving away from the universal 5σ limit.[6]

From the theoretician's perspective, the look elsewhere effect could be reduced even further based on a more immediate reliance on the assumption of the existence of the Higgs particle. The data collected at the LHC until December 2011 did not merely indicate a signal at 125 GeV. It also excluded with 95% confidence a Higgs mass above 127 GeV. This limit had to be treated differently than the 141 GeV limit discussed above in an important respect. Since the 115-141 GeV mass window was established based on different data than the data

---

[6] Since *any* specification of LEE in a particular experimental context must rely on *some* theoretical considerations based on the theory to be tested, the experimentalist position arguably suggests a general tendency to stick to the default position of a rigid universal discovery condition rather than allow for a conceptual analysis of the individual experimental context.

that provided the measured signal itself, the fact that the measured signal was found within that window was not a priori guaranteed and, on that basis, could justify a different treatment of the actual data than what a potential measured signal outside that window would have required. It was possible to say: since the signal was within the mass window, it constituted a serious Higgs candidate; therefore the look elsewhere effect could be reduced to the energy range compatible with a Higgs observation. In the latter case, this line of reasoning did not work since the data that constituted the signal was also part of the evidence that implied the mass limit at 127 GeV. If the overall data had been different and the measured signal had occurred at a different energy, let us say at 135 GeV, that data obviously would not have implied a mass limit of 127 GeV. Therefore, it was not true in the given case that a measured signal above 127 GeV would have been incompatible with a Higgs particle observation. The line of reasoning supporting the reduction of the look elsewhere effect that was presented in the previous paragraphs thus did not work this time. The look elsewhere effect could be reduced to the range between 115 and 127 GeV nevertheless, however, if one was ready to rely on the conviction that the Higgs particle existed with a high probability. Under this assumption, one could adhere to the following line of reasoning:

1) It is likely that the Higgs exists (for the sake of the argument, let us assume the arbitrarily chosen likelihood of 90%).
2) With a probability of 95%, the Higgs does not have a mass below 115 or above 127 GeV.
3) Therefore, with a probability of 0,90x0,95=0,855 there is a scalar field within that mass window, whether or not any other new physics can be found beyond that window.
4) Assuming that there is a scalar within that mass window, only a signal within that mass window can be identified with it. Therefore, under this condition, it is justified to reduce the look elsewhere effect to the range between 115 and 127 GeV. Given the high trust in the existence of the Higgs particle (and once again assuming a width of the signal of 1 GeV), we thus can reduce the look elsewhere effect to a factor close to 12 (with our chosen likelihood we would have 0,855x12+0,145x26≈14.)

The look elsewhere effect is then reduced by another factor 2. In the given case, the line of reasoning is more directly based on the presumption that the Higgs particle exists. While in the earlier argument it was sufficient to assume that we had good reasons to look for the Higgs particle rather than for something else that could be outside the mass window 115-141 GeV (in other words, that it was much more likely that there was an actual signal at a specific energy within that window than at some specific energy outside), the present argument only works if one is really committed to the existence of the Higgs particle. This second step of reducing the look elsewhere effect thus is ‘theoretically contaminated’ to a higher degree than the first one and therefore is even less acceptable from an ‘experimentalist’ position.

Both steps in conjunction reduced the look elsewhere effect by a factor 6. The significance of the data when taking into account the look elsewhere effect thus would have increased from a less than 3σ effect when seen from the ‘experimentalist’ position to an effect well beyond the 3σ limit.

5: The Significance of the Debate

Why is the discussion between the ‘experimentalist’ and the ‘theoretician’s’ position’ of interest? It is not relevant any more for the assessment of the Higgs data since the evidence collected in 2012 amounted to the discovery of a Higgs particle on any account. In order to understand why the relevance of the discussion goes beyond that of an anecdote during the process of Higgs data collection, we have to ask why a debate about the look elsewhere effect arose at all at this stage of the evolution of high energy physics. One reason lies in the specific nature of the Higgs search. Apart from some more technical aspects of the process of data collection (which were responsible for the fact that the question did not arise in a similar way in the case of the top-quark search), a core reason that is of crucial importance from our perspective has to do with time frames. If the process of data collection is a matter of hours, days or even a few weeks, it does not make much sense to start a ‘philosophical’ discussion about the status of an intermediate result of the data collecting process. Even when privately endorsing a ‘theoretician’s position’, the only scientifically reasonable strategy under such circumstances is to wait for a couple of days or weeks until the full data is in and the question is decided empirically on any account. Playing safe and adhering to the ‘experimentalist’ position thus used to be scientific consensus among all participants. The search for the Higgs particle, due to its technical and conceptual complexity, required longer time scales. The announcement of evidence for the Higgs particle came three years after the experiment had started and it took another seven months until a discovery could be announced. This was a sufficiently large time frame for justifying an official announcement of evidence for the Higgs particle that did not yet amount to a discovery, and, consequently, for triggering the described debate between adherents of the ‘experimentalist’ and the ‘theoretician’s’ position.

Now it is important to understand that the long timelines of the Higgs search are characteristic of the overall evolution of fundamental physics. Supersymmetry, which constitutes a possible structural characteristic of high energy physics that would imply the existence of a wide range of new elementary particles and may be found at the LHC within the next decade, would be far more difficult to establish conclusively than the Higgs particle. It would presumably take several years to get from the first indication of supersymmetry to an announcement of its conclusive confirmation. Empirical confirmation of other, more far-reaching theoretical hypotheses in fundamental physics like cosmic inflation or string theory, to the extent it will be achieved at all, must be expected to require even longer time frames and presumably will be based on complex patterns of cosmological data that are less conclusive than the data extracted from collider physics. Fundamental physics thus enters a stage where the presence of significant but inconclusive evidential support for individual theories will constitute the status quo for many years or even decades. The discussion about the status of the Higgs data in this light may be seen as a test case for a far more urgent debate that is due to arise in the foreseeable future. The question will be to what extent it is legitimate to move away from the canonical ‘experimentalist’ position in cases where physics 1) can rely on a strong and cogent theoretical analysis of the overall physical context and 2) must specify the status of its theories based on inconclusive data for periods of time which may approach the length of an individual scientific career. To be sure, not all aspects of the discussion of the Higgs discovery case can be applied directly to other contexts in high energy physics and cosmology. Nevertheless, the Higgs case constitutes a good starting point for the analysis of the described questions. It is on that basis that the debate on the Higgs evidence merits closer philosophical inspection.

6: Bayesianism versus Frequentism

The debate between the experimentalist and the theoretician's position can be understood in terms of a conflict between a rigidly frequentist and a partially Bayesian perspective. Let us first briefly characterize a Bayesian and a frequentist perspective in their pure forms. The Bayesian approach aims at extracting probabilities of the truth of a scientific hypothesis H in the face of empirical evidence E based on the probabilistic relation

$$P(T|E) = P(E|T)\frac{P(T)}{P(E)}$$

where T denotes the statement that hypothesis H is true. Empirical evidence E is taken to confirm H iff it increases the probability that H is true, that is iff

$$P(T|E) > P(T)$$

The prior probability P(T) can be informed by old empirical data, theoretical considerations or prejudice. The posterior extracted from considering one set of data E can serve as a prior for the next step of empirical testing. To get started with the empirical process, however, the scientist has to assume a prior that is not based on empirical tests of the given hypothesis. For that reason, the Bayesian perspective always contains a subjective element.

From a Bayesian perspective, the scientific process can nevertheless be seen as inter-subjectively reliable. This conclusion is based on a specific property of Bayesian theory confirmation: repeated consideration of new empirical data leads to converging posterior probabilities. Starting from very different prior probabilities, after a sufficiently extensive series of empirical tests one ends up with very similar posteriors.

The frequentist approach has a different focus than the Bayesian. While Bayesianism aims at modelling a full and coherent argumentative structure that starts from the assessment of the probability of prior assumptions and ends with the assessment of the probability of a hypothesis in the face of all known data, the frequentist approach focuses on isolating the statistical analysis of numerically well-specified empirical data from vague assessments of priors. The frequentist carries out a statistical analysis of the empirical data within a given framework that is taken for granted without relying on any information about the prior likelihood of the hypothesis to be tested. On that basis, precise statistical analysis can, within a given theoretical framework, determine the probability that empirical data that belongs to a certain class *E* could have arisen if a certain scientific theory were true.

When testing a null hypothesis in a collider experiment, this class *E* of empirical outcomes is specified as the class of outcomes which correspond to event excess rates (compared to the value expected based on the null-hypothesis) that are at least as high as the one actually measured in the experiment. We can then define E as the observation that the collected empirical data belongs to the class *E*, so that $P(E|T_N)$ corresponds to the p-value of the collected data. (T_N denotes the statement that the null-hypothesis is true.) In order to arrive at $P(T_N|E)$, the factor $P(T_N)/P(E)$ would have to be taken into account.

Specifying that factor would require vague and subjective assessments of priors, however, from which the frequentist prefers to abstain. Therefore, she remains content with specifying $P(E|T_N)$ and uses this value as an indicator for the viability of hypothesis H (she calls it the likelihood of H). This is justified, from a Bayesian perspective, by the abovementioned convergence behaviour of $P(T|E)$ under repeated empirical testing. The frequentist relies on the understanding that the impact of priors is eventually 'washed out' by empirical data.

Applied to the Higgs search at the LHC experiments, the frequentist analysis takes for granted the empirically well confirmed parts of the standard model of particle physics plus a wide range of scientific concepts in collider physics as the basis for analysing the data. It does not, however, admit any information that relies on the assessment that the Higgs particle is likely to exist. The frequentist perspective therefore resembles the canonical 'experimentalist' position.

A Bayesian perspective on the Higgs search, to the contrary, would acknowledge that the analysis of specific experiments at the LHC may be based on informed priors for the existence of the Higgs particle. These priors are themselves based on previously collected empirical data and theoretical reasoning relying on that data. Since a long and elaborate scientific process has led up to the emergence of those priors, they must not be understood as entirely subjective, even though the specific values attributed to them by individual scientists may differ to some degree based on subjectively chosen presumptions. In other words, the scientists' trust in the existence of the Higgs particle before the start of LHC experiments can be understood as the result of a scientifically informed assessment rather than of mere subjective prejudice.

A full realization of a Bayesian perspective on the Higgs search would need to extract $P(T_N|E)$ from $P(E|T_N)$ by specifying the prior $P(T_N)$ and $P(E)$. In order to give a quantitative example, let us imagine that physicists collect empirical evidence at the LHC that implies $P(E|T_N)=0{,}01$. The Bayesian now takes into account that we strongly believed in the Higgs particle already before the experiment, and, let us say, attributed a probability of 90% to its existence. Therefore, we excluded the null hypothesis with at least 90% probability from the start. Assuming the simplified scenario that the Higgs hypothesis is the only plausible alternative to the null-hypothesis, we write $P(T_N)\approx 0{,}1$. Given that we strongly expect E if the Higgs exists, we have $P(E)\approx 0{,}9$, which gives $P(T_N|E)=P(E|T_N)P(T_N)/P(E)\approx 0{,}01x(1/9)\approx 0{,}0011$. (For the sake of simplicity, we have ignored LEE at this point. It will enter the picture below.) A Bayesian approach thus would generate a considerably higher degree of trust in the Higgs particle than the frequentist statistical analysis of the numerical data.

Though Bayesian perspectives are sometimes thought about in experimental high energy physics, no physicist would propose to replace the frequentist statistical analysis by a consistently Bayesian line of reasoning. The motives for that restraint are very clear. By establishing the Bayesian analysis of the research process as a fully viable strategy of scientific data analysis, one would permit that a rigid quantitative statistical analysis where the numerical input is well-determined by the empirical data gets adulterated by probability assessments which are vague and subjective. It would put guessing priors on the same footing as rigid and quantitative experimental testing.

Fortunately, there is also a less intrusive way of introducing aspects of Bayesian reasoning into data analysis. This alternative path avoids messing with the statistical analysis itself and rather addresses the question how to interpret the statistical results. It was emphasised already above that the interpretation of the statistical analysis in terms of ‘significant evidence’ and ‘discovery’ is not as rigid and univocal as the statistical analysis of empirical data itself. It relies on the subjective decision of setting specific significance limits. The decision where to put these limits is driven by striking a balance between playing safe and allowing for clear scientific statements. One cannot avoid this subjective element in the interpretation of the data: setting discovery limits somewhere is necessary in order to be able to treat scientific statements as stable elements of a scientific world view. Without defining any limits, physicists could never claim to have discovered a new phenomenon in microphysics and could not define a univocal conceptual basis for the null hypothesis in future high energy experiments.

In the case of the search for the Higgs particle, a Bayesian perspective that addresses the question of interpretation must address the role of LEE. As discussed in Section 3, LEE plays a substantial role in specifying the limits for calling data a discovery of new physics. Moreover, it was shown that an understanding of the size of LEE can be based on prior assessments of the probability that the Higgs particle exists. Taking those assessments into account leads to the analysis that was associated with the theoretician’s perspective in Section 4.

We can account for the LEE in the following way: we define a probability $P(E|T_N)_{LEE}$ that takes into account LEE as

$$P(E|T_N)_{LEE} = P(E|T_N)(L_{T_H}P(T_H) + L_{T_{\sim H}}P(T_{\sim H})$$

$L_{T_H}$and $L_{T_{\sim H}}$denote the LEE factors under the assumption that the Higgs hypothesis is true and under the assumption that it is false, respectively. Of course, $P(T_{\sim H}) \geq P(T_N)$ must be satisfied. Strong trust in the truth of hypothesis H (i.e. high $P(T_H)$) in conjunction with a factor L_{T_H}that is reduced compared to the factor $L_{T_{\sim H}}$ implies a reduced value of $P(E|T_N)_{LEE}$ and therefore a higher overall significance of the data.

The theoretician’s perspective thus amounts to a partial inclusion of Bayesian reasoning by introducing $P(T_H)$ into the assessment of the data. The specific strategy of introducing Bayesian priors solely via influencing the quantitative specification of limits with respect to the significance levels does not reach all the way to a determination of P(T|E), however. Bayesian reasoning is confined to the assessment of LEE and introduces vague and subjective prior probabilities only within the interpretational part of data analysis that is vague and subjective anyway. Therefore, the Bayesian element of human reasoning is accounted for without compromising the methodological purity of scientific data analysis in its core regime. This approach seems less repulsive to many physicists than an all-out Bayesian treatment of statistical analysis and, as described above, has been acknowledged by some as a plausible way of reasoning when analysing the significance of the Higgs data.

In the debate on the December 2011 data, considerations along that line were made in a fairly informal way in order to justify one’s own subjective trust in the data. No-one proposed using the described ‘weak’ deployment of Bayesianism ‘officially’ for the specification of a LEE-dependent discovery condition that could replace the rigid 5σ criterion. In principle, however, an altered context of experimentation where the conceptual background is deemed very trustworthy and the time scales for the collection of conclusive empirical data are very high, might suggest this step as a reasonable way to go.

When looking closer at the characteristics of empirical testing in high energy physics, it is actually possible to go beyond the modest statement that the described elements of Bayesian reasoning seem acceptable. In fact, one finds quite strong arguments against the plausibility of a dogmatic adherence to frequentist principles. In specific contexts, the latter fails to be consistent with elementary scientific intuitions about theory confirmation. In the following section, an example of this kind shall be discussed.

7: Assessing the Merits and Problems of the Two Perspectives

As was pointed out above, the plausibility of the 5σ limit directly depends on the LEE that happens to be relevant in today's high energy experiments. If a significantly higher LEE became typical for experimentation in high energy physics, statistical fluctuations of 5σ significance would start showing up in experimentation. It may be expected that scientists would then be led to introduce a higher limit for the discovery of a particle in order to maintain the trustworthiness of experimental results. To give a specific example, if a new generation of experiments looked for new particles in 10 000 bins, the chances of finding a 5σ fluctuation in such an experiment would be about $3x10^{-3}$. It would clearly be premature in this context to call a 5σ signal a discovery of a corresponding particle. A 6σ limit would be necessary to retain the old level of trustworthiness of experimentation in the new scientific context.

Let us now imagine a theory H_2 that predicts the existence of a scalar particle h with a mass within a parameter interval I_1 and forbids the existence of a scalar particle with a mass within the 100 times larger interval I_2. In order to test H_2, physicists first build an experiment that scans interval I_1 with 100 bins. They succeed in finding a scalar with 5σ confidence level, which, as experimentation still proceeds within the 'old' context of testing smaller energy ranges and therefore of a smaller LEE, constitutes the criterion for calling the data a discovery at the time. Later on, experimentalists develop new techniques which allow them to test far wider energy ranges. On that basis, they build a larger experiment in order to test the second prediction of H_2. In this experiment, they scan the interval I_2 with 10 000 bins (without providing further tests of I_1) and find no further scalar, which is in perfect agreement with H_2. Now physicists proceed to carry out an overall analysis of the entire data collected in both experiments. Since experimentation has entered a stage of testing wider energy ranges, it would be necessary to correct the limit for discovery and shift it, let us say, to 6σ. The data that indicates a scalar field in interval I_1 does not amount to a 6σ effect, however, which means that physicists are not allowed to speak of a discovery of the scalar h anymore. The confirmation of theory H_2 by the observation that no scalar exists in I_2 thus has, in effect, invalidated the discovery of the phenomenon h that is predicted by H_2. An implication of this kind clearly is at variance with our intuitive understanding of confirmation and discovery.

The theoretician's perspective avoids the above implication because the trust in theory H_2 blocks the look elsewhere effect with respect to I_2. The theoretician's perspective implies the following line of reasoning: the experimental testing of I_1 has led to the discovery of scalar h and thereby has established the viability of H_2 with high probability; since H_2 predicts that no scalar particles within the mass range I_2 exist, strong trust in H_2 implies that I_2 has little relevance for LEE with respect to the search for scalar particles; the experimental testing of I_2 therefore does not significantly change the discovery condition for h and cannot invalidate the discovery of h. By avoiding the paradoxical conclusion suggested by the experimentalist perspective, the theoretician's perspective looks decidedly more plausible than the former in the given context.

However, complications can arise with respect to the theoretician's perspective as well. To see this, let us now imagine that theory H_2 has not been developed and scientists right away build an experiment that can test the entire region I_1+I_2 without having any predictions regarding its particle content. Let us further assume that experimentalists run the experiment and measure a 5σ effect in I_1. When assessing the data, scientists must, once again, take into consideration the large LEE that comes into play in the new context of experimental testing and introduce a correspondingly strong discovery criterion of 6σ. This would imply that the data collected is strong evidence for h but does not constitute a discovery of h. Now imagine that some physicist develops H_2 in order to explain the data. Should scientists be ready to acknowledge that the mere creation of theory H_2 implies the discovery of h by providing the basis for reducing the look elsewhere effect along the lines laid out before? It is quite clear that they should not. It would seem absurd to claim that the capability of scientists to build a theory in agreement with a signal that might eventually turn out to be a fluctuation should influence the statistical treatment of the corresponding data. After all, scientists may well find scientific explanations of fake phenomena if they try hard.

The intuitive understanding of the notion of discovery thus suggests two seemingly contradictory conditions for treating LEE. First, LEE should be taken to depend on our knowledge of the theoretical background in order to avoid the withdrawal of a discovery claim based on evidence that confirms the corresponding theory. But second, the development of a theory based on given data should not *ex post* imply the discovery of new physics based on that same data. The first requirement introduces the very same Bayesian elements of reasoning that are rejected in a strikingly similar context by the second requirement.

The tension between the two conditions can be at least partly dissolved by accounting for the chronology of the research process. In the first scenario, theory H_2 predicts the scalar h before indications for it are observed. Moreover, theorists take H_2 to be their best guess. That is, either they have no alternative theory at all or they have reasons to prefer H_2 to the alternatives they see. The fact that scientists take H_2 to be their best guess implies that they have a good reason for testing region I_1 rather than some other region. The empirical confirmation of H_2 by the discovery of h and then again by observing the absence of scalar particles in the regime I_2 further increases the trust in the viability of H_2 and thereby provides a justification for constraining LEE to I_1 even after the extension of the tested regime. In the second scenario, to the contrary, H_2 is developed in order to account for the data E. Beyond the fact that H_2 is in agreement with the data that was the basis for its construction, there is no particular reason for believing in H_2. The core reason for trusting H_2, namely the nontrivial observation that H_2 was able to predict h in the correct regime without being motivated by a prior observation of h, does not apply. Scientists might well have developed a different theory than H_2 if they had tried to explain different data than the one actually collected. Therefore, it does not seem justified to assume the validity of H_2 ex ante in any kind of reasoning that assesses the significance of data collected in order to test H_2. This means that the full LEE has to be taken into account.

Comparing the two analyzed scenarios reveals the importance of the classical distinction between novel confirmation and accommodation. It has been a longstanding matter of debate in philosophy of science whether or not confirming data that did not influence the process of theory construction has a higher confirmation value for the confirmed theory than data that did influence the construction of the theory. In a Bayesian context, initial claims that the explanatory extra value of novel confirmation was at variance with Bayesianism were refuted by a number of works that presented Bayesian models which supported the novel confirmation hypothesis [Maher 1988, Kahn et al. 1992, Barnes 2008]. In our context, the assumption of a difference between accommodation and novel confirmation is crucial for distinguishing between the two stated scenarios. The first scenario, where the range of testing is extended from I_1 to $I_1 + I_2$, largely derives the justification for reducing LEE from novel confirmation, that is from the fact that the theory H_2 was developed without being informed by data E and was empirically confirmed by E afterwards. In the second scenario, we encounter a case of accommodation where no justification of that kind can be extracted. Therefore, LEE must remain large and no discovery can be established. We thus have the interesting case that a specific form of scientific reasoning implicitly requires the assumption of an extra value of novel confirmation for being consistent with the basic scientific intuition about theory confirmation.[7]

Acknowledgements: I am grateful to Robert Cousins for very helpful information and to Luca Moretti for important comments. This work was supported by the Austrian Research Fund (FWF) grant P22811-G17.

---

[7] In other scenarios, the situation is more complex and the intuitions are less clear than in the two cases discussed. (See e.g. the case where I_1 was tested first but H_2 was developed only after the tests of I_1.) An exhaustive and general analysis of the described class of examples must be deferred to future investigations.

References:

Alvarez, L. W. 1968: “Recent developments in particle physics”, Nobel Lecture.

ATLAS Collaboration 2012: ”Observation of a new particle in the search for the standard model Higgs boson with the ATLAS detector at the LHC”, arXiv:1207.7214.

Barnes, E. C. 2008: *The Paradox of Predictivism*, Cambridge University Press.

CMS Collaboration 2012: “Search for the standard model Higgs boson produced in association with W and Z bosons in pp collisions at sqrt(s) = 7 TeV”, arXive:12093937.

Cousins, R. D. 2007: “Annotated bibliography of some papers on combining significances or p-values”, arXiv:07052209v2.

Dawid, R., Hartmann, S. and Sprenger, J. 2013: “The No Alternatives Argument”, The British Journal of Philosophy of Science, forthcoming.

Feldman, G. 2006: Proceedings PHYSSTAT 2005, concluding talk.

Kahn, J. A., Landsberg, S. E. and Stockman, A. C. 1992: ‚On Novel Confirmation‘, British Journal for the Philosophy of Science 43, 503-516.

Lederman et. al. 1976: “Observation of high-mass dilepton pairs in hadron collisions at 400 GeV”, Physical Review Letters 21, 1236-39.

Maher, P 1988: ‘Prediction, Accommodation and the Logic of Discovery’, PSA 1988, 273-285.

Rosenfeld, A. 1974: “The particle data group: growth and operations eighteen years of particle physics”, Annual Review of Nuclear and Particle Science 25.