Testing Inflation

Chris Smeenk*

1 Introduction

Over the last thirty years, inflationary cosmology has been the dominant theoretical framework for the study of the very early universe. The central idea of inflation is that the universe passed through an impressive growth spurt, a transient phase of quasi-exponential ("inflationary") expansion which sets the stage for subsequent evolution described by the standard big bang model of cosmology. This inflationary phase leaves an imprint on various observable features of the universe. Observations can then constrain the fundamental physics driving inflation, typically described in terms of an "inflaton" field. Traces of an inflationary stage left in the form of temperature variations and polarization of the cosmic microwave background radiation (CMB) are particularly revealing regarding the inflationary phase. There are currently many models of inflation compatible with the available data, including the precise data sets generated by assiduous observations of the CMB. Yet there are ongoing debates regarding how strongly this data supports inflation. Critics of inflation argue, among other things, that its compatibility with the data reflects little more than the enormous flexibility of inflationary model-building. These concerns have become particularly pressing in light of the widespread acceptance of eternal inflation, which seems to imply that all possible observations are realized somewhere in a vast multiverse.

Whether inflation can be empirically justified — whether it is "falsifiable" — is a leitmotif of these debates. There has been little agreement among cosmologists about how to define falsifiability, and whether it demarcates science from the rest as Popper intended.[1] The question at issue is how to characterize a theory's empirical success, and to what extent success, so characterized, justifies accepting a theory's claims, including those that extend far beyond its evidential basis. Success defined as merely making correct predictions, merely "saving the phenomena," does not provide sufficient justification, for familiar reasons. False theories can make correct predictions, and predictive success alone is not sufficient to distinguish among rival theories that happen to agree in domains we have access to. Facile arguments along these lines do not identify legitimate limits on the scope of scientific knowledge; instead, they indicate the need for a more careful analysis of how evidence supports theory. Philosophers have long acknowledged this need, and physicists have historically demanded much more of their theories than mere predictive success. Below I will focus on two historical cases exemplifying strong evidential support. The strategies illustrated in these cases generalize, and inspire an account of how theory and data should be related for a theory to meet a higher standard of empirical success. A theory that is successful by this standard arguably makes a stable contribution to our understanding of nature, in the sense that it will be recovered as a valid approximation within a restricted domain according to any subsequent theory.

Both strategies focus on mitigating the risk associated with accepting a theory. In the initial stages of inquiry, a theory is often accepted based on its promise for extending our epistemic reach. Theories allow

*Rotman Institute of Philosophy, Department of Philosophy, Western University. Email: csmeenk2@uwo.ca.

[1]The question of inflation's falsifiability is discussed by several of the contributors to Turok (1997); Barrow and Liddle (1997) argue that inflation can be falsified. For recent discussions from opposing viewpoints see, for example, Ellis and Silk (2014) and Carroll (2014). These debates use Sir Karl's terminology but do not engage in detail with his views about scientific method.

us to use relatively accessible data to answer questions about some other domain; they provide an epistemic handle on entities or phenomena that are otherwise beyond our grasp. Inflationary cosmology allows us to gain access to the very early universe, and high energy physics, in just this sense: if inflation occurred, then observable features of the CMB reflect the dynamical evolution of the inflaton in the very early universe. Using theory to gain access to unobserverable phenomena poses an obvious risk. The theory provides the connections between data and the target phenomena, and the data provide relevant evidence when interpreted in light of the theory. How does one avoid accepting a just-so story, in the form of an incorrect theory that fits the data? Demanding strong evidence at the outset of inquiry would be counter-productive, because the best evidence is typically developed through a period of theory-guided exploration. The detailed quantitative assessment of a theory is a long-term achievement. The discussion of historical cases in §2 illustrates how a theory can be tightly constrained by independent measurements, and subjected to ongoing tests as a research program develops.

These considerations suggest reformulating debates regarding the falsifiability of inflation with an assessment of two questions (§4). To what extent do observations of the early universe provide multiple, independent constraints on the physics underlying inflation? And has inflation made it possible to identify new physical features of the early universe that can be checked independently? Focusing on these questions allows for a clearer assessment of the challenges faced by cosmologists in developing evidence of comparable strength to that in other areas of physics, going beyond compatibility of inflationary models with available observations. I will argue that the main challenge to the program of reconstructing the inflaton field is a lack of independent lines of evidence. But if inflation is generically eternal, I will briefly argue that the challenges are insurmountable: eternality undermines the evidence taken to support inflation, and blocks any possibility of making a stronger empirical case.

2 The Determination of Theory by Evidence

Assessment of the degree of evidential support for theories, drawing distinctions among theories that all "save the phenomena," has long been a focus of epistemological discussions in physics. On one extreme, some theories are merely compatible with the phenomena, in that their success may reflect ingenuity and flexibility rather than accuracy. Although models constructed by fitting the data can be useful for a variety of purposes, they are not regarded as revealing regularities that can be reliably projected to other cases. On the other extreme, the new laws and fundamental quantities introduced by a theory are as tightly tied down by the evidence as Gulliver by the Lilliputian's ropes. Even though such theories make claims about the structure of the natural world that go far beyond the data used to support them, physicists take them to accurately capture the relevant quantities and law-like relations among them, which can then be projected to other cases and used as the starting point for further work. In numerous historical cases physicists regard a given body of evidence as strong enough to determine the correct theory.[2]

Judgments of the strength of evidence are as difficult to analyze as they are central to the practice of physics. To borrow an analogy from Howard Stein, the situation is akin to that in nineteenth century mathematics: the notion of an adequate mathematical proof, despite its significance to mathematical practice, had not yet been given a systematic treatment. Successful mathematical reasoning did not await the development of classical logic, however, just as the evaluation of physical theories proceeds without a canonical inductive logic. Below I will highlight two styles of argument that have been employed effectively in the history of physics to establish that theories have strong evidential support. Although I will not attempt to analyze these in detail, I hold that any proposed systematic account of inductive reasoning should be judged in part by its treatment of cases like these.

[2] My approach to these issues is indebted to the work of several philosophers, in particular Harper (2012), Norton (1993) (from whom I've borrowed the title of this section), Norton (2000), Smith (2014), and Stein (1994).

One style of argument exploits a theory's unification of diverse phenomena, exemplified by Perrin's famous case for atomism. Perrin argued for the existence of atoms based on agreement among 13 different ways of determining Avogadro's number N, drawing on phenomena ranging from Brownian motion to the sky's color. This case is particularly striking due to the diversity of phenomena used to constrain the value of N, and also to the ease of comparison of different results, all characterized in terms of the numerical value of a single parameter. This argument was only possible due to refinements of the atomic hypothesis, and extensions of statistical mechanics, that allowed precise formulations of relationships between the physical properties of atoms or molecules and measurable quantities. Perrin focused on N (the number of atoms or molecules in a mole of a given substance) in particular as a useful invariant quantity, and measured N in a series of experiments on Brownian motion, drawing on theoretical advances due to Einstein and others. (See Nye (1972) for a historical study of Perrin's work.) By roughly 1912, Perrin's arguments had succeeded in convincing the scientific community of the reality of atoms, decisively settling what had previously been regarded as an inherently intractable, "metaphysical" question.

This kind of overdetermination argument has been used repeatedly in the history of physics (see, in particular, Norton 2000). One common skeptical line of thought holds that theories are inherently precarious because they introduce new entities, such as atoms, in order to unify phenomena. Success fitting a body of data, the skeptic contends, merely reflects the flexibility of these novel theoretical constructs. The consistent determination of theoretical parameters from diverse phenomena counters the worry that the theory only succeeds due to a judicious tuning of free parameters. The overdetermination argument shows that, rather than the piecemeal success the skeptic expects, the theory succeeds with a single choice of parameters. The strength of this reply to the skeptic depends on the extent to which the phenomena probe the underlying theoretical assumptions in distinct ways. Furthermore, the diversity of phenomena minimizes the impact of systematic errors in the measurement of the parameters. The sources of systematic error relevant to Perrin's study of Brownian motion have little to do with those related to measurements of N based on radioactivity, for example. As the number of independent methods increases, the probability that the striking agreement can be attributed to systematic errors decreases.

The conclusion to be drawn from the overdetermination argument depends upon how unlikely the agreement is antecedently taken to be. The truth of the atomic hypothesis and kinetic theory implies an equation relating N to a number of quantities measurable by experimental study of Brownian motion, a second equation relating N to radioactivity, and so on. If the atomic hypothesis were false, there is no reason to expect these combinations of measurable quantities from different domains to all yield the same numerical value, within experimental error. This claim reflects an assessment of competing theories: what is the probability of a numerical agreeement of this sort, granting the truth of a competing theory regarding the constitution of matter? The overdetermination argument has little impact if there is a competing theory which predicts the same numerical agreements. In Perrin's case, by contrast, the probability assigned to the agreeing measurements of N, were the atomic hypothesis to be false, is arguably very low. In arguing for a low antecendent probability of agreement, Perrin emphasized the independence and diversity of the phenomena used to determine the value of N. (Obviously this brief account highlights only one aspect of Perrin's argument; see Chalmers (2011); Psillos (2011) for more thorough treatments.)

There is a second respect in which the conclusions to be drawn from an overdetermination argument must be qualified. These arguments typically bear on only part of a theory, namely whatever is needed to derive the connections between theoretical parameters and measurable quantities. Perrin's case is unusual in that the evidence bears directly on the central question in the dispute regarding atomism, unlike other historical cases in which this style of argument was not as decisive. The strength of an overdetermination argument depends on whether there is sufficient evidence to constrain all of a theory's novel components, or at least the ones at issue in a particular debate. The argument only directly supports parts of the theory needed to establish connections between measurable quantities and theoretical parameters. Identifying the distinct components of a theory and clarifying their contribution to its empirical success is often quite challenging, as the acceptance of the aether based on success of electromagnetic theory in the 19th century illustrates.

The second style of argument focuses on evidence that accumulates over time as a theory supports ongoing inquiry. A physical theory introduces a set of fundamental quantities and laws holding among them that provide the means for explicating some domain of phenomena. Accepting a theory implies a commitment to account for phenomena within this domain on the theory's terms, under the pressure of new discoveries and improving standards of experimental and observational precision. Often this involves treating complex phenomena via a series of successive approximations, with further refinements driven by discrepancies between current theoretical descriptions and observations. Resolving discrepancies by adding further details, without abandoning basic commitments, provides evidence that the theory accurately captures the fundamental physical relationships. The evidence is particularly strong when this process uncovers new features of the system that can be independently confirmed.

Newtonian gravitational theory supported centuries of research in celestial mechanics in just this sense. With the benefit of gravitational theory, one could approach enormously complicated orbital motions, such as that of the moon, via a series of idealizations that incorporate physical details thought to be relevant. Throughout the history of celestial mechanics, there have nearly always been systematic discrepancies between observations and trajectories calcuated based on all the relevant details known at a given time. Subsequent efforts then focused on identifying details left out of the calculation that might resolve the discrepancy. Leverrier's inference that an undiscovered planet was the source of discrepancies in Uranus's orbit is perhaps the most famous example of this type of reasoning. But in most cases the physical source that was eventually identified was not as concrete as an additional planet; the secular acceleration of the moon, for example, results from the slowing rotation of the Earth due to tidal friction. The new details are then incorporated in a more elaborate model, and the search for discrepancies continues. By the early twentieth century, calculations of orbital motion included an enormous number of details. The theory was sufficiently precise to reveal very subtle discrepancies, such as systematic errors in determining sidereal time due to a periodic fluctuation in the Earth's rotational speed.

Smith (2014) convincingly argues that the success of this line of inquiry provides much stronger support for Newtonian gravity than is apparent if the theory is simply treated as making a series of successful predictions. Theoretical models of celestial motions had to be in place to even identify the small discrepancies that were the target of analysis, and in that sense the theory itself underwrites its detailed comparison with observation. The core commitments of the theory place stringent constraints on the kind of new physical details that can be introduced to account for discrepancies. Furthermore, these additions could usually be checked using methods that did not depend on gravity — as with the discovery of Neptune in the location predicted by Leverrier, or measurements of the periodic fluctuation in the Earth's rotation, initially detected by astronomers, using atomic clocks. These independent checks on the details incorporated in ever more elaborate models support the theory's claim to have accurately identified the appropriate quantities and laws. It would be an enormous coincidence for a fundamentally incorrect theory to be so useful in discovering new features of the solar system.

These two historical cases illustrate strategies, arguably used throughout physics, to provide an effective response to skepticism regarding theoretical knowledge. This skepticism is inspired by the apparent circularity of relying so heavily on the very theory in question to support detailed comparison with observations. When interpreted with the aid of the theory, the phenomena yield constraints on the parameters of the theory, and discrepancies that can be the target of further work; yet neither the constraints nor the discrepancies are readily available without the guidance provided by theory. How do we guard against accepting a theory that is self-certifying, sufficiently flexible to avoid refutation despite its flaws? In both examples above, the response to this worry relies upon using multiple, independent lines of evidence, while acknowledging the theory-dependence of the reasoning. The prosaic point underlying this response is that using multiple sources of information, dependent on theory in different ways and with different sources of systematic error, minimizes epistemic risk. This response shifts the burden of proof onto the skeptic: if the underlying theory were false, it would be an enormous coincidence for all of the multiple ways of measuring a parameter to coincide, or for new features added to resolve discrepancies to be independently confirmed.

A second skeptical objection regards the nature of the claims supported by these arguments: can they be regarded as stable contributions to our knowledge of the natural world? Perrin made the case for atomism prior to the advent of quantum theory, and the reasoning in celestial mechanics described above precedes Einstein. Are these arguments undermined by quantum mechanics and general relativity, respectively? As a brief reply to this Kuhnian worry, consider the nature of the claims that are supported by the arguments above. These are claims that specific law-like relations hold between different physical quantities within some domain. Perrin's case depends upon the relations between atomic scale properties and macroscopic, measurable properties. The development of celestial mechanics supports a variety of claims about what features of the solar system are relevant to planetary motions. In these two cases, the claims in question are arguably preserved through theory change, in the sense that there are lawlike relations in the successor theory which are approximated, within a restricted domain, by corresponding relations in the preceding theory.[3] This is true in spite of the dramatic conceptual differences between classical mechanics and quantum mechanics, and between Newtonian gravity and general relativity. As a consequence, the reasoning employed in arguing for the atomic hypothesis and in the development of celestial mechanics is validated rather than undermined by the successor theories.

There may be cases in which a new theory recovers only the predictions of an earlier theory, without establishing the validity of its evidential reasoning in this stronger sense. I do not intend to rule out that possibility by fiat; the evidence may simply unravel, as in the case of Aristotelian natural philosophy. But the burden of proof rests with the new theory to explain away the apparent successes of an old theory when the latter is not recovered as an approximation. In the cases described above, no one has imagined a credible alternative theory that matches the successes of the atomic hypothesis or Newtonian gravity without recovering aspects of these theories as limiting cases. This is the qualified sense in which theories supported by arguments like those described above constitute a stable contribution to our understanding of nature.

3 The Standard Model and Inflation

The standard model of cosmology (SMC) is based on bold extrapolations of theories that have been well-tested by Earth-bound experiments and astronomical observations. The interpretation of cosmological data depends, to varying degrees, on a background cosmological model, and hence assumes the validity of extrapolating general relativity to length scales roughly 14 orders of magnitude greater than those where the theory is subject to high precision tests. The SMC describes the contents of the universe and their evolution based on the Standard Model of particle physics, supplemented with two distinctive types of matter — dark matter and dark energy — that have so far only been detected due to their large-scale gravitational effects. Cosmological observations performed over the last few decades substantiate the enormous extrapolations and novel assumptions of the SMC.

[3] Determining how to recover the preceding theory in an appropriate limit is often surprisingly subtle, and I do not have the space to explore the issue fully here. In the case of Perrin's argument, for example, the central assumptions of kinetic theory Perrin used in his study of Brownian motion are good approximations to a quantum statistical mechanical treatment, except in cases of low temperatures or high densities; a full discussion would consider the approximations involved in the other methods for measuring N as well. For celestial mechanics, the claims in question regard the impact of, for example, Neptune on Uranus's orbit. The current practice of modeling the solar system using Newtonian physics with general relativistic corrections presupposes that the Newtonian description is a valid approximation. Finally, this claim regarding continuity does not require that the earlier theory provides a full account of the phenomena. Perrin did not have a complete theory of the nature of molecules, for example; he was well aware that the problem of specific heats identified by Maxwell had not been solved.

The development of a precise cosmological model compatible with the rich set of cosmological data currently available is an impressive achievement. Yet cosmology clearly relies very heavily on theory; the cosmological parameters that have been the target of observational campaigns are only defined within a background cosmological model. The SMC includes several free parameters, such as the density parameters characterizing the abundance of different types of matter, each of which can be measured by a variety of different types of observations.[4] CMB observations, in particular, place powerful constraints on many cosmological parameters. (Inferences to parameter values from observations of the CMB typically require prior assumptions regarding the nature of the primordial power spectrum, and there are several parameter degeneracies that cannot be resolved based solely on CMB observations.) There are a variety of independent ways of measuring the cosmological parameters that depend on different aspects of theory and have different sources of observational error. For example, the abundance of deuterium produced during big bang nucleosynthesis depends sensitively on the baryon density. Nucleosynthesis is described using well-tested nuclear physics, and the light element abundances are frozen in within the "first three minutes." The amplitudes of the acoustic peaks in the CMB angular power spectrum can be used to determine the baryon density at a later time (recombination, at $t \approx 400,000$ years), based on quite different theoretical assumptions and observational techniques. Current measurements fix the baryon density to an accuracy of one percent, and the values determined by these two methods agree within observational error. This agreement (augmented by other consistent measurements) is an important consistency check for the SMC.

The strongest case for accepting the SMC rests on the evidence in favor of the underlying physics in concert with the overdetermination of cosmological parameters. (See, e.g., Peebles et al. (2009), §5.4 for a brief discussion of tests of the SMC emphasizing the importance of independent measurements of the parameters.) The overdetermination argument has a similar structure to Perrin's argument for atomism described above. The case for the SMC does not yet have the diversity of independent lines of measurement that made Perrin's case so powerful; there are unexplained discrepancies among some measurements; individual measurements are not as precise as those available in many other areas of physics; and there are theoretical loopholes related to each measurement. But the essential epistemic point is the same: due to the diversity of measurements, it is unlikely that evaluation of the SMC has been entirely misguided due to incorrect theoretical assumptions or systematic observational errors. Several lines of observational and theoretical work currently being pursued promise to substantially strengthen the evidential support for the SMC.

Several of the cosmological parameters characterize the universe's "initial" state.[5] The SMC describes the large-scale structure of the universe as a perturbed Friedmann-Lemaître-Robertson-Walker (FLRW) model. The FLRW models are homogeneous and isotropic solutions of EFE. The models are topologically $\Sigma \times \mathfrak{R}$, visualizable as a "stack" of three-dimensional spatial surfaces $\Sigma(t)$ labeled by cosmic time t. The worldlines of "fundamental observers," at rest with respect to matter, are orthogonal to these surfaces, and the cosmic time corresponds to their proper time. EFE simplify to two equations governing $R(t)$, the spatial distance between fundamental observers.

One cosmological parameter — the spatial curvature, Ω_k — characterizes which of the FLRW models best fits observations. It is constrained by observations to be very close to zero, corresponding to the "flat"

[4] See Beringer et al. (2012) for a review of evidence bearing on the cosmological parameters. The total number of parameters used to specify a cosmological model varies in different studies, but typically 5-10 fundamental parameters are used to determine the best fit to a given data set. (Specific models often require a variety of further "nuisance parameters" to account for astrophysical processes.)

[5] This is often taken to be the state as specified at the "boundary of the domain of applicability of classical GR" — e.g., at the Planck time, $t \approx 10^{-43}$s. Appropriate initial data for Einstein's field equations (EFE) specify the full solution (for globally hyperbolic spacetimes), so the choice of a specific cosmic time at which to characterize the initial state is a matter of convention. But this conventional choice is significant if a given solution is treated as a perturbed FLRW solution, since dynamical evolution modifies the power spectrum of perturbations.

FLRW model whose spatial hypersurfaces $\Sigma(t)$ have zero curvature. For the flat model, the total energy density takes exactly the value ($\Omega = 1$) needed to counteract the initial velocity of expansion, $\dot{R} \to 0$ as $t \to \infty$. $\Omega =: \frac{\rho}{\rho_c}$, where ρ is the mass-energy density and the critical density is defined as $\rho_c = \frac{3}{8\pi}\left(H^2 - \frac{\Lambda}{3}\right)$. H is the Hubble "constant" (which in fact varies with cosmic time), defined as $H = \frac{\dot{R}}{R}$, and Λ is the cosmological constant. Other parameters characterize perturbations to the underlying FLRW model, which are fluctuations in mass density needed to provide the seeds for structure formation via gravitational clumping. If these fluctuations obey Gaussian statistics, they can be fully characterized in terms of a dimensionless power spectrum $\mathcal{P}(k)$. The power spectrum of the primeval mass distribution in the SMC takes the simple form of a power law, $\mathcal{P}(k) \propto k^{n_s}$. This power spectrum is parametrized in the SMC by two numbers. The first, the spectral index n_s, is equal to unity if there is no preferred scale in the power spectrum; observations currently favor $n_s = 0.96$, indicating a slight "blue tilt" in the power spectrum, with less power on smaller length scales. A second number is needed to specify the amplitude of the perturbations. (There are a few different ways of doing so. For example, σ_8 is the mass variance of the primordial distribution within a given radius (defined in terms of another parameter, the distance scale h: $8h^{-1} \approx 11Mpc$, given current estimates of h), projected forward to the current time using linear perturbation theory.)

The initial state required by the SMC has three particularly puzzling features. First, it is surprising that the simple, uniform FLRW models can be used at all in describing the early universe. These models have a finite horizon distance, much smaller than the scales at which we observe the CMB.[6] The observed isotropy of the early universe — revealed most strikingly by the temperature of the CMB — supports the use of the FLRW models; yet these observations cover thousands of causally disjoint regions. Why did the universe start off with such a glorious pre-established harmony? Second, an FLRW model close to the "flat" model, with nearly critical density at some specified early time is driven rapidly away from critical density under FLRW dynamics; the flat model is an unstable fixed point under dynamical evolution. In order for observations at late times to be compatible with a flat model, the initial state has to be *very* close to the flat model (or, equivalently, *very* close to critical density, $\Omega = 1$). (It follows from the FLRW dynamics that $\frac{|\Omega - 1|}{\Omega} \propto R^{3\gamma - 2}(t)$. $\gamma > 2/3$ if the strong energy condition holds, and in that case an initial value of Ω not equal to 1 is driven rapidly away from 1. Observational constraints on $\Omega(t_0)$ can be extrapolated back to a constraint on the total energy density of the Planck time, namely $|\Omega(t_p) - 1| \leq 10^{-59}$.)

Finally, the perturbations to the flat FLRW model postulated in the SMC are challenging to explain physically. It is not clear what physical processes could account for the amplitude of the perturbations. Suppose, for example, that one takes the "initial" perturbation spectrum to be imprinted at $t_i \approx 10^{-35}$s. Observations imply that at this time the initial perturbations would be far, far smaller than thermal fluctuations. (Blau and Guth (1987) calculate that observations imply a density contrast $\frac{\delta\rho}{\rho} \approx 10^{-49}$ at t_i, nine orders of magnitude *smaller* than thermal fluctuations.) In addition, the perturbations of the appropriate scale to eventually form galaxies would, in the early universe, be coherent at scales that seem to conflict with the causal structure of the FLRW models. A simple scaling argument shows that the wavelength λ of a given perturbation "crosses the horizon" with expansion, at the time when $\lambda \approx H^{-1}$ (where H^{-1} is the Hubble radius). Assuming that the perturbation spectrum is scale invariant, and for a simple model with $R(t) \propto t^n$ ($n < 1$) the wavelength of a given mode simply scales with the expansion $\lambda \propto t^n$. H^{-1} scales as $H^{-1} \propto t$; as a result, the Hubble radius "crosses over" perturbation modes with expansion. Prior to horizon-crossing, the perturbation would have been coherent on length scales greater than the Hubble radius. The Hubble radius is typically regarded as marking the limit of causal interactions, and as a result it is puzzling how normal physics operating in the early universe could produce coherent perturbations at such scales.[7]

[6]A horizon is the surface in a time slice t_0 separating particles moving along geodesics that could have been observed from a worldline γ by t_0 from those which could not. For a radiation-dominated FLRW model, the expression for horizon distance d_h is finite; the horizon distance at decoupling corresponds to an angular separation of $\approx 1^\circ$ on the surface of last scattering.

[7]The Hubble radius is defined in terms of the instantaneous expansion rate $R(t)$, by contrast with the horizon distance, which

Since the late 70s, cosmologists have sought a physical understanding of how such an unusual "initial state" came about. On a more phenomenological approach, the gravitational degrees of freedom of the initial state could be chosen to fit with later observations. Inflation in effect replaces such a specification with a hypothesis regarding the initial conditions and dynamical evolution of a proposed "inflaton" field (or fields). In the simplest inflationary models, a single field ϕ, trapped in a false vacuum state, triggers a phase of exponential expansion. If the inflaton field ϕ is homogeneous, then the false vacuum state contributes an effective cosmological constant to EFE, leading to quasi-de Sitter expansion.[8]

The resulting spurt of inflationary expansion can provide a simple physical account of the SMC's starting point, as emphasized with sufficient clarity to launch a field by Guth (1981). Inflationary expansion stretches the horizon length; for N "e-foldings" of expansion the horizon distance d_h is multiplied by e^N. For $N > 65$ the horizon distance, while still finite, encompasses the observed universe. The observed universe could then have evolved from a single pre-inflationary patch, rather than encompassing an enormous number of causally disjoint regions. (This pre-inflationary patch is larger than the Hubble radius (Vachaspati and Trodden 2000), however, so inflation does not dispense with pre-established harmony.) During an inflationary phase the density parameter Ω is driven *towards* one. (This is apparent from the equation above, given that $\gamma = 0$ during inflation.) An inflationary stage long enough to solve the horizon problem drives a large range of pre-inflationary values of $\Omega(t_i)$ close enough to 1 by the end of inflation to be compatible with observations.

The most remarkable feature of inflation, widely recognized shortly after Guth's paper, was its ability to generate a nearly scale-invariant spectrum of density perturbations with correlations on length scales larger than the Hubble radius.[9] Inflation produces density perturbations by amplifying vacuum fluctuations in a scalar field ϕ, with characteristic features due to the scaling behavior of the field through an inflationary phase. Start with a massless, minimally coupled scalar field ϕ evolving in a background FLRW model. The Fourier modes ϕ_k of linear perturbations to a background solution are uncoupled, with evolution like that of a damped harmonic oscillator. For modes such that $\frac{k}{R} << H$, the damping term is negligible, whereas those with $\frac{k}{R} >> H$ evolve like an over-damped oscillator and "freeze in" with a fixed amplitude. The inflationary account runs very roughly as follows. (This behavior follows from the Klein-Gordon equation in an FLRW spacetime, considering linearized perturbations around a background solution; see Mukhanov et al. (1992) for a comprehensive review of the evolution of perturbations through inflation, or Liddle and Lyth (2000) for a textbook treatment.0 Prior to inflation one assumes a vacuum state, i.e. the modes ϕ_k are initially in their ground state. For $\frac{k}{R} << H$ the modes evolve adiabatically, remaining in their ground states. This account is not sensitive to exactly when a given mode is assumed to be born in its ground state. During inflation the modes scale with the exponential expansion whereas H is approximately constant. Due to this scaling behavior, modes will reach the horizon scale $\frac{k}{R} \approx H$ — "horizon exit". The damping term is then no longer negligible and the modes "freeze in" as they cross the horizon. Modes then "re-enter" the horizon after inflation has ended, because in standard FLRW expansion the scaling behavior is reversed. Finally, these modes are treated as classical density perturbations upon re-entering the horizon. (This is a quantum to classical transition; whether it can be justified by appeals to squeezing of the quantum state and decoherence is contentious.) This evolution leads to a nearly scale invariant spectrum, with the amplitude of the pertubations fixed by the energy scale of inflation at horizon exit (as discussed below). (The spectrum is

---

depends upon the expansion history over some interval (the particle horizon, e.g., depends on the full expansion history). For radiation or matter-dominated solutions, the two quantities have the same order of magnitude.

[8]The stress-energy tensor is given by $T_{ab} = \nabla_a \phi \nabla_b \phi - \frac{1}{2} g_{ab} \left(g^{cd} \nabla_c \nabla_d \phi - V(\phi) \right)$, where $V(\phi)$ is the effective potential; for a homogeneous state, such that $V(\phi) >> g^{cd} \nabla_c \nabla_d \phi$, $T_{ab} \approx -V(\phi) g_{ab}$, leading to $R(t) \propto e^{\xi t}$ with $\xi^2 = \frac{8\pi V(\phi)}{3}$.

[9]This is sometimes regarded as a successful prediction of inflation, since this feature of inflation was initially not known to many researchers (despite early results, including Starobinsky 1979, 1980; Mukhanov and Chibisov 1981). Yet the initial prediction based on specific models under discussion (with inflation driven by a Higgs field in a grand unified theory) was incorrect, as the amplitude of pertubations was far too large. Insuring the correct amplitude leads to one of the "fine-tuning" problems of inflation, since the coupling of the scalar field driving inflation has to be very small.

not *exactly* scale invariant because the Hubble constant is not truly constant throughout inflation.)

To provide an account of the SMC's initial state, the inflationary phase has to be followed by a stage called "re-heating." Any matter or radiation present prior to inflation is rapidly diluted away during inflationary expansion, leaving a universe that is essentially empty except for the inflaton field. Reheating is required to fill the universe with matter and radiation, with temperature and densities appropriate for subsequent evolution within the standard big bang model.

Inflation provides a physical account of otherwise puzzling features of the starting point for the SMC. This is often described as solving "fine-tuning" problems of the SMC. Imagine choosing a cosmological model at random from among the space of solutions of EFE. Even without a well-defined measure on this space of solutions, it seems obvious that an FLRW model (or a perturbed FLRW model) must be an incredibly "improbable" choice. According to the SMC alone, what we observe is incredibly improbable; according to the SMC *plus inflation*, on the other hand, what we observe is to be expected, because "generic" pre-inflationary states lead to an appropriate starting point for the SMC. There are several objections to this line of argument, some of which go back to an incisive early criticism by Penrose.[10] Perhaps the most fundamental objection regards the starting point for the argument: why should we treat the initial state as "generic," a "random choice" from among all possible states? (Possible according to which theory?) It is also not clear that inflation succeeds by its own lights: Penrose, in particular, argued that a pre-inflationary patch with an appropriate state to trigger the onset of inflation should be *less likely* than an initial state for the SMC (without inflation). I won't explore these issues further here, in part because many proponents of inflation apparently regard the emphasis on fine-tuning as part of the initial motivation for inflation that can now be replaced with a more powerful empirical argument (see, e.g., Liddle and Lyth 2000, p. 5), to which I now turn.

4 Assessing Inflation

Inflation provides a promising account of the origins of the initial state for the SMC, and at the same time opens up the prospect of using observations of the CMB and large scale structure to constrain physics at an energy scale of $\approx 10^{15} - 10^{16}$ GeV. Unlike other competing theories, it has not been ruled out as the observational picture of the early universe has come into sharper focus over the last thirty years. Observations have led to a remarkably simple picture of the state of the early universe, which is well-described by a flat FLRW model, with Gaussian, adiabatic, linear, nearly scale invariant density pertubations. Inflation generates primordial fluctuations in the very early universe, at length scales larger than the Hubble radius. As they cross the Hubble radius, they set up coherent oscillations leading to acoustic peaks in the CMB power spectrum. Observations of acoustic peaks support the primordial nature of the fluctuations, contrasting with predictions from competing models of structure formation based on active sources for fluctuations (such as topological defects). (See, e.g., Durrer et al. (2002) for a review of structure formation via topological defects, and its contrasting predictions for CMB anisotropies.) That inflation is compatible with this observational picture of the early universe is an important success.[11] Does this amount to mere compatibility with the data, or does inflation fulfill its promise of providing a physical understanding of the early state? Here I will briefly assess challenges to providing stronger evidence in favor of inflation, based on following the strategies described above.

The inflaton is typically treated as a new field to be added to the Standard Model of particle physics.

[10] See Penrose 2004, Chapter 28 for a recent exposition of the arguments he first made in the early 80s; see also Earman and Mosterin (1999); Albrecht (2004); Hollands and Wald (2002b); Gibbons and Turok (2008); Carroll and Tam (2010) for related discussions.

[11] Here I will not address other conceptual and theoretical problems related to inflation, discussed in, e.g., Earman and Mosterin (1999); Turok (2002); Hollands and Wald (2002b); Brandenberger (2014).

Michael Turner called inflation a “paradigm without a theory” to emphasize the resulting flexibility of inflation. A bewildering variety of different inflationary models have been proposed, so many that theorists complain of the difficulty in finding an unused name for a new model. Many of the models can be characterized in terms of the Lagrangian proposed for the inflaton field:

$$\mathcal{L} = -\frac{1}{2}g^{ab}\partial_a\phi\partial_b\phi - V(\phi) + \mathcal{L}_I(\phi, A_a, \psi, ...), \tag{1}$$

where $V(\phi)$ is the effective potential, and $\mathcal{L}_I$ is an interaction term, specifying interactions with other fields in the Standard Model. Assuming that inflation is driven by a single field with a Lagrangian with this form already reflects some simplifications. Inflationary models with multiple scalar fields have been developed, motivated by proposals in high-energy physics that include many light scalar fields expected to be dynamically relevant in the early universe. But Planck observations support restricting attention to simple single-field models. Planck 2015 data provides strong evidence that the perturbations are adiabatic, which is compatible with simple single-field models; the failure to detect non-Gaussianities further supports the use of single-field models, and the choice of a standard kinetic term (the first term) in the Lagrangian (see, e.g., the discussions in Ade et al. (2015), §10, and Martin (2015)). A model from this class is characterized by a choice of the effective potential $V(\phi)$ and interaction term $\mathcal{L}_I$, along with assumptions regarding initial conditions for the field.

Observations of the CMB and large scale structure constrain the Lagrangian in two main ways. The primordial fluctuations place constraints on the effective potential well before the end of the inflationary phase. Inflation generates scalar and tensor perturbations whose physical properties depend on the features of the effective potential $V(\phi)$ at horizon exit, with $\frac{k}{R} \approx H$. Perturbations relevant to CMB observations typically crossed the horizon at ≈ 60 e-foldings before the end of inflation, whereas those that are re-entering the horizon now were produced a few e-foldings later. (The calculation is model-dependent and depends on assumptions regarding the reheating temperature; this estimate holds for a variety of slow-roll models with plausible further assumptions.) The features of scalar and tensor perturbations amplified through the inflationary phase can be described, in some cases, with equations relating the perturbation spectra to the value of $V(\phi)$ and its derivatives (at the scale when the perturbations crossed the horizon). Equations have been derived for models satisfying the slow-roll approximation, although it is not possible in general to calcuate the perturbation spectra for an arbitrarily chosen $V(\phi)$. “Slow-roll” models feature a flat effective potential, such that (roughly) $V', V'' << V$, leading to a long inflationary phase, sufficient to solve the horizon and flatness problems. (Here $'$ is the derivative $\frac{d}{d\phi}$. The slow roll conditions are constraints on V', V'' which insure that the damping term $(3H\dot{\phi})$ dominates over $\ddot{\phi}$ in the equation of motion for the inflation field: $\ddot{\phi} + 3H\dot{\phi} + V'(\phi) = 0$.) For these models there are simple expressions for the amplitude, spectral index, and “running” of the spectral index for both scalar and tensor perturbations in terms of V, V', V''. The scalar and tensor perturbation spectra are not independent, and a consistency relation, relating the spectral index of the tensor perturbations to the ratio of amplitudes of scalar and tensor perturbations, r, can be obtained by solving to eliminate V. (More generally, in a pertubative treatment there is a hierarchy of consistency relations. There are a few different parametrizations used in relating the effective potential to observable features of the perturbations, including slow roll parameters and Hubble flow parameters.)

A successful account of reheating depends on a different part of $V(\phi)$, along with the interaction term $\mathcal{L}_I$. Early accounts of inflation treated reheating as occurring when the inflaton field oscillated near the true minima of $V(\phi)$, assumed to be much steeper than the flat plateau needed for slow-roll, transferring energy to other particle species. Subsequent work has focused on energy transfer from the inflaton field to other particle species via coherent oscillations with parametric resonance. Observational constraints on the details of reheating are weaker than those related to the generation of primordial fluctuations (see, e.g. Martin et al. 2015).

There are several different approaches to reconstructing the inflaton potential from observations, and evaluating competing inflationary models (see, e.g. Lidsey et al. 1997; Martin et al. 2013). Given the wealth of observational data already available, upcoming observations, and the large variety of inflationary models, these techniques naturally focus on determining which models best fit the data. Martin et al. (2013), for example, adopts a Bayesian approach to analyze 193 single-field slow-roll inflationary models, concluding that 9 models with "plateau"-shaped potentials are preferred. The method relies on statistical tools optimized for determining the best model given the inherent noise and uncertainty of observational data. The ranking weighs the closeness of fit to the data a given model achieves against the complexity of the model, to avoid the pitfall of overfitting the data, but it is not designed to assess physical plausibility of a given model. Although this issue deserves further scrutiny, the epistemic point addressed in the model selection literature is distinct from the question addressed by the historical strategies discussed above. Finding the best fit model, granting the general framework used for interpreting the data, is not the same as evaluating the framework itself, although obviously the existence of a successful model — or lack of one — is relevant to this second task. The historical strategies aim to assess the validity of the underlying framework, to guard against being systematically misled by accepting an incorrect framework that nevertheless accommodates the data.

Turning to the first strategy discussed above, observations do provide independent constraints on the underlying inflationary mechanism for amplifying perturbations. A scale-invariant spectrum of scalar perturbations was proposed well before inflation on general grounds (Harrison 1970; Peebles and Yu 1970; Zel'dovich 1972). But there is not a similar argument in favor of a scale-invariant tensor perturbation spectrum, or any theory-independent reason to expect the two spectra to be linked as reflected in the consistency relation. Furthermore, measurements of the tensor perturbations directly constrain $V(\phi)$ at the point where a given length scale crossed the horizon. Measuring the tensor perturbation spectrum at different length scales, if it were feasible observationally, would give a direct reconstruction of $V(\phi)$.[12] Detection of CMB B-mode polarization, leading to a measurement of r, along with a measurement of the spectral index for tensor perturbations, n_t, directly tests the consistency relation. Measuring r is the target of a number of post-Planck missions, but the follow-up measurement of n_t is particularly challenging for small values of r. The possibility of nailing down the inflationary mechanism for amplifying perturbations in this fashion is certainly one of inflation's most appealing features.

There are, however, several contrasts with overdetermination arguments such as Perrin's. The first contrast regards the target of the argument, the Lagrangian for the inflaton field — and in particular the function $V(\phi)$ and the various couplings included in the interaction term $\mathcal{L}_I$. It is obviously much more challenging to provide a compelling overdetermination argument for the Lagrangian as opposed to a single number N. Furthermore, the existing observational constraints apply to two distinct dynamical regimes of the inflaton's evolution: the amplification of quantum fluctuations at horizon crossing, ≈ 60 e-folds before the end of inflation, compared with the decay of the inflaton and reheating at the very end of the inflationary phase. Inflationary models are a package deal rather than a single ticket: without theoretical constraints on the properties of the inflaton field, one can choose the shape of the potential relevant to amplification of perturbations, and then separately choose the shape of the potential near the true minimum and couplings in the interaction term. As long as this remains a relatively free choice, with weak constraints imposed in either direction, success in these two distinct dynamical regimes does not provide overlapping constraints.

The evidential situation changes substantially for an inflaton Lagrangian that is identified within a spe-

[12] This may even include constraints based on solar-system scale gravitational wave observatories. Boyle et al. (2014) argue that if there is a tilt in the tensor perturbation spectrum, as suggested by the BICEP2 initial results, the proposed Big Bang Observatory would provide a second set of measurements at a scale 10^{18} smaller than those relevant for the CMB, providing an enormous lever arm for more precise tests of inflation. See also Alvarez et al. (2014) for an overview of tests of inflation based on large-scale structure.

cific particle physics model. In such a case, the parameters appearing in the Lagrangian are constrained by cosmological data related to the details of inflation, as well as whatever experimental data is relevant to the particle physics model. This would provide a compelling set of independent constraints. Furthermore, since the inflaton model would be a single ticket item in this case, different cosmological measurements provide overlapping constraints on the Lagrangian. Yet the promise of directly identifying a canonical candidate for the inflaton field has not been fulfilled; instead, there has been a proliferation of toy models of inflation. Constructing physically plausible models for the inflaton has been difficult because $V(\phi)$ has to be very flat. The prospects for re-establishing a tighter link through direct experimental study are extremely bleak: the properties required for an inflaton field in a slow-roll model insure that it can only feasibly be studied observationally through its impact on the early universe.

Even without resolving the identity of the inflaton, the case for inflation can be strengthened by imposing other constraints on the class of allowed models. In practice this is reflected in assessments of the plausibility of different inflationary models, given assumptions about physics at the appropriate energy scale. There have also been proposals to characterize how inflationary predictions depend upon the amount of fine-tuning of the potential, which lead to constraints on the parameter ranges compatible with less finely-tuned potentials. (Boyle et al. (2006), for example, characterize fine-tuning in terms of the number of zeroes appearing in the slow-roll parameter η and its first derivative (with respect to the number of e-folds), which is intended as a measure of the number of "features" added to the effective potential; inflaton models with little fine-tuning in this sense favor specific parameter ranges for n_s, r.) On either of these approaches, considerable weight is put on the further constraints imposed in the name of plausibility or simplicity. Past debates regarding the viability of different types of models make the challenges to achieving consensus on these questions clear.[13]

A final contrast regards the assessment of alternatives. Perrin argued that the agreeing measurements of N would be an enormous coincidence if the atomic hypothesis were false. How likely is the simple early state required by the SMC, if inflation did not occur? Turok (2002), for example, remarks that "*The success of the simplest inflationary models is perhaps more of a success for simplicity than it is for inflation*" (p. 3458, emphasis original). Any early universe theory that generates a nearly scale invariant spectrum of primordial fluctuations will match many of inflation's successes. Theorists have discovered several different ways of generating such fluctuations, ranging from alternative ways of modifying causal structure (varying speed of light theories) to "bounce" models, which replace the Big Bang singularity with a Big Bounce. They treat the primordial fluctuations as generated prior to the bounce, although details of implementation, and the physics used to construct the model, differs substantially among the different models. (See Brandenberger (2014) for an overview of the matter bounce and string gas models, and Lehners (2008) for a review of ekpyrotic and cyclic models.) There is no reason to expect a consistency relation between tensor and scalar perturbations in a "simple" initial state, and this relation also discriminates between inflation and other models for the generation of primordial fluctuations. Further observational work, in particular detection of primordial gravitational waves and tests of the consistency relation, would lead to a much stronger case that the observed properties of the early universe would be an enormous coincidence if inflation were false.

Regarding the second strategy, I am unaware of any case in which inflation has been used to uncover a new feature of the early universe that can be independently checked. There are a variety of ways in which inflationary model-building has become more sophisticated, with a much clearer understanding of what needs to be in place in a full account of inflation. There is no shortage of theoretical innovation in building inflationary models. Inflation has also guided observational work by identifying features that can be used to constrain inflationary models and contrast inflation with competing theories. But the question is whether inflation has allowed cosmologists to identify robust physical features of the early universe that can be tested

[13]For example, inflation is commonly taken to predict a flat universe with $\Omega_0 = 1$. There were heated debates in the mid to late 90s regarding so-called "open inflation" models that yield a value of $\Omega_0 \approx .2 - .3$, which was at that time favored by observations. Insofar as these were regarded as plausible models, inflation no longer predicts flatness, and the value of Ω_0 instead provides a constraint on the parameter space for models (see, e.g., the discussions in Turok 1997).

in ways that do not assume inflationary theory itself. There are no analogs, as far as I am aware, of adding a new physical feature as part of the model that can, like the existence of Neptune, be easily checked by other means. This is in part due to the observational inaccessibility of the early universe, but also to the lack of a canonical choice of the inflaton field. Given a fixed choice for the inflaton field, discrepancies with observations would force theorists to elaborate the model, possibly identifying new features of the early universe in the process. At present the choice of inflationary models is too flexible to support this kind of approach.

Above I emphasized the need for multiple, independent lines of evidence, in order to mitigate the theory-dependence of evidential reasoning. The challenges to pursuing the two historical strategies in cosmology both reflect our lack of accessibility to the early universe and to the energy scales of inflation. The observed state of the universe is compatible with inflationary models, but we have not yet developed a more detailed account of how inflation occurred. In the historical cases described above, it was ultimately the development of detailed accounts of the nature of atoms, and of the motions in the solar system and their causes, that provided confidence in the theories employed along the way. The alternative to regarding these theories as stable contributions to our knowledge of the natural world is to accept that, for example, measurements of the Earth's slowing rotation by atomic clocks simply happened to agree with measurements of the moon's motion, by an astronomical coincidence. A successful, detailed account of inflation, going beyond the initial step of using CMB data to constrain the inflaton Lagrangian, could support an argument of this sort. The challenge to taking the next steps is our lack of access to energy scales associated with the inflaton, and to specific quantities that discriminate among models.

The challenges to the observational program of further constraining the inflaton field have little to do directly with the distinctive features of cosmology, such as the uniqueness of the universe. Neither of the strategies described above require that the system under study can be experimentally manipulated. It is also not essential to consider a repeatable phenomena, with multiple instances subject to study. The inability to conduct relevant experiments, and lack of multiple instances, are often taken to distinguish cosmology from other areas of inquiry, leading to limits on what can be established (e.g. Munitz 1962). To make the contrast between limitations that are inherent to cosmology and problems of accessibility more vivid, imagine that an alien civilization provided us with an accelerator able to probe physics at $10^{16} GeV$. Access to the physics at this energy scale, to determine the properties of the inflaton (if it exists), would enable thorough development and testing of a detalied account of the universe's early history.

There is a more interesting challenge in cosmology regarding how to deal with initial conditions, and potential trade-offs between assumptions regarding initial conditions and dynamics (cf. Smolin 2015). Early discussions of inflation often emphasized its ability to "wash away" dependence on the pre-inflationary state of the universe, doing away with the need for assumptions about the initial state. (Collins and Stewart (1971) noted in response to a precursor to inflation that dynamics cannot completely "wash away" the initial state, however. Given fairly weak assumptions about the dynamics, it follows from standard existence and uniqueness theorems for differential equations that one can always find a pre-inflationary state that will lead to any given post-inflationary value of Ω_0 (for example).) But it is clear that inflation requires assumptions regarding the initial state of the inflaton field (homogeneous, with an appropriate value of $V(\phi)$, in a spacetime region larger than the Hubble radius), along with an appropriate form of the potential $V(\phi)$. These are sometimes called inflation's fine-tuning problems. Assumptions about what is a plausible initial state are also relevant to assessing the account of structure formation. Hollands and Wald (2002b) construct a simple model that produces a similar spectrum of density perturbations without an inflationary phase based on a different *Ansatz* for the initial conditions. Their model describes quantized sound waves in a perfect fluid, with the same "overdamping" of modes with $\lambda >> H^{-1}$ as in inflation. By contrast with inflation, there is no horizon crossing, so it is significant precisely when the modes are taken to be in a vacuum state.[14] The fine-tuning problems of inflation are often thought to be resolved within the context of eternal inflation, to which I now turn.

Many cosmologists hold that inflation is "generically eternal," in the sense that inflation produces a multiverse consisting of "pocket universes," where inflation has ended, even as inflationary expansion continues elsewhere. (See, e.g., Aguirre (2007) for an introduction to eternal inflation.) The mechanism leading to this multiverse structure is also assumed to lead to some variation in the physical parameters among the different pocket universes. The solution to the fine-tuning problems of the original inflationary models is based on invoking an anthropic selection effect: pocket universes featuring observers will be ones in which various necessary preconditions for the existence of life like us hold. These preconditions plausibly include the existence of structures like galaxies; the formation of galaxies depends upon the presence of small fluctuations in an expanding FLRW model; and the small fluctuations themselves are ultimately traced back to an initial state for ϕ and form of the effective potential $V(\phi)$ appropriate to trigger an inflationary phase.

Accepting eternal inflation undermines the observational program of attempting to constrain and fix the features of the inflaton field, in two senses.[15] First, the appeal to anthropic selection undercuts the motivation for introducing a specific dynamical mechanism for generating a multiverse. The anthropic argument is intended to counter the objection that ϕ and $V(\phi)$ probably do not have appropriate properties to initiate inflation. While that may be true in the multiverse as a whole, it is, the argument goes, not the case in the habitable pocket universes, which are expected to have undergone inflation in order to produce galaxies (for example). However, the argument works just as well with other proposed ensembles, as long as the observed universe is compatible with the underlying laws. Rather than the inflationary multiverse, why not simply consider a relativistic cosmological model with infinite spatial sections, and some variation among different regions? By parity of reasoning, even if a region with properties like our observed Hubble volume is incredibly improbable in general, it may be highly probable within the anthropic subset. I don't see a plausible way to refine the argument to draw a distinction between these two cases, so that one can preserve the original motivations for inflation while accepting eternal inflation.

The second challenge raised by eternal inflation regards the prospects for using evidence to constrain theory along the lines outlined above. Briefly put, the two strategies above both rely on the exactness of theory in order to develop strong evidence — either in the form of connections among different types of phenomena, or in the form of rigidity as a theory is extended to give a more detailed account. Eternal inflation is anything but exact. Deriving "predictions" from eternal inflation requires specifying the ensemble of pocket universes under consideration; a measure over this ensemble that is well-motivated; and a specification of the subset of the ensemble within which observers can be located. Each of these raises a number of technical and conceptual problems. But even if these are resolved, there are then several substantive auxiliary assumptions standing between the predictions of eternal inflation and the comparison with observations. Rather than using observations to directly constrain and probe the physics behind the formation of structure, we would instead be delimiting the anthropic subset of the multiverse.

[14] Hollands and Wald (2002a) propose to take the modes to be "born" in a ground state when their proper wavelength is equal to the Planck scale, motivated by considerations of the domain of applicability of semi-classical quantum gravity. The modes will be "born" at different times, continually "emerging out of the spacetime foam," with the modes relevant to large-scale structure born at times much earlier than the Planck time. By way of contrast, in the usual approach the modes at all length scales are specified to be in a ground state at a particular time, such as the Planck time. But the precise time at which one stipulates the field modes to be in a vacuum state does not matter given that the sub-horizon modes evolve adiabatically.

[15] See Ijjas et al. (2013, 2014) for an assessment of related problems, along with the response by Guth et al. (2014). I discuss these issues at greater length in Smeenk (2014).

5 Conclusion

Science often proceeds by making substantial theoretical assumptions that allow us to extend our reach into a new domain. My approach above has been to focus on asking how evidence can accumulate in favor of these assumptions as the research based on them advances. In many historical cases, subsequent research has established that a theory has to be accepted, at least as a good approximation, with the only alternative being to accept an enormously implausible set of coincidences. Based on the methodological insights gleaned from these historical cases, I have argued that the main problems with establishing inflation with the same degree of confidence stem from our lack of independent lines of access. In addition, eternal inflation undercuts the observational program devoted to constraining the inflaton field.

References

Ade, PAR et al. (2015). "Planck 2015. XX. Constraints on Inflation." *arXiv preprint arXiv:1502.02114* .

Aguirre, A. (2007). "Eternal inflation, past and future." *arXiv preprint arXiv:0712.0571* .

Albrecht, A. (2004). "Cosmic inflation and the arrow of time." In *Science and Ultimate Reality* , J. D. Barrow, P. C. W. Davies, and Charles Harper, eds. Cambridge: Cambridge University Press, 363–401.

Alvarez, M. et al. (2014). "Testing inflation with large scale structure: Connecting hopes with reality." *arXiv preprint arXiv:1412.4671* .

Barrow, J. D. and Liddle, A. R. (1997). "Can inflation be falsified?" *General Relativity and Gravitation* 29, 1503–1510.

Beringer, J. et al. (2012). "Review of particle physics." *Physical Review D* 86.

Blau, S. K. and Guth, A. (1987). "Inflationary cosmology." In *300 years of gravitation*, S. W. Hawking and Werner Israel, eds. Cambridge: Cambridge University Press, 524–603.

Boyle, L., Smith, K. M., Dvorkin, C. and Turok, N. (2014). "On testing and extending the inflationary consistency relation for tensor modes." *arXiv preprint arXiv:1408.3129* .

Boyle, L. A., Steinhardt, P. J. and Turok, N. (2006). "Inflationary predictions for scalar and tensor fluctuations reconsidered." *Physical review letters* 96, 111301.

Brandenberger, R. (2014). "Do we have a theory of early universe cosmology?" *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 46, 109–121.

Carroll, S. M. and Tam, H. (2010). "Unitary Evolution and Cosmological Fine-Tuning." *ArXiv e-prints* 1007.1417.

Carroll, S. (2014). "What scientific idea is ready for retirement? Falsifiability." URL http://edge.org/response-detail/25322.

Chalmers, A. (2011). "Drawing philosophical lessons from Perrin's experiments on Brownian motion." *The British Journal for the Philosophy of Science* 62, 711–732.

Collins, C. B. and Stewart, J. M. (1971). "Qualitative cosmology." *Monthly Notices of the Royal Astronomical Society* 153, 419–434.

Durrer, R., Kunz, M.n and Melchiorri, A. (2002). "Cosmic structure formation with topological defects." *Physics Reports* 364, 1–81.

Earman, J. and Mosterin, J. (1999). "A critical analysis of inflationary cosmology." *Philosophy of Science* 66, 1–49.

Ellis, G. F. R. and Silk, J. (2014). "Scientific method: Defend the integrity of physics." *Nature* 516, 321.

Gibbons, G. W. and Turok, N. (2008). "Measure problem in cosmology." *Physical Review D* 77, 063516.

Guth, A. (1981). "Inflationary universe: A possible solution for the horizon and flatness problems." *Physical Review D* 23, 347–56.

Guth, A., Kaiser, D. and Nomura, Y. (2014). "Inflationary paradigm after Planck 2013." *Physics Letters B* 733, 112–119.

Harper, W.L. (2012). *Isaac Newton's Scientific Method: Turning Data Into Evidence about Gravity and Cosmology*. Oxford Univ Pr.

Harrison, E. R. (1970). "Fluctuations at the threshold of classical cosmology." *Phys. Rev.* D1, 2726–2730.

Hollands, S. and Wald, R. (2002a). "Comment on inflation and alternative cosmology." Hep-th/0210001.

— (2002b). "Essay: An alternative to inflation." *General Relativity and Gravitation* 34, 2043–2055.

Ijjas, A., Steinhardt, P. J. and Loeb, A. (2013). "Inflationary paradigm in trouble after Planck 2013." *Physics Letters B* 723, 261–266.

— (2014). "Inflationary schism." *Physics Letters B* 736, 142–146.

Lehners, J. (2008). "Ekpyrotic and cyclic cosmology." *Physics Reports* 465, 223–263.

Liddle, A. and Lyth, D. (2000). *Cosmological Inflation and Large-Scale Structure*. Cambridge: Cambridge University Press.

Lidsey, J. E. et al. (1997). "Reconstructing the inflaton potentialan overview." *Reviews of Modern Physics* 69, 373.

Martin, J. (2015). "The observational status of cosmic inflation after planck." *arXiv preprint arXiv:1502.05733* .

Martin, J., Ringeval, C. and Vennin, V. (2013). "Encyclopaedia inflationaris." *arXiv preprint arXiv:1303.3787* .

— (2015). "Observing inflationary reheating." *Physical review letters* 114, 081303.

Mukhanov, V. F. and Chibisov, G. V. (1981). "Quantum fluctuations and a nonsingular universe." *JETP Letters* 33, 532–535.

Mukhanov, V. F., Feldman, H. A. and Brandenberger, R. H. (1992). "Theory of cosmological perturbations. part 1. classical perturbations. part 2. quantum theory of perturbations. part 3. extensions." *Physics Reports* 215, 203–333.

Munitz, M. K. (1962). "The logic of cosmology." *British Journal for the Philosophy of Science* 13, 34–50.

Norton, J (1993). "Determination of theory by evidence: How einstein discovered general relativity." .

— (2000). "How we know about electrons." In *After Popper, Kuhn and Feyerabend: Recent Issues in Theories of Scientific Method*, R. Nola and H. Sankey, eds. Dordrecht: Kluwer Academic Publishers, 67–97.

Nye, M. J. (1972). *Molecular reality: A perspective on the scientific work of Jean Perrin*. Macdonald.

Peebles, P. J. E. and Yu, J. T. (1970). "Primeval adiabatic perturbation in an expanding universe." *Astrophys. J.* 162, 815–836.

Peebles, P J E, Page Jr, L. A. and Partridge, R. B. (2009). *Finding the big bang*. Cambridge: Cambridge University Press.

Penrose, R. (2004). *The road to reality*. London: Jonathan Cape.

Psillos, S. (2011). "Moving molecules above the scientific horizon: On Perrin's case for realism." *Journal for general philosophy of science* 42, 339–363.

Smeenk, C. (2014). "Predictability crisis in early universe cosmology." *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 46, 122–133.

Smith, G. E. (2014). "Closing the loop." In *Newton and Empiricism*, Zvi Biener and Eric Schliesser, eds. Oxford: Oxford University Press, 262–351.

Smolin, L. (2015). "Temporal naturalism." *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* .

Starobinsky, A. (1979). "Spectrum of relict gravitational radiation and the early state of the universe." *JETP Letters* 30, 682–685.

— (1980). "A new type of isotropic cosmological models without singularity." *Physics Letters B* 91, 99–102.

Stein, H. (1994). “Some reflections on the structure of our knowledge in physics.” In *Logic, Metholodogy and Philosophy of Science, Proceedings of the Ninth International Congress of Logic, Methodology and Philosophy of Science*. 633–55.

Turok, N. (2002). “A critical review of inflation.” *Classical and Quantum Gravity* 19, 3449.

Turok, N., ed. (1997). *Critical dialogues in cosmology*. Singapore: World Scientific.

Vachaspati, T. and Trodden, M. (2000). “Causality and cosmic inflation.” *Physical Review D* 61, 3502–06. Gr-qc/9811037.

Zel’dovich, Ya. B. (1972). “A hypothesis, unifying the structure and the entropy of the universe.” *Mon. Not. Roy. Astron. Soc.* 160, 1–3.