RECEIVED WISDOM

By Phillip Helbig

Science is a self-correcting process. While a consensus implies a majority, what is consensus is not determined by majority vote *per se*, but rather by the majority being convinced (rightly or wrongly!) that something is correct. In some cases, that represents a reversal of an earlier situation. I briefly review several cases in cosmology (and astrophysics) in the last hundred years or so where the received wisdom at a particular time turned out to be wrong, leading to a new consensus. Whether that will happen for two relatively new topics remains to be seen.

Introduction

At least in some cases, there had been a real consensus among people who had thought about a problem, but it later turned out that that consensus had been wrong. I discuss some examples below. Many more details can be found in the references; the point here is not to give a detailed review, but to call attention to several topics which are similar in this respect. Also, my main point is to call attention to those episodes, rather than to explore the reasons involved (though I do comment on them briefly), which were certainly not all exactly the same. As my late history teacher used to say, "Just an observation, not a judgement".

Of course, scientists don't create theories in a vacuum and then see which are contradicted by random experiments. Although eventually a stable consensus is reached, that can be achieved by numerous routes. (I think of that as similar to the no-hair theorem for black holes: the final state is very simple and independent of a possibly very complicated past.) During such earlier phases, both the choice of theories and the choice of experiments depend on many non-scientific factors. The simple view is that a theory in contrast with experiment is falsified[1]. That is true as long as one can be sure that all aspects of the experiment are correct. In practice, while theories cannot be proved correct, only disproved, our faith in them can increase (or our Bayesian prior can be strengthened) if they continue to hold up in the light of new data and experiments. In such a case, the first, correct, reaction to an experiment which appears to disprove a theory is often to check the experiment. That was what Eddington meant by "it is also a good rule not to put overmuch confidence in the observational results that are put forward until they are confirmed by theory". While often used to caricature Eddington as an ivory-tower theorist (who also wrote "well, these experimentalists do bungle things sometimes" [2]), the full quotation is more interesting:

> Observation and theory get on best when they are mixed together, both helping one another in the pursuit of truth. It is a good rule not to put overmuch confidence in a theory until it has been confirmed by observation. I hope I shall not shock the experimental physicists too much if I add that it is also a good rule not to put overmuch confidence in the observational results that are put forward until they have been confirmed by theory. [3]

So it is a mixture of new data and prior belief which forms opinion. At least in part, prior opinion is based on previous data, so a contradiction between theory and experiment could falsify the theory, but it could also mean that the experiment is wrong. If the new experiment tests a previously untested prediction, then the situation is not immediately clear. If it is similar to previous experiments which agree with theory, then most would probably conclude that the new experiment is wrong, or at least should be checked before abandoning the theory (and the previous experiments, for which it needs to explain why they agree with the theory while the new one doesn't). Healthy scepticism is good, which is what Sagan meant by "extraordinary claims require extraordinary evidence" [4].* Shapley noted that one should be sceptical of observations: "No one trusts a model except the man who wrote it; everyone trusts an observation, except the man who made it." However, he also pointed out that "[t]heories crumble, but good observations never fade."

The combination of data and prior belief (also based on things which have a history of working well in science, such as Occam's razor and even a sense of beauty) is what is important. Of course, science can never absolutely confirm anything, only rule it out, but some ideas can be believed in with high confidence, and some of those can turn out to be wrong. That is an honest mistake, as opposed to holding on to a debunked idea or falsified theory for too long for irrational reasons.

Olbers's Paradox

I say little about this old chestnut here, since it has been discussed in detail by Harrison [6–15] (yes, *that* Harrison [16]), including a chapter in his excellent cosmology textbook [17] (chap. 24) and even in an entire book [18] on the topic; see also works by Wesson *et al.* [19] and Wesson [20]. Suffice it to say that the first person to glimpse the proper solution (our Universe is not old enough for it to have been filled with light) to the traditional riddle of why the sky is dark at night appears to have been Edgar Allan Poe [21] and that in this case there was not *one* consensus for a long time but rather various explanations had support at various times. Some of the explanations were wrong even according to the state of knowledge at the time. For example, absorption by dust cannot explain Olbers's Paradox

*Others had said something similar before Sagan; O'Toole [5] traces similar quotations from various people (including Pierre-Simon Laplace) back to 1708.

because the dust would heat up to be as bright as whatever is heating it. Other explanations, such as the hierarchical Universe of Charlier [22,23] (an idea, also in the context of a solution to Olbers's Paradox, going back to Lambert [24]), are valid solutions, but do not apply to our Universe. (Note that I am referring here to the traditional Olbers's Paradox. We now know that most of the photons in the Universe are in the cosmic microwave background (CMB). In that case, the expansion of the Universe explains why the CMB is no longer optically bright, as it is redshifted by a factor of 1000 or so. Expansion also plays a role in the traditional Olbers's Paradox, but is not the dominant effect [19,20].)

The cosmological constant

Whether, as claimed by Gamow [25,26], Einstein actually said that the cosmological constant was the biggest blunder of his life and, if so, what he meant by that is not completely clear [27–30]. However, it is certainly clear that he concentrated on models without a cosmological constant after the discovery of the expansion of the Universe, the most popular being the Einstein–de Sitter model [31]. At least with regard to the question of spatial curvature, though, the Einstein–de Sitter model, which has neither a cosmological constant nor spatial curvature, was a practical approximation for the data available at the time, not intended as anything more than a provisional model [32], though Einstein's belief that the cosmological constant is zero was certainly stronger. Even though Einstein was often wrong in his later years (see below), and even though the 'biggest blunder' story might not be true, his influence is partly to blame for the neglect of the cosmological constant from about 1930 to about 1990, although it was never completely out of fashion (*e.g.,* ref. 33). It was certainly true that observations weren't good enough to detect the cosmological constant until the advent of modern supernova cosmology, making use of the classic m–z relation [34,35], though there had been hints a few years before (*e.g.,* refs. 36–38). Of course, a (nearly) flat universe (perhaps inspired by inflation) and density parameter $\Omega_0 < 1$ ($\Omega = (8\pi G\rho)/(3H^2)$ where ρ is the density, G the gravitational constant, and H the Hubble constant; 0 refers to the current value because, in general, it changes with time) implies a positive cosmological constant; Peebles [39] found much earlier that such models fit the observations well. Some preferred to leave it out if possible to make things simpler, while others, noting that Nature uses all degrees of freedom available, preferred to leave it in with a value to be determined observationally. (Finding a value to be 0 usually indicates that there is some reason, such as a previously unknown conservation law.)

Departures from the Einstein–de Sitter model usually involved a smaller, but sometimes larger (*e.g.,* ref. 40), value for the density parameter Ω_0, usually with the (dimensionless) cosmological constant $\lambda_0 = 0$ ($\lambda = \Lambda/(3H^2)$, where Λ is the cosmological constant). That was probably due to the lack of enthusiasm for the cosmological constant, to the fact that observations weren't good enough to detect the value we believe that λ_0 has today, and perhaps also because calculations are easier than in a flat model ($\Omega_0 + \lambda_0 = 1$) or a non-flat model with (of course) $\Omega_0 > 0$ and $\lambda_0 \neq 0$. Since in the $\lambda_0 = 0$ case a spatially closed universe ($\Omega_0 > 1$) implies a universe which will collapse in the future and *vice versa*, the terms 'open universe' and 'closed universe' were sometimes used to refer to one of those aspects, sometimes the other, sometimes both. More general cases require more precise terminology.

Even after evidence which is now seen to support the idea of a positive cosmological constant had become available, for a while it was possible to find alternative explanations. For example, an early argument in favour of $\Lambda > 0$ was the combination of $\Omega_0 = 0.3$ or so and a nearly flat Universe (motivated by inflation); such a Universe would also be old enough to accommodate the oldest objects within it without an unreasonably low value for the Hubble constant. However, some argued that Ω_0 had been underestimated and/or that the increase in the measured value with scale could be extrapolated all the way up to $\Omega_0 = 1$ (corresponding to a flat Universe without a cosmological constant). The combination of better observations for traditional measurements and new tests, such as the m–z relation for type-Ia supernovae, eventually led to a new consensus (although there are still a few who still believe that $\Lambda = 0$, perhaps even some reading this article).

Certainly for some, one reason to believe that $\Lambda = 0$ is that some estimates from quantum mechanics predict a huge value: many, many orders of magnitude larger than astronomical limits. It seemed more likely that some unknown mechanism results in an exact cancellation, rather than an almost exact cancellation leaving a very small but non-zero value. It would be interesting to know to what extent that argument was responsible for leaving Λ out of various analyses. On the other hand, General Relativity (GR) says nothing about quantum mechanics and the Λ in GR could be different from that resulting from quantum mechanics; perhaps the latter is cancelled exactly and the former is what is observed. Still others might have rejected the quantum-mechanics estimate since it is obviously very wrong in some sense. (See also the section below on the cosmological-constant problem.)

In any case, the belief that the cosmological constant $\Lambda = 0$ was a relatively strong consensus which has been completely overturned. However, that consensus was probably due mainly to lack of data, perhaps strengthening a prejudice in favour of $\Lambda = 0$. The current consensus is certainly due mainly to more and better data, not only from the m–z relation for type-Ia supernovae but also from the CMB.

The Einstein–de Sitter universe as the standard model

So although the cosmological constant never went completely out of fashion, there was almost a consensus that there is no cosmological constant (nor spatial curvature) as the Einstein–de Sitter model to a large extent became the standard model of cosmology from about 1970 (*e.g.,* ref. 41 and references therein), after the demise of the Steady State theory[42,43]. Observationally, the Einstein–de Sitter model, in which quantities of interest are easy to calculate, was a good working model, since observations were not yet good enough to detect significant departures from it, though observers had almost always favoured a value for the density parameter $\Omega_0 < 1$ (*e.g.,* ref. 41 and references therein) (1 being the value in the Einstein–de Sitter model).

By the 1990s, many had pointed out that observations were better fitted with a low-density model with a positive cosmological constant; values $\Omega_0 \approx 0.3$ and $\lambda_0 \approx 0.7$ were a good fit to the observations (*e.g.,* refs. 36–38). Although today the cosmic-microwave-background (CMB) data alone provide tight constraints on λ_0 and Ω_0, already 20 years ago the combination of constraints from the m–z relation for type-Ia supernovae[34,35] (which did indicate accelerated expansion but, at that time, essentially measured $\Omega_0 - \lambda_0$) and the CMB (which, at that time, essentially measured $\Omega_0 + \lambda_0$) shifted the consensus to the standard model which we still have today. My guess is that the estimated values of λ_0 and Ω_0 will not change much and that 'dark energy' will never be shown to be anything other than a cosmological constant (even if, in contrast to now, there were strong theoretical reasons for believing so, that could probably never be detected observationally, at least not *via* traditional observational cosmology[44,45]).

The Einstein–de Sitter model was a consensus, though not universal. Overbye[46] gives a fascinating glimpse into the thinking of the time:

> [The late David] Schramm, a neutrino advocate, was one who claimed to know the big picture. He liked neutrinos because they could "close" the universe and make omega a perfect 1.0 [In a footnote, Overbye discusses the misleading nature of that oft-repeated statement.] That was the answer required by inflation. If you understood anything at all about grand unified theories and inflation, you would realize that omega had to be 1.0. A true physicist could think no other way; it was the paradigm. A cosmologist's job was to reconcile the observations with that number.
>
> We were outside his office talking to one of his graduate students, a young woman who was doing numerical simulations of the universe. She was doing great, he said, but why was she doing these runs with omega equal to 0.2?
>
> That was what Davis and White did, she answered.
>
> Schramm told her to get rid of those runs. "You're thinking like an astronomer instead of like a physicist," he snorted, adding, "Simon White never understood inflation."

Ironically, when considering the observations, Schramm himself had co-written a good and important and influential paper on the topic[47], favouring a low value for Ω_0. While it is perhaps not surprising that theorists interested in inflation favoured the Einstein–de Sitter model, I was somewhat surprised that Allan Sandage, often thought to be a good example of an empirical observational cosmologist ("We never really did talk about cosmological models" he said when remembering his time with Hubble[48]), seemed very convinced that the Einstein–de Sitter model must be correct[49]. Of course, without a cosmological constant, the Hubble constant would have to be low (as Sandage had always claimed, even before the idea of inflation was popular or even invented) in order to reconcile the calculated and observed ages of the Universe were $\Omega_0 = 1$.

The Einstein–de Sitter model implies a low Hubble constant, but the reverse is not necessarily true. My impression from hearing his lectures, which in that respect were more detailed than in the written version, is that he had latched on to the inflation-inspired Einstein–de Sitter model at least partly because it was compatible with his relatively low value for the Hubble constant, but also perhaps because it demanded a low Hubble constant.

Observationally, non-detection of spatial curvature implies only that the radius of curvature is too large for us to detect; it does not imply that the Universe must be exactly flat nor arbitrarily close to it, any more than the fact that we rarely notice the curvature of the Earth in day-to-day life implies anything other than that the radius of curvature is larger than the lengths we normally deal with. The idea that the Universe must be described as (almost) exactly flat stems from misguided ideas about stability (*e.g.,* ref. 41 and references therein) and/or from the belief that inflation must have made the radius of curvature of the Universe very much larger than we could detect.

Of course, the Einstein–de Sitter model is a special case of models without a cosmological constant. Although not completely accurate, to some extent one could say that while most believed that $\Lambda = 0$, observers thought that Ω_0 was to be determined observationally, and such observations tended to indicate a low value for Ω_0 (*e.g.,* ref. 50), while theorists preferred the Einstein–de Sitter model. As mentioned above, however, Sandage, an observer, was heavily influenced by theorists in his later years, becoming a firm believer in the Einstein–de Sitter model, whereas when younger he had been searching for two numbers, H_0 and q_0 (if $\lambda_0 = 0$, $q_0 = \Omega_0/2$; q is the deceleration parameter)[51]. Originally preferred for its simplicity, as originally intended, with time it became so popular that other models, even with $\Lambda = 0$, were considered almost fringe, as indicated by the story involving Schramm mentioned above. The fact that Peebles favoured the Einstein–de Sitter model was certainly influential, though I wonder whether his attempts to still make it work when it was becoming more and more obvious that it couldn't fit the observations were at least in part due to playing devil's advocate. So in this case the original reason for the consensus was relatively benign, but it held on longer than it should have due to continued support from influential figures in the field. Another reason, somewhat overlapping, is that some observations (*e.g.,* refs. 52–54) did seem to support, or at least be compatible with, the Einstein–de Sitter model, though uncertainties were large and the claim was too enthusiastic.

Gravitational waves

Einstein first thought that gravitational waves were real[55], then that they weren't, then again that they were real. The claim that they weren't was submitted to *Physical Review*, but was rejected. It was then submitted to the *Journal of the Franklin Institute*, but before publication it was revised to change the conclusion[56]. See the investigation by Kennefick[57] for some background; the incident led to Einstein never, with the exception of a letter to the editor in in response to a criticism of his unified-field-theory work[58], submitting to *Physical Review* again, since he objected to his paper being refereed.

Not only Einstein was confused. The confusion was not completely resolved until 1959[59], though momentum had been building since Feynman (using the pseudonym Mr. Smith) described a thought experiment at a conference in 1957. As this piece is not a review but is intended merely to call attention to some historical consensus shifts, I point readers to the excellent historical overview by Mauro *et al.*[60], which everyone interested in the topic should read. See also section 4 in the article by Trimble[61], which contains several references related to this story. The main ingredients in the confusion were the issue of coordinate *vs.* real effects (comparable to the difference between the singularity at the horizon of a black hole, a mere coordinate singularity in Schwarzschild coordinates, and the real (within the context of GR) singularity at the centre) and doubt on the part of influential figures such as Einstein and Eddington. Essentially all remaining sceptics were convinced by the agreement of theory and observation regarding energy loss *via* gravitational waves in the famous binary pulsar[62], which can be seen as the first indirect detection of gravitational waves, which in the meantime have now been detected directly[63]. In this case, it was neither an honest mistake nor hanging on to an outmoded idea for irrational reasons (though the opinions of Einstein and Eddington probably played a small role), but rather it was a topic which had not been investigated in enough detail by enough people.

Black holes

Einstein also believed that black holes cannot form, at least not *via* astrophysical processes[64], primarily because he imagined the formation *via* a series of stationary states and argued that collapse would be prevented because particles would have to exceed the velocity of light. (He considered only circular orbits, but believed the result to be more general.) Penrose[65] demonstrated that, with some relatively uncontroversial assumptions, GR can produce singularities provided that gravity is strong enough (to create a so-called trapped region). (A related theorem by Hawking[66–68] states that, with similar assumptions, the Universe must have begun with a singularity, the Big Bang.) Penrose received the 2020 physics Nobel Prize for this work, a three-page paper in which the statement "The space-time manifold is incomplete" is expounded upon in a footnote: "The 'I'm all right, Jack' philosophy with regard to the singularities would be included under this heading!" While there was never a clear consensus that black holes cannot exist, evidence for bodies which are too compact to be anything else, such as the compact objects in X-ray binaries and the massive objects at galactic centres, has led over time to a consensus that they do. Similar to the case of gravitational waves, scepticism about black holes was due primarily to the topic not having been investigated fully. To some extent, the dislike of true (as opposed to merely coordinate) singularities might have made some sceptical of black holes in general. Today it is clear that such singularities follow from GR but also suspected by most that GR must be modified under such extreme conditions in order to be compatible with quantum theory.

Cosmological horizons

There was real confusion regarding the concept of cosmological horizons, perhaps in part due to the fact that in the early days the de Sitter model (a flat universe with no matter and a cosmological constant which expands exponentially; $\lambda_0 = 1$ and $\Omega_0 = 0$) was often used as a fiducial model and in that model the event horizon corresponds to the Hubble radius (at which the recession velocity is the speed of light) and is also constant in time, while in general neither is the case. For modern expositions of the topic of cosmological horizons, see Chapter 21 of the textbook by Harrison [17] and the papers by Davis & Lineweaver [69] and van Oirschot *et al.* [70].

The long-standing confusion was cleared up in a famous paper by Rindler [71], but, as Rindler himself recalls [72], it was sparked by further confusion documented in a debate in these pages. Whitrow [73] had claimed that the existence of (what was later to be known as) an event horizon in the Steady State model was an argument against that model. (Of course, we now know that the Steady State model is incorrect, but also that the existence of an event horizon is not *per se* an argument against it nor against any other model.) Bondi & Gold [74] responded, concluding with

> We do not know what hurricanes will be directed against our cosmological edifice; but we are a little aggrieved to think that it is being credited with so little structural strength that Dr. Whitrow's puff could make it shudder.

Whitrow [75] was not amused, and clarified his claim. He attempted to clear it up [76] but appears to me to have been more confused; Hoyle [77] agrees: "When I wished to refer to an observer with an idealized telescope I referred to an observer with an idealized telescope; and when I did not wish to refer to an idealized telescope I did not refer to an idealized telescope." Annoyed, Whitrow told Rindler to "do something" [72], which ultimately led to Rindler's definitive paper [71]. To some extent the confusion was similar to that regarding gravitational waves and black holes, but on the other hand, as Whitrow's arguments demonstrate, also due to trusting one's intuition in a regime for which it is not suited.

Rindler noted that there are two types of cosmological horizons. One, dubbed the particle horizon, is, at a given time, the greatest (proper) distance (measured at the current time) from which a signal could have reached us. By symmetry, the corresponding sphere represents the furthest distance to which a signal emitted from our position could have reached. One can think of an expanding sphere of light emitted at the earliest possible time; its edge is the particle horizon with respect to the centre. While of course light travels *locally* at the speed of light, the expansion history of the universe affects how the distance to the horizon changes with time. The other, dubbed the event horizon, represents the greatest distance from which a signal could travel to our position in the future. By symmetry, that corresponds to the greatest distance which a signal sent from our position could ever reach. Big-bang models have a finite particle horizon; models which asymptotically expand exponentially have a finite event horizon. There are thus models with a particle horizon, an event horizon, both, or neither (*e.g.,* Einstein's original static model, and the relativistic equivalent of the Milne model with $\lambda_0 = 0$ and $\Omega_0 = 0$). The particle horizon is often thought of as the spatial surface at a given time (such as the present) while the event horizon is often thought of as a surface in spacetime, though both can be thought of in both ways. (That is probably because two questions are considered more important than others: what are the furthest objects which can be seen now, and which events (in the sense of relativity) can affect us in the future.) In the spacetime sense, the particle horizon is the backward lightcone now and the event horizon the greatest possible extent of the backward lightcone in the future. Note that the particle horizon always increases in terms of proper distance and that the event horizon in the case of pure exponential expansion is at a fixed proper distance (it also corresponds to the Hubble sphere in that case, which is not true in general). See the paper by Rindler [71] and Chapter 21 in the book by Harrison [17] for more details.

On-going: the cosmological-constant problem

The cosmological-constant problem refers to the question as to why Λ is so small compared to the expectation from quantum field theory. That vacuum energy can produce a gravitational effect was noted already by Nernst [78]. Depending on definitions, the discrepancy can be as high as 120 orders of magnitude (*e.g.,* refs. 79,80) and has even been called "the worst theoretical prediction in the history of physics" [81]. Interestingly, most seem to believe the prediction from quantum field theory and postulate some mechanism which cancels the large expected value — or, since we now know that $\Lambda > 0$ though much smaller than that estimate, *almost* cancels it (which is perhaps conceptually more difficult, as it would seem to involve fine-tuning, whereas an exact cancellation could be due to some unknown symmetry principle). Of course, in GR there is no concept of any sort of vacuum energy from quantum field theory; Λ is simply a constant like the gravitational constant G (and indeed can appear in purely Newtonian theory). Schrödinger [82] seems to have been the first to suggest that a fluid with equation of state $p = -\rho$ would behave like a cosmological constant. Einstein [83] acknowledged that, but didn't see the point, preferring to keep the term on the 'geometric' rather than the 'matter' side of his field equation.

But if one, why not both? Weinberg [84] postulated that the prediction from quantum theory is correct, but that there is also a 'bare' or 'geometrical' cosmological constant which is negative and slightly smaller in absolute value than the 'matter' one, so that the net result is an effective cosmological constant which is slightly positive, the fine-tuning being explained by an appeal to the weak Anthropic Principle: an effective cosmological constant which is too negative would have led to a universe which collapsed after a short time, while a too positive one would have led to a universe which expanded too quickly for structure to form. Despite the influence of Weinberg, most still believe that there is a real cosmological-constant problem, perhaps out of distaste for arguments based on the Anthropic Principle.*

Bianchi & Rovelli[94] (in an expanded version of another article[95]) take a different point of view and dispute the reasoning leading to the prediction of an extremely large cosmological constant from quantum field theory. Even if one disagrees with them on that point, they still present several other interesting arguments, largely independent of that specific argument. They point out that what Einstein later thought of the cosmological constant is irrelevant and that the 'coincidence problem' (the fact that the energy densities due to matter and the cosmological constant are approximately equal at the current epoch, even though they have a different dependency on time) is overstated, giving three arguments against it. They also support the idea that a spatially closed Universe makes more sense (see below) — all in all, one of the most sensible papers in modern cosmology. Their argument against the conventional derivation of something like a cosmological constant is too detailed to be summarized here, but their conclusion to the corresponding section is worth quoting:

> To trust *flat-space* QFT telling us something about the nature of a term in Einstein equations which implies that spacetime cannot be flat, is a delicate and possibly misleading step. To argue that a term in Einstein's equation is "problematic" because flat-space QFT predicts it, but predicts it wrong, seems a *non sequitur* to us. It is saying that a simple explanation is false because an ill-founded alternative explanation gives a wrong answer.

as is part of their general conclusion:

> But to claim that dark energy represents a profound mystery is, in our opinion, nonsense. "Dark energy" is just a catch name for the observed acceleration of the universe, which is a phenomenon well described by currently accepted theories, and predicted by these theories, whose intensity is determined by a fundamental constant, now being measured. The measure of the acceleration only determines the value of a constant that was not previously measured. We have only discovered that a constant that so far (strangely) appeared to be vanishing, in fact is not vanishing. Our universe is full of mystery, but there is no mystery here.

The cosmological-constant problem is similar to the flatness problem (see the penultimate section) in that most people who believe that it exists have probably not looked into in detail themselves, but have absorbed it from textbooks (which usually just mention it rather than elucidate it) and so on as part of the 'lore'. It is thus good that Rovelli mentions the argument above in his recent book[96]

*Straumann[85] mentions (without citation) that Zel'dovich made a similar claim in 1967, though without explicitly invoking the Anthropic Principle. However, I haven't been able to find it explicitly mentioned in the literature, though it is perhaps implicit in a paper by Sakharov[86] on the idea of explaining the cosmological constant *via* vacuum fluctuations, something also explored by Zel'dovich[87,88] (see (also) the corresponding English translations and commented reprints[89–93]).

(reviewed in these pages[97]), even though it concentrates on just the essentials; of course, avoiding misinformation should be considered essential.

There are two constants of nature in the Einstein equation, Λ and G. Why do many see it as a puzzle that we don't understand the ultimate origin of Λ, nor its value, while the same questions are rarely asked about G? Of course, even if there is no contribution to an effective cosmological constant from vacuum energy, its value could still be determined by the Anthropic Principle. For that matter, the value of G could also be determined by the Anthropic Principle. I am not aware of any calculations of their values from deeper principles. Thus, within the context of GR, the value of Λ is no more puzzling than the value of G; the problem arises when trying to explain the small observed value despite the prediction of a large value from quantum mechanics.

For those who believe in the large value of Λ predicted by quantum mechanics and also willing to accept Weinberg's argument, then the problem is solved. However, due to its reliance on the Multiverse and the Anthropic Principle, not all are willing to accept it. Unless the prediction from quantum mechanics is fundamentally wrong for some unknown reason, then the problem still exists.

A detour into particle physics

For many, the belief that $\Lambda = 0$ exactly was based on the idea that, since the particle-physics expectation was that it should have an extremely large value, some unknown mechanism must cancel that contribution, and would cancel it exactly, because having Λ small but non-zero would imply a fine-tuned cancellation mechanism. (As mentioned above, Weinberg explained the small value by invoking the Multiverse and the Anthropic Principle, though not all accept his explanation.) Similarly, the observed small mass of the Higgs boson could be explained by a cancellation mechanism involving supersymmetry. In the former case, a small but non-zero value for $\Lambda = 0$ is observed, so we now know that no such cancellation mechanism exists (not that there had been a convincing argument for one). In the latter case, (at least most theories of) supersymmetry predicted that the *Large Hadron Collider* should have detected at least some of the predicted supersymmetric partners of the known particles, but that didn't happen. Thus, what appeared to many to be a convincing cancellation mechanism was ruled out, but since the mass of the Higgs boson is small, that small mass is now a puzzle. The two cases are similar in that the purported cancellation mechanism has been ruled out, but in the former case *via* the observed absence of cancellation and in the latter case *via* falsified predictions of the theory, leaving the observed small Higgs mass without an explanation. In both cases, vague belief in naturalness motivated both expectations, though supersymmetry at least had some additional motivation.

Not all refinements are paradigm changes

Of course, not all modifications of a working model indicate some sort of important change. The current cosmological standard model with $\Omega_0 \approx 0.3$ and $\lambda_0 \approx 0.7$ is at least approximately spatially flat (since $\approx 0.3 + \approx 0.7 = \approx 1$), but, contrary to what one sometimes hears, that flatness is not an important ingredient, for several reasons. First, the detection of non-zero spatial curvature would not involve any new physics. Second, in contrast to the discovery of the acceleration of the Universe or (were it to happen) the discovery that 'dark energy' is something other than the cosmological constant, there would be no qualitative, and only a small quantitative, change in our understanding of the Universe. Third, assuming flatness (as is sometimes done) is a practical matter: if a certain cosmological test is relatively insensitive to spatial curvature, then it makes sense to assume it as a prior in the analysis, since the observational evidence from other tests that the Universe is at least approximately flat is good; even if the test is sensitive to spatial curvature, in practice it probably cannot distinguish between perfect flatness and flatness to a very good approximation, though if there is a chance of detecting spatial curvature then one should make use of it. (Of course, perfect flatness can never be proved observationally, merely ruled out.) Fourth, even inflation does not predict a perfectly flat Universe, only one with a large radius of curvature. (Though, to be sure, it is not really clear what inflation robustly predicts; when a model with $\Omega_0 \approx 0.3$ and $\lambda_0 = 0$ was still viable, there were many papers on 'open inflation', though to be fair such models were probably more contrived than the typical inflation model. While they did lend some credence to the 'inflation can predict anything' argument, it is fair to say that such models were not part of mainstream inflation and that most believed that (near) flatness was a robust prediction. The first convincing 'detection' of a positive cosmological constant was belief in the inflation-inspired flat universe together with observational evidence that Ω_0 is substantially less than 1[39].) Fifth, the origin of a universe with positive spatial curvature, and hence spatially finite, described approximately by a Friedmann–Robertson–Walker model, is perhaps easier to understand and/or more likely to be true because a spatially infinite FRW model is spatially infinite even at the Big Bang, hence its origin is more difficult to understand with respect to causality.

The flatness problem

Another on-going debate concerns the flatness problem. Two important aspects of the flatness problem are the fine-tuning problem (why was Ω so close to 1 in the early Universe?) and the time-scale problem (why is Ω so close to 1 today?). Briefly, there is no fine-tuning problem since all non-empty FRW models with a Big Bang have Ω arbitrarily close to 1 at the Big Bang, and there is no instability problem because arguments claiming that there is are based on a false analogy[98,99]. As many readers of this *Magazine* are aware, for the past several years I've been trying to convince the community that the flatness problem, as it is usually understood*, is based on a misunderstanding (*e.g.,* ref.

*By that I mean the flatness problem as applied to a Friedmann model containing dust, a cosmological constant, both, or neither, which is usually what is assumed in the literature actually discussing the flatness problem, but perhaps not by those who mention it in passing or just think about it. Certainly when a specific time is introduced, such as setting initial conditions at a specific temperature, at least one aspect of the flatness problem is changed and some of the arguments against it no longer apply.

41 and references therein). Apart from obviously wrong responses†, one point of criticism, probably just thought more often than spoken as well, is that I must be wrong because I don't agree with the consensus in the community. In this particular case, I don't think that there is a consensus as much as the flatness problem having become part of the lore. Very similar descriptions appear in dozens of textbooks (which almost always ignore the technical literature on the flatness problem but not on other topics mentioned, even if only briefly); like typos in reference lists, one can almost track who has copied from whom without actually checking very much.

Three other misconceptions are common. One is that denying the existence of the flatness problem would be some sort of paradigm shift as discussed above. However, that wouldn't change any observational conclusions. In the past, it could have weakened the belief in the flatness of our Universe, but that is now an observational fact ($\Omega_0 + \lambda_0 = 1$ to within a per cent or so). It could, and should, though, urge people to consider that non-zero spatial curvature might be detectable, unless there is a *robust* prediction (probably from inflation) that any such curvature would be undetectable. (Despite 'open inflation' models which were considered at some point, I think it is fair to say that near flatness is a robust prediction of inflation, but one needs to know exactly how near before one gives up the possibility of detecting spatial curvature.) Another is that it doesn't matter because we know that inflation happened. Without using near-flatness itself as evidence, while there is some evidence to support inflation, it is not yet a proven theory to the same degree as, day, the Big Bang. A third is that it doesn't matter whether there is a flatness problem in classical cosmology because there is in quantum cosmology and the early Universe must certainly be described by some theory which takes quantum mechanics into account. It *does* matter for the understanding of classical cosmology.

With regard to the possibility or usefulness of observationally trying to detect curvature, to some extent, one's expectations play a role. If one sees curvature as an additional parameter, one can ask whether a better fit to observations justifies allowing an additional parameter: since an additional parameter can be used to make the fit better, the advantage must be large enough to outweigh such a 'disadvantage' (*e.g.,* ref. 100). On the other hand, one's default expectation could be the general case and one would require evidence for vanishing curvature.

Conclusions

Throughout the history of cosmology, there have been phenomena for which there was a consensus with regard to their explanation and which was overturned in a non-trivial way. (Of course, refinement is part and parcel of science and

† "You just can't be right because I work on inflation" is something I have actually heard. Even after I had explained that inflation either happened or not independently of whether there is a flatness problem to solve, I still couldn't make any headway. It does seem, though, that many who have actually thought about it in detail agree with me. Whether we are right remains to be seen; I'm the first to admit that I haven't convinced everyone. One thing that we can agree on is that there are different opinions in the literature, with several well known cosmologists arguing against the traditional understanding of the flatness problem (*e.g.,* ref. 41 and references therein).

most advances do not involve any sort of revolution[101].) Nevertheless, if the *only* reason why one believes something is that there seems to be a consensus, one should be sceptical, because sometimes such consensuses exist despite the facts, having taken on lives of their own in the manner of urban legends. As discussed above, some of those have been replaced with a new consensus on what is probably the correct explanation. Current examples which I think might be overturned in the future are the flatness problem and the cosmological-constant problem; it will be interesting to see for how long they will still be regarded as problems. Interestingly, the consensus in those cases is in regard to the problems themselves, rather than their solutions, particular with respect to the cosmological-constant problem, for which there is no consensus regarding its solution. While one could see inflation as a consensus solution for the flatness problem, there are other arguments for inflation, and even universal belief in the non-existence of any sort of flatness problem would not be sufficient to rule out inflation. Understandably, most debate is not in fields where there is a perceived consensus, but rather in fields where it is clear that there is not a consensus, *e.g.,* regarding the existence of the Multiverse[102–104], fine-tuning[105–108], or the Anthropic Principle[109–113] (note that various combinations of those topics are often discussed together; they are not necessarily related, but can be). To some extent, the fact that cosmology is now a data-driven science, in contrast to the old days when there were only two-and-one-half or nine facts[114], has caused some to lose sight of such important debates (assuming they ever had them in sight at all). We need the data-driven science, but also discussion about fundamental principles.

Acknowledgements

I thank the anonymous referees for helpful comments which have greatly improved this work. This research has made use of NASA's Astrophysics Data System Bibliographic Services.

References

(1) K. Popper, *The Logic of Scientific Discovery* (Basic Books), 1959.
(2) A. S. Eddington, *The Nature of the Physical World* (Cambridge Univ. Press), 1928.
(3) A. S. Eddington, *New Pathways in Science: Messenger Lectures 1934* (Cambridge Univ. Press), 1935.
(4) C. Sagan, *Broca's brain: Reflections on the Romance of Science* (Random House), 1979.
(5) G. O'Toole, 'Extraordinary claims require extraordinary evidence', `https://quoteinvestigator.com/2021/12/05/extraordinary/`, 2021.
(6) E. R. Harrison, *Nat.*, **204**, 271, 1964.
(7) E. R. Harrison, *MNRAS*, **131**, 1, 1965.
(8) E. R. Harrison, *Phys. Today*, **27**, 30, 1974.
(9) E. R. Harrison, *Amer. J. Phys.*, **45**, 119, 1977.
(10) E. R. Harrison, *Mercury*, **9**, 83, 1980.

(11) E. R. Harrison, in A. G. W. Cameron (ed.), *Astrophysics Today, Readings from Physics Today* (American Institute of Physics), 1984, p. 296.
(12) E. R. Harrison, *Sci.*, **226**, 941, 1984.
(13) E. R. Harrison, *Nat.*, **322**, 417, 1986.
(14) E. R. Harrison, in S. Bowyer & C. Leinert (eds.), *The Galactic and Extragalactic Background Radiation. Proceedings of the 138th Symposium of the International Astronomical Union, held in Heidelberg, FRG, June 12–16, 1989* (Kluwer Academic Publishers), 1990, p. 3.
(15) E. R. Harrison, in B. Bertotti, R. Balbinot & S. Bergia (eds.), *Modern Cosmology in Retrospect* (Cambridge Univ. Press), 1990, p. 33.
(16) E. R. Harrison, *Phys. Rev. D*, **1**, 2726, 1970.
(17) E. R. Harrison, *Cosmology, the Science of the Universe*, 2nd edn. (Cambridge Univ. Press), 2000.
(18) E. R. Harrison, *Darkness at Night: A Riddle of the Universe* (Harvard Univ. Press), 1987.
(19) P. S. Wesson, K. Valle & R. Stabell, *ApJ*, **317**, 601, 1987.
(20) P. S. Wesson, *ApJ*, **367**, 399, 1991.
(21) E. A. Poe, *Eureka: A Prose Poem* (Putnam), 1848.
(22) C. V. L. Charlier, *Medd. Lund. Astron. Observ. Ser. I*, **38**, 1, 1908.
(23) C. V. L. Charlier, *Medd. Lund. Astron. Observ. Ser. I*, **98**, 1, 1922.
(24) J. H. Lambert, *Cosmologische Briefe über die Einrichtung des Weltbaues* (Klett), 1761.
(25) G. Gamow, *Sci. Am.*, **195**, 136, 1956.
(26) G. Gamow, *My World Line* (Viking Press), 1970.
(27) C. O'Raifeartaigh & S. Mitton, *Physics in Perspective*, **20**, 318, 2018.
(28) C. O'Raifeartaigh *et al.*, *Eur. Phys. J. H*, **43**, 73, 2018.
(29) S. Weinberg, *Physics Today*, **58**, 31, 2005.
(30) J. D. Barrow, in K. Chamcham, J. Silk, J. D. Barrow & S. Saunders (eds.), *The Philosophy of Cosmology* (Cambridge Univ. Press), 2017, p. 83.
(31) A. Einstein & W. de Sitter, *Proc. Natl. Acad. Sci. USA*, **18**, 213, 1932.
(32) P. Helbig, *The Observatory*, **141**, 117, 2021.
(33) S. M. Carroll, W. H. Press & E. L. Turner, *ARA&A*, **30**, 499, 1992.
(34) A. G. Riess *et al.*, *AJ*, **116**, 1009, 1998.
(35) S. Perlmutter *et al.*, *ApJ*, **517**, 565, 1999.
(36) J. P. Ostriker & P. J. Steinhardt, *Nat.*, **377**, 600, 1995.
(37) L. M. Krauss & M. S. Turner, *Gen. Rel. Grav.*, **27**, 1137, 1995.
(38) L. M. Krauss, *ApJ*, **501**, 461, 1998.
(39) P. J. E. Peebles, *ApJ*, **284**, 439, 1984.
(40) M. J. Rees, *The Observatory*, **89**, 193, 1969.
(41) P. Helbig, *Eur. Phys. J. H*, **46**, 10, 2021.
(42) H. Bondi & T. Gold, *MNRAS*, **108**, 252, 1948.
(43) F. Hoyle, *MNRAS*, **108**, 372, 48.
(44) S. Castello, S. Ili c & M. Kunz, *Phys. Rev. D*, **104**, 023522, 2021.
(45) S. Castello, in E. Augé, J. Dumarchez & J. T. T. Van (eds.), *Proceedings of the 56th Rencontres de Moriond,2022 Cosmology session* (ARISF), 2022, p. 79.

(46) D. Overbye, *Lonely Hearts of the Cosmos* (HarperCollins), 1991.
(47) J. R. Gott III *et al.*, *ApJ*, **194**, 543, 1974.
(48) A. R. Sandage, 'Interview of Allan Sandage by Bert Shapiro on 1977 February 8, Niels Bohr Library & Archives, American Institute of Physics, College Park, MD USA', online, 1977, `https://www.aip.org/history-programs/niels-bohr-library/oral-histories/32867`.
(49) A. R. Sandage, in B. Binggeli & R. Buser (eds.), *The Deep Universe*, vol. 23 of *Saas-Fee Advanced Course* (Springer), 1995, p. 1.
(50) R. G. Carlberg, in J. T. Thanh, Y. Giraud-Heraud, F. Bouchet, T. Damour & Y. Mellier (eds.), *Fundamental Parameters in Cosmology: Proceedings of the XXXIIIrd Rencontres de Moriond* (Éditions Frontiers), 1998, p. 423.
(51) A. R. Sandage, *Physics Today*, **23**, 34, 1970.
(52) E. D. Loh & E. J. Spillar, *ApJ*, **307**, L1, 1986.
(53) A. Dekel *et al.*, *ApJ*, **412**, 1, 1993.
(54) A. F. Heavens & A. N. Taylor, *MNRAS*, **275**, 483, 1995.
(55) A. Einstein, *Sitzungsb. Kön. Pr. Akad. Wiss.*, 154–167, 1918.
(56) A. Einsten & N. Rosen, *Journal of the Franklin Institute*, **223**, 43, 1936.
(57) D. Kennefick, *Phys. Today*, **58**, 43, 2005.
(58) A. Einstein, *Phys. Rev.*, **89**, 321, 1952.
(59) H. Bondi, F. A. E. Pirani & I. Robinson, *Proc. Roy. Soc. Lond. A*, **251**, 519, 1959.
(60) M. D. Mauro, S. Esposito & A. Naddeo, in R. Ruffini & G. Vereshchagin (eds.), *The Sixteenth Marcel Grossmann Meeting* (World Scientific), 2022.
(61) V. Trimble, *Eur. Phys. J. H*, **42**, 261, 2017.
(62) R. A. Hulse & J. H. Taylor, *ApJ*, **195**, L51, 1975.
(63) B. P. Abbott et. al. (LIGO Scientific Collaboration and Virgo Collaboration), *Phys. Rev. Lett.*, **116**, 166102, 2016.
(64) A. Einstein, *Ann. Math.*, **40**, 922, 1939.
(65) R. Penrose, *Phys. Rev. Lett.*, **14**, 57, 1965.
(66) S. W. Hawking, *Proc. Roy. Soc. Lond. A*, **294**, 511, 1966.
(67) S. W. Hawking, *Proc. Roy. Soc. Lond. A*, **295**, 490, 1966.
(68) S. W. Hawking, *Proc. Roy. Soc. Lond. A*, **300**, 187, 1967.
(69) T. M. Davis & C. H. Lineweaver, *PASA*, **21**, 97, 2004.
(70) P. van Oirschot, J. Kwan & G. F. Lewis, *MNRAS*, **404**, 1633, 2010.
(71) W. Rindler, *MNRAS*, **116**, 662, 1956.
(72) W. Rindler, 'Roundtable discussion: recollections of the astrophysics revolution', `https://youtube.com/watch?v=iH8btReqv4c`, 2013, special event at the 27th Texas Symposium on Relativistic Astrophysics, Dallas, Texas.
(73) G. J. Whitrow, *The Observatory*, **73**, 205, 1953.
(74) H. Bondi & T. Gold, *The Observatory*, **74**, 36, 1954.
(75) G. J. Whitrow, *The Observatory*, **74**, 37, 1953.
(76) G. J. Whitrow, *The Observatory*, **74**, 173, 1954.
(77) F. Hoyle, *The Observatory*, **74**, 253, 1954.
(78) W. Nernst, *Verhandlungen der Deutschen Physikalischen Gesellschaft*,

18, 83, 1916.
(79) S. Weinberg, *Rev. Mod. Phys.*, **61**, 1, 1989.
(80) R. J. Adler, B. Casey & O. C. Jacob, *Amer. J. Phys.*, **63**, 620, 1995.
(81) M. P. Hobson, G. P. Efstathiou & A. N. Lasenby, *General Relativity: An Introduction for Physicists* (Cambridge Univ. Press), 2006.
(82) E. Schrödinger, *Physikalische Zeitschrift*, **19**, 20, 1918.
(83) A. Einstein, *Physikalische Zeitschrift*, **19**, 165, 1918.
(84) S. Weinberg, *Phys. Rev. Lett.*, **59**, 2607, 1987.
(85) N. Straumann, *Eur. J. Phys.*, **20**, 419, 1999.
(86) A. D. Sakharov, *Dokl. Akad. Nauk SSSR Ser.Fiz.*, **177**, 70, 1967.
(87) Y. B. Zel'dovich, *ZhETF Pis'ma*, **6**, 883, 1967.
(88) Y. B. Zel'dovich, *Uspekhi Fizicheskikh Nauk*, **95**, 209, 1968.
(89) A. D. Sakharov, *Sov. Phys. Dokl.*, **12**, 1040, 1968.
(90) A. D. Sakharov, *Gen. Rel. Grav.*, **32**, 365, 2000.
(91) Y. B. Zel'dovich, *JETP Lett.*, **6**, 316, 1967.
(92) Y. B. Zel'dovich, *Soviet Physics Uspekhi*, **11**, 381, 1968.
(93) Y. B. Zel'dovich, *Gen. Rel. Grav.*, **40**, 1557, 2008.
(94) E. Bianchi & C. Rovelli, 'Why all these prejudices against a constant?', arXiv:1002.3966, 2010.
(95) E. Bianchi & C. Rovelli, *Nat.*, **466**, 321, 2010.
(96) C. Rovelli, *General Relativity: The Essentials* (Cambridge Univ. Press), 2021.
(97) P. Helbig, *The Observatory*, **142**, 70, 2022.
(98) P. Helbig, *MNRAS*, **421**, 561, 2012.
(99) M. Holman, *Found. Phys.*, **48**, 1617, 2018.
(100) A. R. Liddle, *MNRAS*, **351**, L49, 2004.
(101) I. Asimov, *Skept. Inq.*, **14**, 35, 1989.
(102) B. J. Carr (ed.), *Universe or Multiverse?* (Cambridge Univ. Press), 2007.
(103) S. Friederich, *Multiverse Theories: A Philosophical Perspective* (Cambridge Univ. Press), 2021.
(104) P. Helbig, *The Observatory*, **141**, 267, 2021.
(105) M. J. Rees, *Astrophys. Space Sci.*, **285**, 375, 2003.
(106) G. F. Lewis & L. A. Barnes, *A Fortunate Universe: Life in a Finely Tuned Cosmos* (Cambridge Univ. Press), 2017.
(107) P. Helbig, *The Observatory*, **137**, 243, 2017.
(108) F. C. Adams, *Phys. Rep.*, **807**, 1, 2019.
(109) B. Carter, in M. S. Longair (ed.), *Confrontation of Cosmological Theories With Observational Data* (Reidel Publishing Co.), 1974, p. 291.
(110) G. F. R. Ellis, *Gen. Rel. Grav.*, **43**, 3213, 2011.
(111) M. J. Rees & B. J. Carr, *Nat.*, **278**, 605, 1979.
(112) J. D. Barrow & F. J. Tipler, *The Anthropic Cosmological Principle* (Oxford Univ. Press), 1988.
(113) B. R.-W. Williams, *Because we are here: a new approach to the history of the anthropic principle*, Master's thesis, Iowa State University, Ames, Iowa, 2007, `https://lib.dr.iastate.edu/rtd/15019`.
(114) M. Longair, *QJRAS*, **34**, 157, 1993.