

Anthropic arguments outside of cosmology and string theory

Milan M. Ćirković¹

*Astronomical Observatory of Belgrade, Volgina 7,
11000 Belgrade, Serbia*

Abstract. Anthropic reasoning has lately been strongly associated with the string theory landscape and some theories of particle cosmology, such as cosmological inflation. The association is not, contrary to multiple statements by physicists and philosophers alike, necessary. On the contrary, there are clear reasons and instances in which the anthropic reasoning is useful in a diverse range of fields such as planetary sciences, geophysics, future studies, risk analysis, origin of life studies, evolutionary theory, astrobiology and SETI studies, ecology, or even strategic studies and global policy. The association of anthropic reasoning with string theory and particle cosmology has not only become the standard wisdom, but has been often construed in a negative way, in order to demonstrate or insinuate that such reasoning is too abstract or even belongs to “fringe science”, remote from run-of-the-mill research practices in any other more “mundane” and less theoretical scientific discipline. The purpose of this paper is to (i) analyse some of the counter-examples to the standard wisdom, (ii) suggest that the anthropic reasoning is more flexible, more general, and less fashion-driven than the detractors state. In addition, we consider some historical and/or extrascientific motivation for this persistent prejudice.

Keywords: history and philosophy of physics – epistemology – anthropic reasoning – planetary science – risk analysis – bioethics

1. Introduction: Why anthropic reasoning?

Anthropic reasoning, understood in a loose enough sense, has been present in the history of ideas since ancient times.² In modern science, it appears in Boltzmann’s famous 1895 retort to Zermelo on the topic of the origin of the Second Law of Thermodynamics, or the temporal asymmetry of the universe (Boltzmann 1895, p. 415):

If we assume the universe great enough, we can make the probability of one relatively small part being in any given state (however far from the state of thermal equilibrium), as great as we please. We can also make the probability great that, though the whole universe is in thermal equilibrium,

¹ Also at Future of Humanity Institute, Faculty of Philosophy, University of Oxford, Suite 8, Littlegate House, 16/17 St Ebbe's Street, Oxford, OX1 1PT, UK. E-mail: mcirkovic@aob.rs.

² For some specific instances, see Ćirković (2004).



our world is in its present state. It may be said that the world is so far from thermal equilibrium that we cannot imagine the improbability of such a state. But can we imagine, on the other side, how small a part of the whole universe this world is? Assuming the universe great enough, the probability that such a small part of it as our world should be in its present state, is no longer small.

If this assumption were correct, our world would return more and more to thermal equilibrium; but because the whole universe is so great, it might be probable that at some future time some other world might deviate as far from thermal equilibrium as our world does at present. Then the aforementioned H-curve would form a representation of what takes place in the universe. The summits of the curve would represent the world where visible motion and life exist.

The issue at hand was whether the thermodynamical asymmetry (popular “arrow of time”) is a local or truly global, universe-wide phenomenon. The idea behind Boltzmann’s – or Boltzmann-Schuetz’s – approach was to explain the local thermodynamical disequilibrium by appealing to the size of the universe and the conditions necessary for our existence as living beings and observers. What can be „more anthropic“ than this? The whole polemic is interesting not only from the point of view of history of science (Steckline 1983), but also since Boltzmann's ideas could be recast in the modern multiverse context (Ćirković 2003). However, it went largely unnoticed for quite a long time, and Boltzmann’s inception of the anthropic reasoning was recognized only about the time of the publication of the comprehensive monograph of Barrow and Tipler (1986).

In the orthodox history of science, it was Dicke's (1961) rejoinder to Dirac's (in)famous “large-number hypothesis” (LNH) which first takes into account explicit condition for existence of observers at present.³ The locution “anthropic principle” comes from the study of Carter (1974), which defined anthropic principles and went some way to impose some order on the nature and quality of anthropic arguments used in cosmology thus far. For most people and sources, Carter’s study is the ultimate origin of anthropic reasoning, although it is already clear from the above (and detailed, although hardly unbiased, historical survey could be found in Barrow and Tipler 1986) that it only followed from previous considerations.

From the start the debate revolved about the infamous problem of fine-tuning of fundamental constants of nature and cosmological parameters. While Dirac argued that some numerical relationships between those (understood as pure numbers, in order to avoid confusion stemming from man-made units of measurement) hold in the absolute sense throughout the history of the universe, Dicke has turned the explanatory issue “upside down” by asking when and how such approximate relationships could be *observed*. For example, we observe stars of age similar – within an order of a magnitude – to the age of the universe itself, since we are observers dependent on stars and stellar evolution for our existence. Either much earlier or much later, it would be impossible that observers like us evolve in the universe, thus there is no surprise that we observe approximate relationship between these timescales. This conclusion of Dicke could be generalized to other fine-tuning “coincidences”, as has been subsequently clarified. In any case, Dicke pioneered an approach which is *evolutionary* in nature, relying on the concept of observership, rather than (at least partly) metaphysical concepts of absolute laws.

So, historically, anthropic reasoning *has* indeed emerged from cosmological considerations (for further historical summary see Ellis 2011). But it went much farther than this circumscribed domain



³ See Dirac (1937, 1961).

in the 40+ years since Carter's seminal work. And there is really no reason why it should not have, since what Carter himself understood well (e.g., Carter 1983, 1993), but what subsequent critics *and* supporters often failed to understand and appreciate, anthropic reasoning deals with *observation-selection effects*, and has no teleological meaning whatsoever. It simply applies whenever the number or other properties of observers come into play and influence particular scientific result or claim. This disteleological nature of the anthropic reasoning has been best explicated in the seminal monograph by Bostrom (2002).

Thus, it follows from the above outlined general argument that in many fields of science in which observation selection is likely to play some role, it is possible to use anthropic reasoning. An obvious example is biological evolution, which is itself the process through which observers have emerged so far.⁴ Obviously, the number and properties of observers at present impose constraints – possibly stringent constraints – on processes such as abiogenesis and subsequent unfolding of evolution. Another relevant domain is computer science, including the fields such as complexity theory, artificial intelligence (AI). About the same time when Dicke replied to Dirac's LNH, the great pioneer of computer science, Norbert Wiener, published his famous *Cybernetics* (Wiener 1961). It contains a beautiful discussion of the observation-selection effects in the problem of the arrow of time:

*Our observations of the stars are through the agency of light, of rays or particles emerging from the observed object and perceived by us. We can perceive incoming light, but can not perceive outgoing light, or at least the perception of outgoing light is not achieved by an experiment as simple and direct as that of incoming light. ... This being the case, we can see those stars radiating to us and to the whole world; while if there are any stars whose evolution is in the reverse direction, they will attract radiation from the whole heavens, and even this attraction from us will not be perceptible in any way, in view of the fact that we already know our own past but not our future. Thus the part of the universe which we see must have its past-future relations, as far as the emission of radiation is concerned, concordant with our own. The very fact that we see a star means that its thermodynamics is like our own... Within any world with which we can communicate, the direction of time is uniform.*⁵

Today, post-Barrow and Tipler and post-Bostrom, it is not difficult to recognize anthropic reasoning in Wiener's discourse. This is not a "Whiggish" interpretation of history – to forestall such criticism – since the issues Wiener discusses are still parts of open research programs and heterodox strategies should be not only permitted, but desirable; there is no established view which could enforce misreading of the historical evidence.

Therefore, in stark contrast to the prevailing belief, anthropic arguments have an important tradition, and a role to play *outside* highly speculative fields of string theory and particle/quantum cosmology. In what follows I review some examples and argue for the role for anthropic arguments and explanatory reasoning much larger than hitherto assumed. The primary goal of this study is to dispense with the prejudice that anthropic reasoning is limited to extremely general and "abstract" or "remote" fields of particle cosmology and string theory. Detractors of anthropic reasoning (and they are legion; e.g., Maddox 1984; Earman 1987; Gould 1987; Wilson 1994; Pagels 1998; Manson

⁴ Without prejudice that this is how all or even majority of observers emerge. For instance, our current fledgling AI efforts might be crowned by success in the future and it is quite possible that most of the observers in the universe are of machine origin (e.g., Dick 2003).

⁵ Wiener (1961), pp. 34-35.



2000; Mosterín 2000, 2005; Klee 2002; Smolin 2007; to mention just a few examples⁶) have traditionally argued that anthropic reasoning is quite limited in scope, and while most of them have not pressed this scope limitation as the main reason for rejecting the anthropic reasoning, the meme has spread around. To show that it is wrong, therefore, and that the anthropic reasoning is much broader in scope as an explanatory strategy, leads to weakening the entire sceptical case. In particular, it includes weakening the sceptical strategy in the domain of fundamental physics, where currently the greatest battles are waged.

Let us consider just a couple of examples of the sceptical discourse. Earman (1987) writes in his well-publicized – and largely outdated – critique of the anthropic reasoning:

But to be legitimate, the anthropic reasoning must be backed by substantive reasons for believing in the required worlds-within-worlds structure... Neither classical general relativity nor quantum mechanics provide any firm grounds for taking worlds-within-worlds models seriously, and while various speculative versions of the new inflationary cosmology may eventually provide such ground, the verdict is at present very much in doubt.⁷

Note that not only is the application of the anthropic reasoning strictly circumscribed in Earman's account, but specific advances in cosmology are required as *prerequisites* for anthropic explanations.

Pagels (1998) goes along:

My own view is that although we have not yet discovered the most basic physical laws, if we do, the possibility of life in a universe governed by those laws will in some sense be written into them. The existence of life is not a selective principle acting on those laws; rather, it is a consequence of them. Whether or not I am right, it is simply premature to invoke the anthropic principle until the origin of the universe is much better understood.

So, we have a clear pattern here: it is alleged that the anthropic reasoning makes at least some sense in cosmology, but is strictly rejected in other fields. This limitation is then used to diminish and denigrate the usage of anthropic reasoning *even in cosmology*. Note that Pagels' argument is at best a special pleading: that the possibility of life is allowed by the most basic laws is a truism, since what else could it be (so long as we are metaphysical naturalists)? At worst, it is an outright straw-man argument, since we would be hard pressed to find anybody suggesting that the existence of life is not a consequence of the fundamental laws of physics. The indefinite article betrays Pagels' crucial omission – what he fails to address is the question how could life be a consequence of the most fundamental laws and not *the* consequence of them. The most fundamental laws – whatever precisely they are – almost certainly permit both universes without life and those with life; since we are obviously in the latter category, it is uncontroversial that our existence acts as a “selective principle” (to use Pagels' charged terminology) not on laws themselves, but on *instantiations of these laws*; not on the equations of dynamics, but on the multiple solutions of those equations. To believe that in contrast to all other theories ever devised in physics these equations will have just a single

⁶ Not to mention the cottage industry of anti-anthropic web sites or blogs, e.g., <http://www.math.columbia.edu/~woit/wordpress>; <http://motls.blogspot.rs/> .

⁷ Here, one should note that what Earman calls “worlds-within-worlds” is what we nowadays call the multiverse.



solution, smacks of unsupported quasireligious faith. If they have multiple solutions then the question 'Why do we observe solution A, rather than solutions B, C, ...?' is entirely legitimate, scientific, pertinent, and unavoidable problem. So, Pagels' criticism is, in the final analysis, based on a *quid pro quo* (cf. Barnes 2012).

Further, Mosterín (2005) in a spectacularly hostile review of the anthropic thinking argues:

The universe (as far as we can know it) is something unique ('uni-verse' and 'uni-que' come from the same root 'uni', one). We can learn a posteriori how the universe is, but it makes no sense to speculate on how it should be on the basis of a priori statistical considerations. This is the reason why John Leslie's (1989) firing squad argument is flawed. He compared our existence to the survival of a sentenced man, because each of the guns in his execution squad misfires. Has someone (God?) tinkered with the guns beforehand? Of course, there have been lots of firing squads and seldom have all the guns misfired. There is a grim statistics of firing squads. But the universe is a unique historical fact. There are no statistics of universes. Besides, the components of the firing squad are people with the intention of shooting, but there are no intentions in the fabric of the universe. At least in usual language, fine tuning implies intentionality and multiplicity of cases.

Here we can follow the reasoning of the detractors a few steps further: since it is not *really* acceptable *even* in cosmology, then we do not need to consider anything further in declaring the anthropic reasoning completely bogus and devoid of sense. Some of the gaping holes in Mosterín's argument are obvious (metaphysical commitment to a single universe, *quid pro quo* reference to the "usual language" in describing fine tunings, etc.), others are somewhat more subtle (commitment to a verificationist account of truth), yet others stem from ignorance or suspicion toward recent developments in science, especially particle and quantum cosmology and string theory. The best recent rebuttals of the claims of Mosterín and comp. are given by Carroll (2006) and Barnes (2012) in a technical manner, and by the great Leonard Susskind in his popular book (Susskind 2006). Part of the arguments Mosterín uses against the anthropic reasoning is from an auto-ironically imprecise paper by Klee (2002); see the comprehensive rebuttal in Walker and Ćirković (2006). However, the detractors gained much traction in science, philosophy, and popular media. For example, anti-anthropic papers are much easier to publish than pro-anthropic ones. Even corrections of obvious *numerical errors* in the anti-anthropic literature are lacking (as testified by the never corrected errors regarding the size and mass of the Milky Way in Klee's paper), while detractors like Martin Gardner or Pagels relished in pointing out even trivially simple mistakes/typos in the encyclopaedic monograph of Barrow and Tipler.

Nevertheless, the criticism often sounds better than it really is, since it invokes abstract and remote issues of cosmology and fundamental physics. A few people feel comfortable discussing "the most basic physical laws" or the issues of cosmological origins, since there is no intuitive grasp of them. Therefore, it pays off for the opponents of the anthropic reasoning to promote the dogma that this reasoning is inextricably linked to fundamental physics and cosmology. So, there is not only rhetorical but a game-theoretical motivation for the anti-anthropic campaign. In the following, I shall use examples from planetary science and risk analysis in order to refute the dogmatic prejudice that the anthropic reasoning is limited to highly abstract and non-intuitive domains of fundamental physics



and cosmology. Emancipation from the dogma leads not only to our increased explanatory freedom, but also to some quite practical consequences, as the examples from the domain of risk analysis will show. Therefore, the whole of this debate has profound ethical ramifications, which is too often conveniently ignored by the mainstream philosophers. I also speculate on the origin of hostile prejudices about anthropic reasoning in the concluding section.

2. Anthropic arguments in planetary sciences and astrobiology

In the domain of planetary sciences which experienced explosive growth in the last couple of decades there is a substantial room for anthropic argument. To a large degree, the development of planetary sciences, both within Solar system, and as applied to the large number of extrasolar planets, has been fuelled by the quest for *habitability* (Horneck and Rettberg 2007; Hanslmeier 2009). According to NASA's *Astrobiology Roadmap* (Des Marais et al. 2008, p. 720):

We must move beyond the circumstances of our own particular origins in order to develop a broader discipline, "Universal Biology." Although this discipline will benefit from an understanding of the origins and limits of terrestrial life, it also requires that we define the environmental conditions and the chemical structures and processes that could support life on other habitable planets. Thus we need to exploit universal laws of physics and chemistry to understand polymer formation, self-organization processes, energy utilization, information transfer, and Darwinian evolution that might lead to the emergence of life in planetary environments other than Earth. ... Some conditions that support chemistry that is sufficiently rich to seed life might be detrimental to self-organization of biological structures. Conversely, conditions that promote the emergence of biological complexity might be unfavorable to organic chemistry. Thus an integrative, interdisciplinary approach is necessary to formulate the principles underlying universal biology. The perspectives gained from understanding these principles will markedly improve our ability to define habitability and recognize biosignatures.

And subsequently (p. 727):

Viewing Earth's ecosystems in the context of astrobiology challenges us to consider how "resilient" life really is on a planetary scale, to develop mathematical representations of stabilizing feedbacks that permit the continuity of ecosystems in the face of rapidly changing physical conditions, and to understand the limits of these stabilizing feedbacks. Ideally this consideration will provide insight into the potential effects of environmental changes that are abrupt as well as those changes that unfold over time scales ranging from seasonal cycles to millions of years.

Therefore, probability of observing specific properties of an environment – *any* environment in the Galaxy! – are linked with the number of observers, whose density is a consequence of complex evolutionary processes which could not be reasonably accounted for in the conventional causal manner. While planetary (to simplify things from the start!⁸) parameters are certain to play the determining role in the degree of habitability, they cannot be said to cause it in proximate sense. The relationship is similar to the one between fundamental physical constants like Planck constant, coupling constants of forces, or masses of quarks and leptons in the Standard Model on one side, and



⁸ Not necessarily rejecting the possibility that there are different and more radical kinds of habitats (asteroids, small icy bodies, molecular clouds, etc.).

properties of the ^{12}C nucleus exhibiting famous fine-tuning discovered by Sir Fred Hoyle (Dunbar et al. 1953; Hoyle 1954; Barrow and Tipler 1986) on the other side. In theory, fine-tunings of energy levels in the ^{12}C nucleus are causally explained by properties (and, consequently, fine-tunings) of fundamental constants. If we accept ontological reductionism there is no other mystical factor there – the ^{12}C nucleus consists of quarks and gluons obeying dynamical laws prescribed by quantum physics and nothing else, so its properties, including “controversial” ones, are reducible to the properties of the constituents. In practice, however, the ^{12}C nucleus is too complex and too strongly-interacting quantum system to be understood in terms of these basic ingredients.⁹ So we cannot track the fine-tuning in the ^{12}C nucleus down to the fine-tuning of fundamental constants. In practice we are treating the fine-tuning in ^{12}C as a distinct phenomenon.

The analogy with other complex systems which are also prerequisites for life as we know it is obvious. In particular, habitable planets are complex outcomes of many lower-level factors of planetary sciences in a manner still more complicated than the ^{12}C nucleus being complex outcome of many lower-level factors of quantum/particle physics. So we again cannot hope to track down apparently fine-tuned high-level properties to a set of basic, low-level planetary properties such as the total mass, chemical composition, distance from the parent star, etc. Hence, the very concept of habitability is inextricably linked to the anthropic reasoning – it summarizes physical prerequisites for particular spacetime density of observers. It is unimportant here that we tend to take this concept as unnecessarily narrow – it is not really required that the life we expect to emerge in habitable locales be exactly like the terrestrial one, or even very similar. We speak about potential for habitability on Jupiter's moon Europa below the thick ice crust, although we can be pretty certain that any European forms of life will have to be very different from anything we have encountered on Earth. So, instead of being anthropocentric, as criticisms directed at astrobiology often charge, the notion of habitability actually represents a scientific attempt to liberate ourselves from anthropocentrism while retaining the advantages of scientific method as developed among terrestrial biologists.

In order to do so, understanding of preconditions leading to life, complex life and observership are necessary – and it is exactly the essence of anthropic reasoning such as contained in the works of Dicke or Bostrom or Barnes (and parallel with the one developed in cosmology and fundamental physics by Carter, Linde, Susskind, Tegmark, Hogan, Carroll, and others). An additional difficulty which makes the application of anthropic reasoning here more difficult lies in the feedback loops created by even very simple lifeforms, in which they influence their physical environment. This is in sharp contrast with the situation in fundamental physics and cosmology, where there is no feedback loop between life and observers on the one side and physical systems under study on the other (spatial and temporal scales are either too small or too large for observers to have meaningful impact). It is well-known, for example, that abiogenesis as currently understood could have only occurred in an anaerobic conditions; the present amount of oxygen in the atmosphere is solely a consequence of the evolution of photosynthetic life forms and their establishing dominance over methanogen and other contemporaries (e.g., Horneck and Rettberg 2007). In turn, the aerobic environment predominant on Earth since the “Great Oxygenation Event”, roughly 2.3 billion years before present, enabled emergence of a huge number of complex organisms, some of which have been the cause of further feedback loops in the environment. While these feedbacks play crucial role in the classic Gaia hypothesis and are still very much in play as explanatory vehicles (a great example

⁹ For modern attempts to do so, see for example Freer and Fynbo (2014).

being recent study of Chopra and Lineweaver 2016), we need not endorse any such general theory in order to perceive how habitability is highly emergent property for whose explanation all naive and greedy reductionist schemes simply fail. In their stead, we need at least something akin to the anthropic reasoning.

A specific model is provided by the “rare Earth” hypothesis of Ward and Brownlee (2000). In a nutshell, it is a probabilistic argument suggesting that, while simple microbial life is probably ubiquitous throughout the Galaxy, complex biospheres, like the terrestrial one, are very rare due to the exceptional combination of many distinct requirements for high habitability. These requirements have become familiar to this day even in popular science accounts: **Circumstellar habitable zone** (a habitable planet needs to be in the very narrow interval of distances from the parent star in order to possess liquid water on surface), **“Rare Moon”** (having a large moon to stabilize the planetary axis is crucial for long-term climate stability), **“Rare Jupiter”** (having a giant planet (‘Jupiter’) at the right distance to deflect much of incoming cometary and asteroidal material enables a sufficiently low level of impact catastrophes), **“Rare nuclides”** (radioactive r-elements – especially ^{238}U and ^{232}Th – need to be present in the planetary interior in sufficient amounts to enable plate tectonics and the functioning of the carbon–silicate cycle), etc. There are other items on the list as well – in this sense, „rare Earth“ hypothesis is an open theoretical system, since everyone is free to add items pertaining to a particular area of relevance and expertise. However, the general reasoning is that all these requirements are mostly independent and *a priori* unlikely, so that their combination is bound to be incredibly rare and probably unique in the Milky Way. Without entering into pros and cons of this important hypothesis (and there are many items on both sides of the ledger!), we need to understand that there is nothing contradictory between the conclusion that complex biospheres are extremely rare and the empirical observation that we are living in a complex biosphere. Contradiction is removed by – lo and behold! – the anthropic reasoning. It is the anthropic reasoning which accounts for the Bayesian probability shift: since we require complex biosphere to exist as observers with an extremely high probability,¹⁰ we need to update our probabilistic beliefs accordingly.

In addition, the rare Earth paradigm assumes something about the *duration of time* necessary for the processes leading to complex biospheres to complete. This is the topic of another empirical success of the anthropic reasoning, namely Carter’s relation, which predicts the probable length of time the Earth will remain a habitable planet from the number of improbable steps that occurred in the evolution of intelligent life on Earth (Carter 1983). If the total habitable lifetime of Earth, both past and future is denoted by T and the time necessary to evolve intelligent observers like us is t_e , Carter argues that it should hold that $T - t_e \approx T / n$, where n is the number of “critical steps” (or “improbable steps” or “key transitions”) in the evolutionary process. Clearly, a relation between these numbers is neither obvious nor trivial – without going into whether it is verified or falsified, *it certainly is an empirical claim*. Yet Carter’s relation is a consequence of the traditional “weak” anthropic principle: self-selection from our nature as intelligent observers which evolved on a habitable planet.



¹⁰ Not exactly necessity of unity probability, since we have to account for a minuscule probability of our being Boltzmann brains or our observations of complex biosphere around us being deceitful.

Finally, one might argue that the first anthropic argument in planetary sciences appeared centuries ago, in the work of French philosopher and naturalist Bernard Le Bouvier de Fontenelle (1657-1757), published in 1686, a single year before the great scientific revolution inaugurated by Newton's *Principia*.¹¹ It deals with habitability of *planet Earth*, and is contained in a single paragraph of his *Conversation on the Plurality of the Worlds*. It reads:¹²

In the next places, the reason why the planes of their [comets'] motions are not in the plane of the ecliptic, or any of the planetary orbits, is extremely evident; for had this been the case, it would have been impossible for the Earth to be out of the way of the comets' tails. Nay, the possibility of an immediate encounter or shock of the body, of a comet would have been too frequent; and considering how great is the velocity of a comet at such a time, the collision of two such bodies must necessarily be destructive of each other; nor perhaps could the inhabitants of planets long survive frequent immersions in the tails of comets, as they would be liable to in such a situation. Not to mention anything of the irregularities and confusion that must happen in the motion of planets and comets, if their orbits were all disposed in the same plane.

Thus, to the question: *why are (observed) orbits of comets highly inclined, in contradistinction to the coplanar planetary orbits?* Fontenelle offers a deceptively simple answer. We would not be here – to contemplate on the peculiarities of cometary trajectories – if these orbits were different (that is, similar to those of planets). In modern rendition, our existence in a stable environment selects some planetary+cometary configuration out of the entire set of all such configuration possible under the dynamical laws. Therefore, there is no reason to believe that the observed planetary+cometary configuration will be typical or average. This passage of Fontenelle had been published 8 years before celebrated Halley's suggestion of December 12, 1694, that comets might collide with planets (published in Halley 1704):

This is spoken to Astronomers: But, what might be the Consequences of so near an Appulse; or of a Contact; or, lastly, of a Shock of the Coelestial Bodies, (which is by no means impossible to come to pass) I leave to be discuss'd by the Studios of Physical Matters.

This famous idea has been followed up by such luminaries as Newton, Wright, Laplace, Lagrange, and others, in the vein of what is usually (and only partially justifiable) called “Biblical catastrophism” of the pre-Cuvier epoch.¹³ Fontenelle wrote the passage 18 years before Newton wondered (in *Opticks*):¹⁴

Whence is it that planets move all one and the same way in orbs concentrick, while comets move all manner of ways in orbs very excentric... blind Fate could never make all the planets move one and the same way in orbs concentrick, some inconsiderable irregularities excepted, which may have risen from the mutual actions of comets and planets upon one another, and which will be apt to increase, till this system wants a reformation. Such a wonderful uniformity in the planetary system must be allowed the effect of choice.

Thus Newton, as a promoter of the Design argument in natural philosophy, failed to understand the power of the Fontenelle's argument, and went deeper into a blind alley (from the modern point of view) of seeking the supernatural Design and/or supernatural regulating mechanism. The same

¹¹ More details in Ćirković (2002).

¹² According to the 1990 translation by H. A. Hargreaves.

¹³ For colourful pieces of its history, see Clube and Napier (1990).

¹⁴ Quoted from the famous fourth edition (Newton 1730).



tension between the apparent design and the explanatory “filtering” through various observation selection effects persists to this day, and is the source of innumerable debates and confusions.¹⁵

And the issue of perception of peculiarity of the inhabited subset of planets in the entire set is a legitimate target of physical inquiry – the one which is directly guided by anthropic reasoning. Again, like in the cosmological case, we have three possible options of explaining the peculiar (non-planar) nature of cometary orbits in our Solar system. The first is to deny the validity and meaningfulness of the question; this is the standard theistic answer which forbids further discussion. In the less direct and allegedly more philosophical manner, we could speak about this state-of-affairs as being a brute fact. The theistic answer can hardly be further discussed, and the explanatory nihilism/brute-factism seems irrational today, since we know that the Solar system forms a tiny part of a much larger whole. As to the origin and properties of this larger whole (the Galaxy) we do have different (and working!) explanations – it would be truly strange to expect that the property of minuscule planetary system like ours cannot, even in principle, be explained.¹⁶

From the other two options, one—causal—entails the idea that there is a law-like reason (presumably to be derived from the future “Theory of Everything” or some other high-level physical theory) for atypical or surprising structure of the early Solar system. In other words, an enormous amount of information necessary for description of the atypical initial conditions can be encoded in some new law(s) of nature and consequent law-like correlations of various matter and vacuum fields. This option is still viable for cosmology, but hardly for planetary cosmogony. *Cosmogonical* initial conditions are not privileged in any way over initial conditions for any other physical process; we do not seek an explanation for (say) the formation of chemical elements or the formation of the ozone layer or for origin of ice ages in a future unified field theory. We seek it much lower down on the epistemological ladder – and higher in terms of complexity of the explanans. Moreover, the very existence of such higher-level theory to which the explanatory work could be outsourced seems highly dubious, again in the light of the necessity for such theory to explain other planetary systems in the Galaxy as well. Some of them – and perhaps most of them – seem to be very different from the Solar System (e.g., Kepler-16b; Doyle et al. 2011).

The third—anthropic—option is to avoid giving a specific description through embedding those conditions into a sufficiently symmetric (“typical”) background. Again, stated in terms of information, the same long description of what we perceive as atypical initial conditions arises—as so often in physics!—from the process of symmetry breaking. The overall description is simple enough, and may be reduced (in the extreme case) to a rule similar to “All possible combinations of initial conditions exist.” That such a high degree of symmetry can indeed completely reproduce the situation in our particular cosmological domain becomes an immediate consequence (cf. Tegmark 1996, 2008; Collier 1996). Observation selection effect then accounts for features of the observed system required for the local observers to exist.

Let us now think of the application of this mode of thinking to the particular cosmogonical issue. We would similarly say “planetary systems with all possible configurations of planetary and cometary orbits exist” as a part of the larger whole (say our Galaxy). Now, we ask: are all such configurations



¹⁵ See also Halley (1726), for example of the curious mixture of teleological and physical thinking characteristic of the early modern era.

¹⁶ I shall return to the issue of explanatory nihilism in the section 4 below.

compatible with our existence (on a planet!)? And the answer, intuitively clear even to Fontenelle, with his rudimentary understanding of preconditions for complex life and intelligence, seems to be negative. Planetary “irregularities and confusions” would have likely resulted in the absence of all observers, contrary to the empirical evidence. There is only a subset of all configurations leading to the emergence of us as intelligent observers, namely the one in which collisions between planets and smaller bodies (comets and asteroids) are not too frequent. Thus, our existence acts as an observational selection effect (cf. Bostrom 2002), or “filter” selecting those sites – in this case planetary systems – where configurations of cometary orbits are in some sense atypical. There is nothing inherently problematic in this methodological approach – or at least nothing *more* problematic than in explaining properties of stellar surveys by the Malmquist bias, or paleontological extinctions by the Signor-Lipps effect, etc. All of these are just manifestations of the observation selection-effects, which need to be studied and corrected for in the context of the anthropic reasoning.

3. Anthropic arguments in risk analysis

Another scientific field which by default deals with variable number of observers is risk analysis – in particular when applied to large catastrophic risks (Bostrom and Ćirković 2008). In the previous section, we have seen that one of the “rare Earth” requirements is that a habitable planet is not subjected to impact and supernova risks more severe than a particular threshold value. Let us now focus on Earth and consider big natural hazards: how do we actually measure the level of such *historical* catastrophic risks as impacts or supervolcanoes?¹⁷ Usual procedure (e.g., Woo 1999) is to fit a distribution to a sufficiently big database of the events of well-defined type in terms of both severity and time. From calculated moments of this distribution one can compute further values of interest, for instance the probability of an earthquake of magnitude larger than S hitting within next T years, or the probability a supereruption emitting more than V cubic kilometers of ash and dust over the same interval of time. The database is in most cases built up from *historical evidence* (or even indirectly, from the fossil record). Empirical predictions are, in general, moments of the distribution function of any particular risk.

The usual procedure does not take into account the tally of observers. Therefore, it is subject to a specific bias called the *anthropic shadow* in Ćirković, Sandberg, and Bostrom (2010): a part of the parameter space of large catastrophic events is not adequately sampled, due to its incompatibility (or weaker compatibility, so to speak) with our existence as observers at present. In other words, the usual procedure, which does not distinguish between the directions of time (the distribution function is assumed to be the same, or similar, or obtained by the same procedure in the future as in the past) fails to account for the selection effects. Consequently, it leads not only to *undersampling* of the total empirical dataset, but to *underestimates* of the predicted risks as well. The past record is unreliable, due to our presence as observers in the present; and it is unreliable in the consistently misleading direction – it makes our environment seem safer than it really is. This is a quintessentially anthropic

¹⁷ The risk from close supernovae/ γ -ray bursts is somewhat different case, since these dramatic explosions will tend to leave few traces in the historical record. While in some cases it was considered an *advantage* (e.g., Schindewolf 1962), the anthropic selection effect I wish to discuss here does not apply to such processes (but see Bishop and Egli 2011).



effect with a wide range of consequences, from risk analysis to epistemology of historical science to ethics and politics of risk.

Anthropic shadow is easiest to understand through a particular example of a specific risk whose magnitude we assess from historical data: terrestrial impact craters. Craters are consequences of impacts of small Solar System bodies (asteroids and comets) upon Earth's surface. Impacts produce physical effects ranging from local and small (most of meteorites, like the famous Murchison meteorite which fell in Australia in 1969), to regional and severely destructive (e.g., the event which created Barringer crater in Arizona about 50,000 years ago), to global and cataclysmic (e.g., the Chicxulub impact, 65 million years ago at the K/Pg boundary). There are 188 confirmed impact craters on Earth at the time of this writing (February 2016).¹⁸ There is, obviously, a relationship between the size of the impactor and the size of the crater created, although the relationship is non-linear and complicated, needing to account for properties such as chemical composition of the impactor, incident angle, etc. (Hughes 2003). But if we do not bother with such complications, we can gauge the severity of an impact catastrophe by the size of an impact crater. So, in the long term, the list of known impact craters will give us the distribution function of impact catastrophes – and by reconstructing it we can predict or retrodict the catastrophe of a given size in a given time interval; for example, we can answer the question how probable is a Chixhulub-size impact within next 1000 years from now. Of course, we need to deal with “classical” selection effects, such as erosion, incompleteness of our crater list, etc.

However, this neglects another key selection effect: the fact that our existence as observers as present does impose limitations on the possible risk distribution functions. Not all distribution functions *compatible with all the historical evidence* are compatible with our existence as observers (and our number) at present. So, those historical trajectories are falsely taken into account – in essence, *subsumed into* – in constructing the distribution function. We need to unravel this biased sampling in order to get to the *real* (or *a posteriori*) distribution function, giving us the *real* magnitude of risk. Surprisingly enough, as shown by Ćirković et al. (2010), the correction leads to underestimating the risk or *overconfidence*. In our example of impact risk, it is clear that there is no chance whatsoever to find a 100km-size crater younger than 1 million years, and that a probability of find such huge crater younger than 10 million years is minuscule. This is because such catastrophe would have caused the extinction of humankind, if it happened, or would have cut the evolutionary chain leading to humans if further in the past. On the other hand, the probability of such crater being formed in the *next* million years is certainly not zero, and the probability of its formation in the next 10 million years is sizeable (e.g., Grieve and Shoemaker 1994). But since the moments of the same distribution function must be the same in both past and the future, we can only infer that the real risk is greater than that calculated by the standard procedure.

¹⁸ See the Earth Impact Database at <http://www.passc.net/EarthImpactDatabase/index.html>.

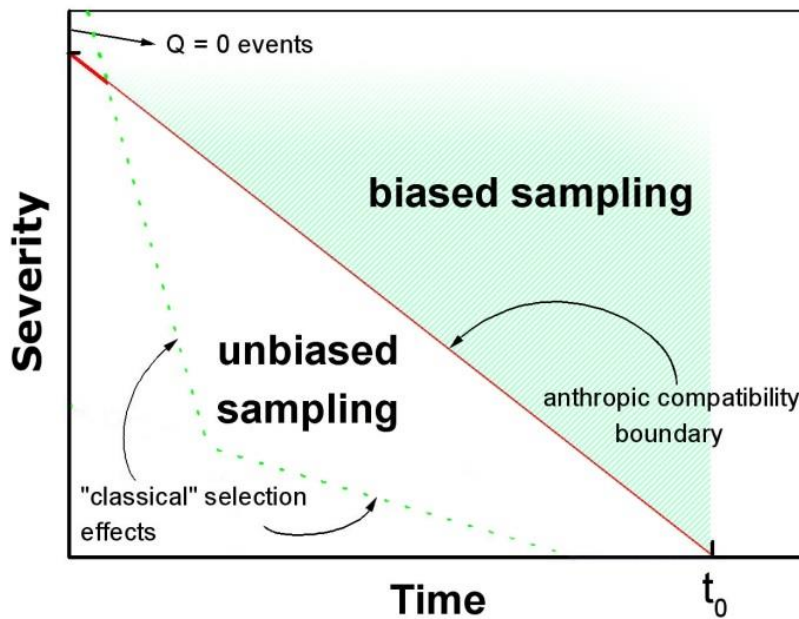


Figure 1: A sketch of the anthropic bias: we do not fairly sample the entire time-severity plane of large risks, only a region compatible with our existence at this particular epoch (the rest is in the “anthropic shadow” – shaded region, see text). The current epoch is denoted by t_0 and we count time from the formation of our planet (from Ćirković et al. 2010). Even if the amount of bias is small – that is, if evolution is convergent and robust against perturbations – this is still a legitimate effect which clearly demonstrates the importance of anthropic thinking in the field of risk analysis. $Q = 0$ events are completely destructive events which leave no survivors.

What happens as we increase – in a thought experiment if not actual historical record – the severity of the catastrophic event whose probability we wish to calculate? Catastrophic events exceeding some threshold severity eliminate all observers and all ecological conditions necessary for subsequent emergence of observers, and are hence unobservable. Some types of catastrophes may also make the existence of observers on a planet impossible in some subsequent interval, the size of which might be correlated with the magnitude of the catastrophe. Because of this bias, the events reflected in our historical record are not sampled from the full events space but rather from the part of the events space that lies beneath the “anthropic compatibility boundary” (illustrated in Figure 1). The part of the parameter space above the boundary lies in what can be called *anthropic shadow*: the observation selection effect implicit in conditioning on our present existence prevents us from sharply discerning magnitudes of extreme risks close (in both temporal and evolutionary terms) to us. This shadow is the source of bias which must be corrected when we seek to infer the objective chance distribution from the observed empirical distribution of events.

So much about anthropic selection effects in our *past*. What about the future? Global catastrophic risk studies uniformly indicate that the biggest threats for humanity’s future now belong to anthropogenic risks, such as nuclear winter, anthropogenic global warming or misuse of biotechnology or nanotechnology. Overconfidence following neglect of the anthropic bias applies here as well – if we have a bias in calculating the distribution function of impact catastrophes from the past leading us to underestimating impact risk, this is not dissimilar from reasoning that since there was no nuclear war in the last cca. 60 years, the risk of its occurring is smaller than we would α

priori expect. The lack of our prior historical experience does not diminish expected risk – it enhances it; while this may sound counterintuitive, it is a direct consequence of the anthropic reasoning.

This introduces overconfidence in our probabilistic estimates of the magnitude of risk above given severity, leading to possibly detrimental neglect of research and mitigation of particular possibly risky processes. Therefore, what is essentially a philosophical question about justification and discrimination between explanatory hypotheses becomes a very serious and very practical issue: *do we underestimate the risks we are unavoidably facing?* If the answer is affirmative, this leads straight to the situation in which methodological biases and neglects lead to increasing likelihood of harm for billions of humans and uncountable number of other living beings. (This can be generalized to other hypothetical intelligent communities in the Galaxy through another anthropic argument, as has been done by Tegmark and Bostrom 2005.)

So, in this case the downplaying of anthropic reasoning might be not only bad science and bad philosophy, but might be ultimately *unethical* as well. Since at least some of the processes subject to anthropic shadow could lead to the ultimate harm (Persson and Savulescu 2008), only minor auxiliary assumptions are necessary to conclude that it is our moral duty to increase our understanding and, ultimately, resilience to such processes. If anti-anthropic dogma stands in the way of better understanding, then this shows not only cognitive bankruptcy of such dogma, but the moral bankruptcy as well.

4. Discussion

It is quite easy to realize that the anthropic reasoning is, in fact, applicable to all fields and problems where processes could lead, or indeed have historically led, to a significant change in the number of observers. While such fields and problems are still not numerous, they delineate an area much wider than what is usually circumscribed to cosmology and the “bottom” of fundamental physics. To those mentioned above, one may add many other contribution from various other fields; for instance, Reinganum’s (1986) “financial proof” of the impossibility of time travel also belongs to this category, since it depends on observation selection against observers belonging to a particular reference class (those time travelling from the future to the present). The present discussion could be understood as a provocation and call for more similar examples from different scientific disciplines. This could help in defending the whole province of explanatory work from interference by mysticism and religion-inspired abuse.

Moreover, any research involving situations – possible or actual – with changing number of observers, are *necessarily* a subject to the anthropic bias. It is mandatory, not optional, to account and correct for this bias. Conversely, it is *bad science* not to do it, no matter what philosophical and methodological prejudices might be in play. In my view, it is also a case of *bad philosophy* to call for rejection of the question or explanatory nihilism *without previously investigating all explanatory options and finding them unacceptable*. Unfortunately, the nihilistic approach still has many adherents when it comes to explanation of very general phenomena of cosmology and fundamental physics (e.g., Callender 2004; Mosterín 2005); the idea is that the issues such as the initial entropy of the universe or the strength ratios of low-energy effective forces are so general that everything else, including entities used in a potential explanation, is contingent upon those fundamental “almost

brute” facts. (The qualification “almost” pertains to the lip service given to possible future advances in physics, which is often not taken seriously at all. Thus Mosterin claims, for instance, that “the most desirable epistemic situation would be that the values of the fundamental constants... could one day be derived from some fundamental physical theory. In the mean time (and this mean time can last *forever*) they *have to be accepted as brute facts*.” [emphasis M. M. C.] Barring the decision of some Central Committee of Science or a Supreme Philosophical Leader, I completely fail to see any reason whatsoever for such compulsion.) So we cannot formulate a causal explanations for those ultra-deep empirical facts about the physical universe, since we do not know in terms of what to formulate such an explanation. We should accept them as primitive happenstances and proceed further from that.

Clearly, the nihilistic strategy fails in clearly circumscribed fields such as planetary science or risk analysis considered above. It would be unacceptable – not to mention rude – to reject further discussion in these fields even in cases where due to historical loss of information we do not have insight in causal agents acting in distant past. While rudeness has not been, however, below some of the anthropic detractors' discourse – e.g., abovementioned numerical errors in the paper by Klee (2002) criticising anthropic theorists for laxity with numbers – it is obvious that those parts of the battleground are effectively lost. We cannot consciously avoid discussing the impact of habitability on the number of observers or selection effects downsizing huge physical risks we are facing. In the latter case, declining to engage in the discussion has adverse morale consequences as well. So, examples like those discussed in the present paper act to decrease our confidence in explanatory nihilism even in cosmology or string theory, quite independently of actually affirming the value of the anthropic reasoning.

We conclude with a subjective and speculative outline of what could be an interesting project in sociology and psychology of science. It is a telling phenomenon that the existence of wide applications of anthropic reasoning has been a “well-guarded secret” for quite some time – and that the prejudice that only relevant conclusions could be drawn from it in string theory and cosmology is so widespread. I suspect that a major factor in resistance to anthropic arguments and explanations is the fact that they go against received wisdom of two otherwise strongly opposed broad camps, which, for the lack of better words, I will call *naive reductionism* and *naive teleology*. In both cases, the qualification “naive” is central, since anthropic accounts – as discussed above – do not really undermine the role of reductionism and, to a limited extent, teleology in explanations. It is one of those prejudices which is strongly supported by proponents of two otherwise opposing views – and therefore it is much harder to reject. In the course of the Cold War, for instance, both sides have been in agreement, for very different reasons, of course, that USSR and its satellite dictatorships are countries of “real socialism”. Many honest and well-intentioned social-democrats or non-Marxist socialists in the West have had a hard time in trying to dispel this notion, since it was promoted by both powerful propaganda apparati. The position of anthropic theorist is similarly uncomfortable, being attacked by both naive reductionists and naive teleologists; the metaphor of Scylla and Charybdis immediately comes to mind. Both naive positions strive to contain anthropic reasoning – if they cannot stamp it out completely – as a contagious heresy, which must not pollute “normal” fields of inquiry. Hence completely nonsensical repeated claims that the anthropic reasoning deals in tautologies, which cannot tell us anything we don't already know; quite to the contrary, those are *quite informative* tautologies, which tell us much about the world.

Naive reductionism, usually sprinkled with remnants of logical positivism, cannot ever come to terms with the idea that extremely complex structures such as intelligent observers may play a role in explanation on any level. Naive reductionists tend to reject or downplay any explanation which is not mechanist and atomic explanation and implicitly hold the dogma that there must be a unique underlying micro-level causal account for any observed macro-level phenomenon. Obviously, the reference to observers and their *historical experience* contained in the very heart of the anthropic reasoning is anathema to naive reductionists. It is very similar to suspicions and misunderstandings characterizing the response of many *secular* scientists and philosophers to neodarwinian evolutionary paradigm; in both cases I suspect deeply rooted suspicion toward using extremely complex and inevitably historical entities in explanatory accounts.¹⁹ (And other modes of explanation, in particular geometric or topological ones, fare no better at hand of naive reductionists; this only confirms that naive reductionism should be fought against and overcome as much as possible.)

Similarly, naive teleologists are dismayed by the *disteology* of anthropic reasoning. What could be a splendid opportunity for prolonged mystical musings on the “relationship of mind and the universe”, “the role of mind [or Mind] in the overall scheme of things”, and similar theatrical clichés, is suddenly discounted as the consequence of simple observation selection effects and could be quantitatively modelled in Bayesian terms! How terribly inconvenient. And not only creationists and similar backward scarecrows feel uncomfortable. While naive teleology is, at first glance, mostly absent from serious literature in at least last several decades and therefore seemingly presents no serious threat to science, one should remain highly vigilant, since anti-Copernican, teleological, and anthropocentric thinking is still powerfully present in popular culture (including many instance of popular science and philosophy), as well as in large segments of the media, politics, arts, and humanistics. Especially when coupled with more virulent strains of social constructivism, epistemic relativism and postmodernism, it still has the capacity to do great and hard-to-repair harm. But even in the relatively benign form exhibited, for instance, by Midgley (1985) or the anthropic detractors cited in the introduction such as Earman or Pagels or Mosterin, it can produce confusion and lead to closed-mindedness on many sensitive scientific issues, from animal sentience, to artificial intelligence, to SETI projects.

That the double fallacy of both naive reductionists and naive teleologists has some traction is testified, in a bizarre way, by perhaps the *least* naive of the naive teleologists of the 20th century, Sir Fred Hoyle, who in his popular book *The Intelligent Universe* proclaimed that “[t]he same nihilistic belief that no aspect of the Universe can be thought of as a consequence of purpose underlies both Darwinism and the anthropic principle.” (Hoyle 1983) His caricature of Darwinism amounts to what I have above labelled naive reductionism (specifically in biology). It sounds funny and outright bizarre today, but Hoyle seems to have been fully aware that there are altogether *three* positions, rather than two, and that the third one, which is the anthropic reasoning, is bound to be regarded with suspicion and hostility by the adepts of the other two.

The *expanded mandate* of the anthropic reasoning, as dealing with any field with variable number of observers, could easily be branched into other domains. Elaboration of the properties of observers requires multidisciplinary collaboration between mathematics and physics of complexity,

¹⁹ The book of Fodor and Piattelli-Palmarini (2010) is an example of this.

evolutionary biology and astrobiology, neurosciences, risk analysis, and philosophy. There is a wide front along which progress could be made by better understanding and building quantitative models of physical, chemical, planetological, etc. preconditions for observership. Both work in artificial intelligence and in zoopsychology and in SETI studies should give us insight into properties of observers different from ourselves. Fields which we have on purpose avoided here, cosmology and fundamental physics, give us tantalizing glimpses of the multiverse, and the interest in physical properties of other universes and their habitability grows almost by day.²⁰ Prospects for further research are bright indeed. Diametrically opposed to the dark, scholastic, uninformative picture painted by the detractors, the anthropic reasoning is indeed liberating, future-oriented and unifying.

Acknowledgements. I wish to thank Jelena Dimitrijević, Nick Bostrom, Anders Sandberg, Branislav Vukotić, Slobodan Popović, Slobodan Perović, Ivana Kojadinović, Karl Schroeder, Aleksandar Obradović, Jelena Andrejić, Goran Milovanović, George Dvorsky, Zoran Knežević, Jacob Haqq-Misra, the late Robert Bradbury, and the late Branislav Šimpraga for many pleasant and useful discussions on the topics related to the subject matter of this study. This is an opportunity to thank KoBSON Consortium of Serbian libraries, as well as NASA Astrophysics Data System. This research has been supported by the Ministry of Education and Science of the Republic of Serbia through the project ON176021.

References

- Adams, F. C. 2008, "Stars in other universes: stellar structure with different fundamental constants," *Journal of Cosmology and Astroparticle Physics* **08**, 010.
- Adams, F. C., Coppessa, K. R., and Bloch, A. M. 2015, "Planets in other universes: habitability constraints on density fluctuations and galactic structure," *Journal of Cosmology and Astroparticle Physics* **09**, 030.
- Aguirre, A., Johnson, M.C., and Shomer, A. 2007, "Towards observable signatures of other bubble universes," *Physical Review D* **74**, 063509[17].
- Barnes, L. A. 2012, "The Fine-Tuning of the Universe for Intelligent Life," *Publications of the Astronomical Society of Australia* **29**, 529-564.
- Barrow, J. D. and Tipler, F. J. 1986, *The Anthropic Cosmological Principle* (Oxford University Press, New York).
- Bishop, S. and Egli, R. 2011, "Discovery prospects for a supernova signature of biogenic origin," *Icarus* **212**, 960-962.



²⁰ Among rapidly growing literature on this topic, let us mention Harnik, Kribs, and Perez (2006); Aguirre, Johnson, and Shomer (2007); Adams, (2008); Schellekens (2013); Adams, Coppessa, and Bloch (2015).

- Boltzmann, L. 1895, "On Certain Questions of the Theory of Gases," *Nature* **51**, 413-414.
- Bostrom, N. 2002, *Anthropic Bias: Observation Selection Effects in Science and Philosophy* (Routledge, New York).
- Bostrom, N. and Ćirković, M. M. (eds.) 2008, *Global Catastrophic Risks* (Oxford University Press, Oxford).
- Callender, C. 2004, "Measures, Explanation and the Past: Should 'Special' Initial Conditions Be Explained?" *British Journal for the Philosophy of Science* **55**, 195-217.
- Carroll, S. M. 2006, "Is Our Universe Natural?" *Nature* **440**, 1132-36.
- Carter, B. 1974, "Large number coincidences and the anthropic principle in cosmology," in M. S. Longair (ed.) *Confrontation of cosmological theories with observational data; Proceedings of the Symposium, Krakow, Poland, September 10-12, 1973* (D. Reidel Publishing Co., Dordrecht), 291-298.
- Carter, B. 1983, "The anthropic principle and its implications for biological evolution," *Philos. Trans. R. Soc. London A* **310**, 347-363.
- Carter, B. 1993, "The Anthropic Selection Principle and the Ultra-Darwinian Synthesis," in *The Anthropic Principle, Proceedings of the Second Venice Conference on Cosmology and Philosophy*, ed. by F. Bertola & U. Curi (Cambridge University Press, Cambridge), 33-66.
- Ćirković, M. M. 2002, "On the First Anthropic Argument in Astrobiology," *Earth, Moon, and Planets* **91**, 243-254.
- Ćirković, M. M. 2003, "The Thermodynamical Arrow of Time: Reinterpreting the Boltzmann-Schuetz Argument," *Foundations of Physics* **33**, 467-490.
- Ćirković, M. M. 2004, "The Anthropic Principle and the Duration of the Cosmological Past," *Astronomical and Astrophysical Transactions* **23**, 567-597.
- Ćirković, M. M. 2006, "Too Early? On the Apparent Conflict of Astrobiology and Cosmology," *Biology and Philosophy* **21**, 369-379.
- Ćirković, M. M., Sandberg, A., and Bostrom, N. 2010, "Anthropic Shadow: Observation Selection Effects and Human Extinction Risks," *Risk Analysis* **30**, 1495-1506.
- Chopra, A. and Lineweaver, C. H. 2016, "The Case for a Gaian Bottleneck: The Biology of Habitability," *Astrobiology* **16**, 7-22.
- Chyba, C. F. and Hand, K. 2005, "Astrobiology: The Study of the Living Universe," *Annu. Rev. Astron. Astrophys.* **43**, 31-74.
- Clube, S. V. M. and Napier, W. M. 1990, *The Cosmic Winter* (Basil Blackwell Ltd, Oxford).
- Collier, J. 1990, "Intrinsic information," in *Information, Language and Cognition: Vancouver Studies in Cognitive Science*, P. P. Hanson, ed. (Oxford University Press, Oxford), 390-409.

- Des Marais, D. J., Nuth, J. A. III, Allamandola, L. J. et al. 2008, "The NASA Astrobiology Roadmap," *Astrobiology* **8**, 715-730.
- Dick, S. J. 2003, "Cultural Evolution, the Postbiological Universe and SETI," *Int. J. Astrobiology* **2**, 65-74.
- Dicke, R. H. 1961, "Dirac's Cosmology and Mach's Principle," *Nature* **192**, 440-441.
- Dirac, P. A. M. 1937, "The Cosmological Constants," *Nature* **139**, 323.
- Dirac, P. A. M. 1961, "Dirac replies [to Dicke's letter]," *Nature* **192**, 441.
- Doyle, L. R., et al. 2011, "Kepler-16: A Transiting Circumbinary Planet", *Science* **333**, 1602-1606.
- Dunbar, D. N., Pixley, R. E., Wenzel, W. A., and Whaling, W. 1953, "The 7.68-Mev State in C¹²," *Physical Review* **92**, 649-650.
- Earman, J. 1987, "The SAP also rises: a critical examination of the anthropic principle," *American Philosophical Quarterly* **24**, 307-317.
- Ellis, G. F. R. 2011, "Editorial note to: Brandon Carter, Large number coincidences and the anthropic principle in cosmology," *General Relativity and Gravitation* **43**, 3213-3223.
- Fodor, J. and Piattelli-Palmarini, M. 2010, *What Darwin Got Wrong* (Farrar, Straus, and Giroux, New York).
- Fontenelle, B. B. [1686] 1990, *Conversation on the Plurality of the Worlds*, (trans. by H. A. Hargreaves, University of California Press, Berkeley).
- Freer, M. and Fynbo, H. O. U. 2014, "The Hoyle state in ¹²C," *Progress in Particle and Nuclear Physics* **78**, 1-23.
- Grieve, R. A. and Shoemaker, E. M. 1994, "The Record of Past Impacts on Earth," in T. Gehrels (ed.) *Hazards due to Comets and Asteroids* (University of Arizona Press, Tucson), 417-464.
- Halley, E. 1704, "Astronomiae Cometicae Synopsis, Autore Edmundo Halleio apud Oxonienses. Geometriae Professore Saviliano, & Reg. Soc. S.," *Philosophical Transactions of the Royal Society of London* **24**, 289-304.
- Halley, E. 1726, "Some Considerations about the Cause of the Universal Deluge," *Philosophical Transactions of the Royal Society of London* **33**, 118-123.
- Hanslmeier, A. 2009, *Habitability and Cosmic Catastrophes* (Springer, Berlin).
- Harnik, R., Kribs, G., and Perez, G. 2006, "A universe without weak interactions," *Physical Review D* **74**, 035006[15].
- Horneck, G. and Rettberg, P. (eds.) 2007, *Complete Course in Astrobiology* (Wiley-VCH, Weinheim).

- Hoyle, F. 1954, "On Nuclear Reactions Occuring in Very Hot Stars. I. the Synthesis of Elements from Carbon to Nickel," *Astrophysical Journal Supplement* **1**, 121-146.
- Hoyle F. 1983, *The Intelligent Universe* (Michael Joseph Limited, London).
- Hughes, D. W. 2003, „The approximate ratios between the diameters of terrestrial impact craters and the causative incident asteroids," *Monthly Notice of the Royal Astronomical Society* **338**, 999-1003.
- Gould, S. J. 1987, *The Flamingo's Smile: Reflections in Natural History* (W. W. Norton & Company, New York).
- Klee, R. 2002, "The Revenge of Pythagoras: How a Mathematical Sharp Practice Undermines the Contemporary Design Argument in Astrophysical Cosmology," *Brit. J. Phil. Sci.* **53**, 331-354.
- Leslie, J. 1989, *Universes* (Routledge, Oxford).
- Maddox, J. 1984, "New Twist for the Anthropic Principle," *Nature* **307**, 409.
- Manson, N. A. 2000, "There Is No Adequate Definition of 'Fine-tuned for Life,'" *Inquiry* **43**, 341-352.
- Midgley, M. 1985, *Evolution as a Religion: Strange Hopes and Stranger Fears* (Routledge, London).
- Mosterín, J. 2000, "The Anthropic Principle in Cosmology: A Critical Review," *Acta Institutionis Philosophiae et Aestheticae* (Tokyo, Japan) **18**, 111-139.
- Mosterín, J. 2005, "Anthropic Explanations in Cosmology," in Hajek, Valdés & Westerstahl (eds.), *Proceedings of the 12th International Congress of Logic, Methodology and Philosophy of Science* (North-Holland Publishing, Amsterdam).
- Newton, I. 1730, *Opticks* (William Innys, London).
- Pagels, H. R. 1998, "A Cozy Cosmology", in J. Leslie (ed.), *Modern Cosmology & Philosophy* (Prometheus Books, Amherst), 180-86.
- Persson I. and Savulescu J. 2008, "The Perils of Cognitive Enhancement and the Urgent Imperative to Enhance the Moral Character of Humanity", *Journal of Applied Philosophy* **25**, 162-177.
- Reinganum, M. R. 1986, "Is Time Travel Impossible? A Financial Proof," *The Journal of Portfolio Management* **13**, 10-12.
- Schellekens, A. N. 2013, "Life at the interface of particle physics and string theory," *Rev. Mod. Phys.* **85**, 1491 –1540.
- Schindewolf, O. 1962, "Neokatastrophismus?" *Deutsch Geologische Gesellschaft Zeitschrift Jahrgang* **114**, 430–445.
- Smolin, L. 2007, "Scientific alternatives to the anthropic principle," in *Universe or Multiverse?*, ed. by Bernard Carr (Cambridge University Press, Cambridge), pp. 323-366.
- Steckline, V. S. 1983, "Zermelo, Boltzmann, and the recurrence paradox," *Am. J. Phys.* **51**, 894-897.

- Susskind, L. 2006, *The Cosmic Landscape: String Theory and the Illusion of Intelligent Design* (Back Bay Books, New York).
- Tegmark, M. 1996, "Does the universe in fact contain almost no information?" *Found. Phys. Lett.* **9**, 25-41.
- Tegmark, M. 2008, "The Mathematical Universe," *Foundations of Physics* **38**, 101-150.
- Tegmark, M. and Bostrom, N. 2005, "Astrophysics: Is a doomsday catastrophe likely?" *Nature* **438**, 754.
- Walker, M. A. and Ćirković, M. M. 2006, "Anthropic Reasoning, Naturalism and the Contemporary Design Argument," *International Studies in the Philosophy of Science* **20**, 285-307.
- Wiener, N. 1961, *Cybernetics* (John Wiley and Sons, New York).
- Wilson, P. A. 1994, "Carter on Anthropic Principle Predictions," *Brit. J. Phil. Sci.* **45**, 241-253.
- Woo, G. 1999, *The Mathematics of Natural Catastrophes* (Imperial College Press, London).