***Self-Preservation and Information:***

***A Dual-Goals Hypothesis of Extraterrestrial Behavior***

**William C. Lane**

**The Nexus Group**

**2987 District Avenue**

**Suite 313**

**Fairfax, VA 22031**

**lane2125@yahoo.com**

Developments that suggest the universe is full of life make the Fermi paradox increasingly pressing, but our search for an extraterrestrial technological civilization (“ETC”) is handicapped by our ignorance of its probable nature and behavior. This paper offers a way around this problem by drawing on information theoretical concepts, including game theory and Bayesian probability. It argues that, whatever its ultimate goals, an ETC would have the same instrumental goals as other intelligent agents. Generically, these are self-preservation and the acquisition of resources. For an advanced ETC, these generic goals imply the objectives of removing existential threats and acquiring strategic and non-strategic information. Since the most problematic existential threats for any ETC would be from other civilizations and planets hosting such civilizations would also be copious sources of information, these objectives would lead it to gather information from these locations and to protect itself against them. This *dual-goals hypothesis* resolves the Fermi paradox, redirects the search for extraterrestrial intelligence and makes testable predictions regarding ETC’s behavior on Earth if it is here.

***Keywords***

SETI; existential risk; instrumental goals; Fermi paradox; dark forest; semantic information

***Article***

1. *Introduction – The Fermi Paradox*

The universe is vast and various. Recent discoveries suggest it is full of life. On average, 0.18 Earth-sized planets orbit in the habitable zone of each F, G and K class star in our galaxy [1]; many of these possess surface water [2] and some may be more habitable than Earth [3]. It seems increasingly likely that life originates widely [4-7]. Fitness landscapes are highly navigable [8], the trend of biological evolution is toward greater complexity [9-10] and life seems capable of shaping its environment to suit its needs [11-12]. The claim that complex life requires rare conditions seems increasingly improbable [13-14]. Complex brains and intelligence seem to be a convergent trait [15], and language, the only uniquely human capacity, evolved like other traits [16-17]. The discovery of massive ancient galaxies [18] lengthens the time available for the evolution of complexity and – because galaxies were closer together in the earlier universe – may increase the likelihood of life’s intergalactic spread. In this context, the failure of the search for extraterrestrial intelligence (“SETI”) to show positive results has made Fermi’s question – “Where is everybody?” – increasingly pressing [19].

Cabrol’s [20, p. 667] observation that, “to find aliens, we must … understand the many ways they could manifest themselves in their environment and communicate their presence” describes an obstacle to this search: An extraterrestrial technological civilization (“ETC”) may be a society of life forms unlike those on Earth [21-22], a synthetic, post-biological system [23], or a symbiotic combination. Whatever it is, an ETC would be the “strangest stranger” we will ever encounter [24]. For Cabrol [20, p. 665], this follows from the “principle of the coevolution of life and environment.” The interaction between life and its planetary environment will “dictate the uniqueness of each planetary experiment … and will do so not only when (or if) life reaches the stage of technological advancement. It will start from the very first moment, as it did on Earth.” Yet, how are we to find ETC unless we know what to look for? Relying on naïve expectations and Earthbound analogies to guide our search is likely to lead us astray. This paper therefore proposes an information theoretical basis for predicting ETC’s goals, and through them its likely pattern of behavior, even absent any knowledge of its intrinsic nature. Since the proposal is that ETC, like other intelligent agents, would pursue two instrumental goals, it can be called the *dual-goals hypothesis*. On this hypothesis, ETC would neither transmit a persistent “lighthouse” signal nor engage in widespread colonization, resolving the paradox. This paper begins by defining the features and motives likely to describe any ETC. From these, it turns to its likely capabilities, the objectives its motives and these imply, and its likely behavior.

1. *Features and motives*

The following features and motives seem likely to belong to an ETC, regardless of its specific nature.

● *It would be a technologically capable physical reality originating from off Earth but existing within our spacetime.* This parsimonious assumption excludes merely metaphysical possibilities and restricts our search to entities of a type we know to exist.

● *It would be a teleonomic agent.* It would act with apparent purpose upon and in its environment. Humans and animals are teleonomic agents; so are corporations and governments. So would be an artificial general intelligence (“AGI”). The claim that ETCs are agents does not imply the absence of divisions or disputes within them [25]; it only means an ETC can respond to its environment as an entity.

● *It would be capable of higher-order cognition*. “Cognition can be defined as the activity, acquisition, organization and usage of knowledge inherent in every living organism” [24, p. 701]. Higher-order cognition is cognition that operates effectively in a wide range of environments and situations. Without defining intelligence, we can define an *intelligent agent* as a teleonomic agent that uses higher-order cognition to regulate its behavior and shape its environment. An ETC would be an intelligent agent in this sense. Because humans are intelligent agents, we can appeal to human behavior as a check on the plausibility of our conclusions, but not as the basis for them.

● *The technology of a modal ETC would be markedly in advance of our own* [26]. Our galaxy is over 13 billion years old and its *currently* potentially habitable planets have a median age ~2 billion years greater than Earth’s. Given our extreme youth, as a planet and a scientific civilization, there are vastly more chances for this claim to be true than for it to be false. Hereafter, “ETC” should be read to mean a technologically advanced ETC.

● Any ETC would have explicit or implicit *ultimate* goals, ends-in-themselves that are intrinsically valuable to it. We can know nothing of these. We can, however, know this: *Whatever its ultimate goals, it would also have instrumental goals: goals that are valuable to it because their achievement is essential to the achievement of its ultimate goals* [27-28].

●*Regardless of its ultimate goals, an intelligent agent would possess two convergent instrumental goals: self-preservation and the amassing of resources* [27-29]. Bostrom offers the example of a computer programmed to produce paper clips, a seemingly harmless ultimate goal that could lead it to resist being turned off (for then there would be fewer paperclips) and also to amass control over as much of the world’s matter as possible, with the aim of turning it all into paperclips. The notion that any agent would act to ensure its continued existence seems self-evident, while human and non-human animals seek to amass resources whenever the foreseen benefit outweighs the probable cost [30].

● *Since intelligent agents will converge on the instrumental goals of self-preservation* *and the amassing of resources and ETC is presumed to be an intelligent agent*, *it would have these goals*. These are generic goals whose specific meaning will depend on the agent’s capabilities and situation. For humans (as an example), “resources” can include money, power, prestige, time and (as with all animals) semantic information. Unless ETC’s ultimate goals expressly involve less advanced civilizations, they seem unlikely to affect us; we will, therefore, take these convergent instrumental goals to be ETC’s only goals for purposes of this analysis.

*3.* *Capabilities*

An ETC’s technology would differ from ours. First, because it would know more science than we do; second, because its technology will reflect the contingencies of its history and its resultant way of thinking [20, 31]. Compare the Polynesian outrigger to the Viking longship or the Chinese to the European wheelbarrow. The complex technologies developed by isolated, technological civilizations should differ more widely than these simple technologies, for there are more degrees of freedom in the design of complex technologies than simpler ones. These differences make imagining ETC’s technological capabilities difficult. Yet, if we can’t do that, how can we make its goals more specific? A group of hunter-gatherers and a spacefaring civilization would both seek to secure their survival and to acquire resources, but these generic goals would imply different behaviors for each.

Rather than trying to describe *technologies* ETCs might have, we can describe *constraints* they would likely have overcome, basing the list on those that we may overcome in the not-too-distant future. Taking this approach, ETC would not be constrained by:

*Its energy supply*. Either nuclear fusion or solar power could in principle supply all of Earth’s energy needs for the indefinite future. In the longer term, antimatter may be a nearly limitless energy source [32]. It is often assumed that economic prosperity requires increasing energy consumption, but more efficient technologies are breaking that link [33-34]. There is no reason to think an ETC would require (or desire) the ever-increasing energy supplies the Kardashev scale implies, for miniaturization, lightweight materials, virtual reality and quantum computing would reduce its energy needs [35]. The apparent nonexistence of any Kardashev Type III civilization [36] supports this claim.

*Aging and death*. Humanity is gradually escaping these constraints. Removing them would mean that, except for sudden events that destroy vital organs, individuals could live as long as they wished, perhaps in purpose-built bodies [37]. A society with this capacity is likely to have a long planning horizon and a correspondingly low implicit discount rate. The discount rate we use to value future costs and benefits is grounded in our biological nature; it would be lower if we expected to live much longer than we do. The higher implicit discount rate used by older adults (with a shorter expected lifespan) than by all adults demonstrates this point [38]. In animals, there is a strong correlation between longer lifespans and lower rates of propagation [39], suggesting that ETC would have a stable population.[[1]](#footnote-1) More speculatively, Smart [40] suggests that an advanced ETC would locate near a black hole, slowing its proper time with respect to ours and dramatically lengthening its time horizon with respect to events outside the black hole region.

*Its home star*. An ETC would be technologically capable of travel to other planets or locations in interstellar space, making colonization possible.[[2]](#footnote-2) Even intergalactic colonization would be feasible for a highly advanced civilization [41]. Minimally, it could pack information about itself into an escape vehicle, then use local material to construct a replica of itself at its destination. In at least this sense, it could escape the destruction of its home star. Left alone, an ETC could become very old indeed, perhaps by surviving multiple civilizational resets [42].

*Material resources*. Advanced manufacturing techniques [43], other forms of high technology, low-cost energy and a stable population would reduce an ETC’s need for physical resources. It could likely satisfy that need by substitution of materials, recycling and mining around its home star [44-45]. It would therefore have little need for resources from other stars, which would be expensive to obtain and transport [45].

*Limits on its observations*. We continue to build more powerful space- and ground-based telescopes and have already launched interstellar probes; more rapid modes of interstellar travel are under study [47]. The size and energy cost of probes could be minimized by the use of nanotechnology and lightweight materials, two rapidly advancing technologies. An ETC could program small, super von Neumann probes to create new probes or other systems using materials found at or near their destination [48-50].[[3]](#footnote-3)

*Openness to observation.* The Committee on SETI Nomenclature defines intelligence as “the quality of being able to deliberately engineer technology which might be observable using astronomical observation techniques” [24, p. 700]. It might as properly define it as the ability to engineer technology that benefits the designer but *cannot* be observed by others [see 51]. ETC would be capable of effective concealment both locally and at a planetary level [52].

*Native powers of cognition*. In 2020, 72 active projects were aimed at producing AGI [53]. Within a century or less, an AI system is likely to become “superintelligent” [27]: that is, to surpass its human creators in computing capacity and a broad range of cognitive capabilities [54]. “Superexponential” increases in computational efficiency [55] will vastly increase the quantity of information a civilization can process and the speed of computation. Superintelligent systems, whether still controlled by their builders or not, would be more cognitively capable – including being more capable of persuasion, deception and other sociolinguistic tasks – than the builders themselves, though bio-engineered brains might match them [37].

This is not an inclusive list; ETC is likely to have other capabilities. These may include a capacity for faster-than-lightspeed (“FTL”) travel [56-57], though its physical possibility remains uncertain.

In one important way, an ETC’s situation would resemble our own. Like us, it could never know what further scientific and technological progress in fields important to its survival might be possible or what forms it might take. This claim is grounded on the historical unpredictability of scientific and technological discovery; including the observation that “the larger and more important the discoveries, the less predictable they would have been” [58, p. 74]. Natural languages are potentially infinite because there is no end of ways a sentence can be constructed; the range of technologies a given understanding of physics can possibly support seems limitless for the same reason. An ETC’s belief that it has the “final theory” of physics could never be held with certainty, for any such belief would be an empirical scientific claim.

*4. Dual Objectives*

For an ETC with the characteristics described in section 2, armed with the capabilities described in section 3, the dual goals of self-preservation and amassing resources would imply the more specific objectives described in this section. Together, its capabilities and these objectives frame the dual-goals hypothesis.

*4.1. Existential Risk*

Philosopher Derek Parfit [59, pp. 453-454] offers this thought experiment:

Compare three outcomes:

1. Peace.
2. A nuclear war that kills 99 per cent of the world’s existing population.
3. A nuclear war that kills 100 per cent.

2 would be worse than 1 and 3 would be worse than 2. … [But] I believe that the difference between 2 and 3 is very much greater. The Earth will remain habitable for at least another billion years. Civilization began only a few thousand years ago. If we do not destroy mankind, these few thousand years may be only a tiny fraction of the whole of civilized human history. The difference between 2 and 3 may thus be the difference between this tiny fraction and all of the rest of this history.

A threat that could destroy 100% of the population with no hope of recovery is an existential threat. “What makes existential catastrophes especially bad is … that they would destroy the future” [60, p. 17].

Until recently, there was little we could do to protect against extrinsically caused existential threats; however, that is changing. Our discovery of a meteor’s role in the extinction of the dinosaurs has led us to develop an effective defense against that threat [61]. Despite a planet-killing event’s low probability, this is money well spent. Bostrom and Ćirković [60, pp. 18-19] calculate “the expected value of reducing existential risk by a mere *one millionth of one percentage point* [to be] at least a hundred times the value of a million human lives.” This assumes that Earth will remain habitable for just a billion years; to an ETC capable of outlasting its home star, an existential risk would have much greater disvalue.

An advanced ETC would know more about possible threats to its existence, be more attuned to their disvalue, and be better equipped to avoid or remove them than we are. Some threats would have a natural source, but other advanced societies could also pose existential threats. On Earth, the active SETI debate revolves around that potential.

Science fiction writer Liu Cixin [62, p. 484] describes how the perception of risks from other societies could affect ETC behavior:

The universe is a dark forest. Every civilization is an armed hunter stalking through the trees like a ghost, gently pushing aside branches that block the path and trying to tread without sound…. The hunter has to be careful because everywhere in the forest are stealthy hunters like him. If he finds another life — another hunter, angel, or a demon, a delicate infant to tottering old man, a fairy or demigod — there’s only one thing he can do: open fire and eliminate them.

This is a plausible scenario. Even large distances between ETCs are unlikely to preclude eventual contact between extremely long-lived civilizations. Nor would the projected time to contact necessarily affect its significance to an ETC. For a sufficiently long-lived society, a risk expected to arise 1,000,000 years in the future might be evaluated much as we would evaluate one expected to arise 100 years in the future.

The history of military technology suggests that weapons will only become more effective, stealthy and deadly. Empathy evolves in the context of relations with conspecifics [63], and is unlikely to extend to wholly unrelated organisms. Tit-for-tat, a strategy that often leads to cooperation among competitors, breaks down when actors cannot know what the actions or statements of their adversaries mean; in that case, “a single mistake about the intentions of the adversary can lead to retaliation and start an endless string of counterstrikes” [64, p. 1075]. Tit-for-tat, like other strategies that build cooperation among competitors, also depends on the possibility of repeated communication, but that would likely result in disclosure of ETC’s location, the very piece of information it would be most loath to disclose. An advanced ETC would know all this, and would likely choose to act rather than be acted upon.

To start with, it would adopt a policy of stealth. In a “dark forest” scenario, providing any information about oneself to another civilization, would always be a mistake, even if you have a technological advantage. Human technology, Liu writes [62, p. 483],

developed over the course of three hundred years. On the scale of the universe, that’s not development. It’s an explosion! … And it might be that my knowledge of your existence and the information I received from our communication was the perfect spark to set off [another] explosion. This means that even though I’m just a newborn or growing civilization I’m still a big danger to you.

AGI dramatically increases this danger by increasing the tempo of a society’s technological development. As Kurzweil [65, p. 97] explains, “once the ‘knee of the curve’ is achieved and the exponential growth explodes, the linear models break down.” Imagine an ETC 100 light years from Earth picking up signals emitted in the 2020’s. How could it know we were not already a threat at the time of their reception? Or that we might not become one during the time it would take to respond? From our perspective, this may seem unlikely, but an ETC may not be so sanguine.

The threat might be indirect. Since there is no ceiling on technological advancement, no ETC, however advanced, could ever be completely sure some *equally or more advanced* ETC was not within striking range of it. The ongoing possibility of entry into our galaxy from elsewhere [41] makes certainty on this point impossible. A less advanced society that learned of ETC’s existence and location might broadcast that information, thereby alerting this more potent foe. The less advanced society might do this accidentally or intentionally, in the latter case using a broadcast message as a weapon [62]. But even a wholly innocent broadcast could create an existential risk.

Analyzing Liu’s “dark forest” scenario using game theory, Yasser [66] concludes, “[I]t’s Pareto Optimal and even Nash Equilibrium to destroy any civilizations, those one knows of, and not share existence information out of fear of being demolished by a more potent civilization – or even a weaker one at a future turn of the game.” The claim that an ETC would respond decisively to any potentially existential threat comports with statements by those responsible for national security. In 2001, Vice President Dick Cheney warned, “If there's a 1% chance that Pakistani scientists are helping al-Qaeda build or develop a nuclear weapon, we have to treat it as a certainty in terms of our response” [67]. Thus, the dark forest scenario transforms the general goal of self-preservation into an objective of destroying potentially dangerous competitors.

In doing so, it or something close to it (*e.g.,* the dual-goals hypothesis outlined here) resolves the Fermi paradox [68]. The main objection to other sociological solutions to the paradox is that they must be quasi-universal to be effective [69]. But if just one ETC in our galaxy came to believe in the dark forest scenario, others *would* inhabit a dark forest whether they believed or not.

*4.2.* *Amassing Information*

For us, the term “resources” connotes material resources. But because the cost of transporting meaningful quantities of resources or products from another star system would exceed the cost of producing or manufacturing them at home [43-45] a demand for material resources is unlikely to motivate ETC’s involvement with other star systems. Rather than material resources, *semantic information* would be the resource it would wish to extract from interstellar sources [46]. Because information can be highly compressed, the energy and material cost of transporting or communicating it across interstellar distances would be small.

The acquisition and processing of semantic information (hereafter “information”) is an essential life process [22, 70]. The value of social information – information derived from the experience of others – likely drove the evolution of human language [16], and the quantity/diversity of social information a society requires increases with its technological sophistication. Contemporary civilization values information provided by scientists, programmers, reporters, historians, doctors, engineers, chefs and many others. Many major companies have value only because of the information they obtain and distribute. Countries expend huge resources to obtain strategic information about possible enemies (see section 5.1). Information alerts us to risks.

But we value semantic information even when it has no immediate use. Curiosity, a “demand for information that has no instrumental benefit,” looks to be “indispensable” to any complex system that needs to survive a real-world environment [71, p. 48]. Novel discoveries and inventions have sparked unexpected advances, often in unrelated fields [72]. Curiosity is so deeply embedded in our nature that we may regard its satisfaction as an ultimate value, but it is instrumental and convergent in this sense. A desire for novel experience is a form of curiosity that many animals share [73]. As these considerations appear to be universal, we should expect an ETC to seek out information of all sorts. More complex life forms embody more semantic information than simpler forms, and technological civilizations are likely to embody vast amounts of it [74].

*5. A Dual-Goals Account of ETC Behavior*

One way to reason from an ETC’s capabilities and objectives to its probable behavior is to imagine its equivalent of SETI – its search for extraplanetary intelligence [25] – being led by two groups. *Defenders* would be primarily interested in its survival; *Explorers* would acknowledge the crucial importance of survival but would mainly be interested in obtaining information, both for instrumental reasons and to satisfy their wide-ranging curiosity. Defenders would veto any suggestion that the ETC should send the sort of intentional signal SETI practitioners hope to receive. To the contrary, they would insist that ETC should conceal itself to the greatest extent possible. But the groups would agree to gather *strategic* and *non-strategic* information by *passive* and *covert* means.

Information is strategic if it can be used to shape or support the competitive strategy or aims of an entity, especially as against the source of the information [75]. All other forms of information, useful or not, are non-strategic. Passive information gathering, for example by astronomical instruments, emits no traceable signal; it is, therefore, the safest means of information gathering. Covert information gathering by probes or similar means entails a risk of discovery but information asymmetry [76] would be maintained if discovery of a probe would yield little or no strategic information. Because disclosure of ETC’s home location would create the greatest risk, it would guard that secret closely, while seeking out other civilizations.

*5.1. Strategic Information Gathering*

The purpose of strategic information gathering is to gain a decision advantage in a possible or ongoing contest [77-78]. Sun Tzu [79, p. 57] summarized the objective as “foreknowledge*.*” An intelligence professional explains, “[I]nformation creates the opportunity for our side to act before events limit our choices” [80].

Secrecy is essential to achieving this goal. In even the simplest games (*e.g*., rock, paper, scissors), knowledge of your opponent’s next move has value only if they do not know you know and cannot modify their strategy in the light of that knowledge [81]. If strategic information-gathering operations are discovered, they may be destroyed or (worse) subverted to the opponent’s ends [79]. Covert information-gathering for strategic purposes (espionage) has therefore been ubiquitous in human history, arising many times independently.

The principles of espionage are not specific to humanity; they are grounded in game theory. It follows that an ETC would adhere to them. Even if it intended to destroy a target civilization imminently, it would spy before attacking. Otherwise, it might be unpleasantly surprised by what Donald Rumsfeld called “unknown unknowns.” It would also spy if (for reasons discussed below) it did not intend an imminent attack. In that case, close monitoring would ensure the target did not become dangerous unexpectedly, either through rapid technological development or by allying itself with some stronger power. Large and powerful countries routinely spy on small, weak ones for reasons of this sort.

These considerations suggest that ETC would engage in a large-scale search for pre-technological, and therefore potentially threatening, civilizations, and would covertly gather information from any it discovered. Super von Neumann probes that use resources found at their destination to create other specialized probes for surveillance, communication, weaponry, or other purposes would likely play the leading role in this effort. Probes situated near a target planet but concealing their presence – sometimes called “lurkers” – would have significant advantages over longer-distance observation [82-83]. Even a highly advanced ETC could not observe a target in detail at astronomical distances. Sophisticated probes could obtain physical samples, view features of the planet unavailable to astronomical observation, and tap communications. And because probes can multiply at no cost to the sending civilization, they would be highly economical. Finally, the time needed to react to threatening developments at the target could be greatly shortened by the maintenance of lurking probes, so long as they had authority to act.

A disadvantage of stationing probes near a target is that they might disclose ETC’s existence. If Murphy’s law applies to ETCs – and why should it not? – lengthy observation could result in some probes being observed or even captured. However, a technological gap between ETC and its target would make reverse engineering difficult. A captured probe is unlikely to create a serious threat unless it discloses ETC’s home location.

*5.2.*  *Non-strategic information gathering*

To see why continuous surveillance of a target might be preferred to prompt destruction, the value of non-strategic information must be considered. For our purposes, scientific, commercial, cultural and many other forms of information are non-strategic regardless of how they are treated within the target’s economy. The question, then, is whether a target planet would typically possess or produce non-strategic information the discovering ETC would value. If it did, and if the ETC could extract it without divulging strategic information, that would be a reason not to destroy the target immediately.

In accordance with Cabrol’s principle of the coevolution of life and environment, geology [84], biology, culture and technology will differ significantly from one planet to another. Thus, even a target far less advanced than the discovering ETC could almost always supply something new. Earth’s biosphere (for example) is a vast storehouse of information. Everything about it may be new to a wholly different form of life, or specific bits may have particular value. Just consider George de Mestral, who invented Velcro® after burdock seeds clung to his woolen socks and coat. On a planet without hooked seeds, an equivalent product might never be invented.

On Earth, indigenous societies often possess forgotten, bypassed or undiscovered techniques and information unknown to societies with higher technologies [85-86]. On the principle of the coevolution of life and environment, this observation should hold with greater force among interstellar civilizations. The differences among planetary technologies may sometimes be fundamental. Imagine an ETC that employs purpose-grown biologics to perform many tasks performed on Earth by machines. A wide range of mechanical inventions might be new and useful to it.

Practicality aside, the art, music, literature, and history of a target civilization may appeal for their novel experiential qualities. In these fields, there seems to be no hierarchy of more and less advanced societies, only differences that spark new ways of thinking when encountered [87]. A target civilization would also provide data an ETC could use to model the patterns of economic, political, and technological evolution that govern societies generally [see 85-86]. The resulting enhancement to ETC’s ability to predict the behavior and course of development of other target or predator civilizations could confer a strategic advantage. Finally, old societies are, as David Bohm [90, p. 22] notes, prone “to accumulate all sorts of misinformation,” information and opinions that are functional in certain ways but also limiting. Often, he writes: “The society is blocked because misinformation is held *rigidly*.” Nothing seems more likely to reduce rigidity than exposure to a wholly different view.

We cannot know which products, features or technologies of one planet would be useful or interesting to another, but a planet large and varied enough to produce a target civilization would possess many candidates. An ETC would be incentivized to obtain any knowledge that might be of future benefit; anything left behind might be just the bit of information needed to address a future problem or even to defend its existence. Conquerors have taken goods and artifacts throughout human history; an ETC could achieve a similar result in a covert, non-destructive way. Everything of potential value could be reduced to its information content and sent to the ETC’s home planet. There ETC science could test its theories against this new set of facts, perhaps refining them in the process.

We should not imagine the extraction of information from a target civilization as a “one and done” event. Any target is likely to be a long-lived “producing well” of valuable, stimulating or merely interesting knowledge. Jane Goodall’s 60-year study of chimpanzees was lengthy from a human perspective, but a two millennia study of a target planet might be very brief for an ETC.

Even if a target posed no imminent threat, the extraction of non-strategic information would still proceed stealthily. If alerted to ETC’s presence, a target civilization might try to prevent the acquisition of its information, demand knowledge in return, or even attempt to locate ETC’s home. Disclosure could also affect the path of the target’s development, making its non-strategic information less unique and therefore less valuable [see 91-92].

This line of thinking resembles Ball’s [91] “zoo hypothesis” but avoids a major objection to it. That hypothesis presumes a diversity of extraterrestrial cultures, so it would take just one to breach the agreement of non-interference [93]. On the dual-goals hypothesis, only one advanced civilization would be active at the same time in the same region; if another were to supplant it, the motives and behavior of the new entity would, *ex hypothesi*, be the same as the old.

*5.3. One Possible Strategy*

Any attempt to describe ETC’s strategy in more detail than this must be undertaken with humility. We do not know what they know or how they think. We can, however, outline one possible strategy while acknowledging that others may appeal to them more. Much as NASA plans to deploy constellations of satellites in low Earth orbit to detect and deflect potentially dangerous astronomical objects [61], an ETC could deploy constellations of super von Neumann probes to detect potentially threatening developments. “First-order” probes could be stationed at or near astronomical bodies that would provide energy and resources but would host no interfering life forms. The diameter of our galaxy’s stellar disc is ~ 105 light years. Given an ETC near the galactic center [40, 94] with a capacity to scatter super von Neumann probes at an average speed (including regeneration time) of 0.01 c, the scattering of first-order probes at throughout the disc could be completed in ~ 5 million years.[[4]](#footnote-4) Deployed probes would begin to return information much sooner. This is not an implausible time frame. As noted above, the modal habitable planet in our galaxy is ~2 *billion* years older than Earth, the lifespan of a surviving ETC could easily exceed that of its original star, and its sensitivity to future threats and readiness to address them would far exceed our own.

After assembly on site, each first-order probe could passively surveil and assess planets within its sphere, then dispatch second-order probes to those that seemed to require it. Ideally, these would be dispatched *before* a technological civilization arose on the target planet. An ancient and advanced ETC would become familiar with the planetary characteristics most conducive to the emergence of technological civilizations; on Earth, Eurasia’s size and shape seems to have played a leading role [88]. Evidence of seasonal agriculture and related activities could provide a further warning [96-97], with signatures of industrialization providing later alerts [98]. Even evidence that a planet is replete with complex life might spark a preliminary investigation. Chemical indications of life have been observed on a planet 124 light years from Earth [7]; ETC could observe from a greater distance. The probes employed in this project could be controlled by AGI or by life forms created on site [99].

A second-order probe’s initial mission would be to determine the target’s dangerousness. If it was not imminently dangerous, it would be mined for non-strategic information for as long as safety allowed. If a target advanced too far, scientifically and technologically, ETC’s interest in self-preservation would come to outweigh its informational value; at that point, the threat would be terminated.

 *6. Predictions of the Hypothesis*

Whether ETC adopts the foregoing strategy or some better one, the dual-goals hypothesis predicts the potentially observable consequences described in this section.

*6.1.* *We will receive no intentional “lighthouse” signal*.

Cocconi and Morrison [100, p. 844] suggested that ETC had, “established a channel of communication that would one day become known to us, and … they look forward patiently to the answering signals … which would make known to them that a new society has entered the community of intelligence.” However, this “lighthouse” [101, p. 1] theory attributes its authors’ motives and values to an unknown entity, and thereby falls prey to what intelligence analysts call the “mirror image fallacy” [102]. An ETC that sees itself as a denizen of the “dark forest” would not emit a lighthouse signal. While others may emit such signals, the time that elapses between the moment a civilization gains the capacity to signal and the moment it becomes a threat to some monitoring civilization may (as humanity’s advance since the 1950’s suggests) be short in cosmic terms. We are unlikely to observe signals sent over such a short time.

*6.2.* *We will observe no evidence of widespread colonization*.

Discussions of ETC behavior often assume that some would engage in *widespread colonization*, defined as colonization beyond the minimum needed to hedge against local risks, including the death of its home star. Yet, no evidence for this behavior has been found and no compelling motive for it has been advanced.

Hanson and colleagues [103, p. 6] justify their colonization model by reference to Earthly examples of “competing species, cultures, and organizations … expand[ing] into new territories and niches … when[ever] such new territories offer supporting resources that can aid reproduction.” In this model, resources are not extracted from the new territories for the benefit of the metropole; the territories simply provide the soil in which colonies grow. But this story conflates the migration of settler groups – plant, animal, or human – into nearby territories with the sort of large-scale program needed to support interstellar colonization. A true colonization program would require support from the metropole, and a metropole would have no reason to provide support unless it or a leading group in it expected to benefit.

Human metropoles have never undertaken large-scale colonization programs for the benefit of the colonists. Instead, such programs have been driven by two primary motives: acquiring resources (spices, minerals, etc.) and enhancing the metropole’s power, prestige, and security. These motives comport with the convergent, instrumental goals described above, but they are unlikely to motivate interstellar colonization. As noted above, the cost of acquiring and transporting material resources from distant planets would almost surely exceed the cost of producing them at home, while novel information could be obtained without colonization. Russia’s perennial quest to colonize neighboring lands provides one analogy of colonization for a security motive. However, invaders had to fight their way through Russia’s colonized lands, which therefore provided a buffer. A stealthy, technologically superior invader might pass among colonized planets without notice or interference.

A third motive for human colonization has been the removal of undesirable individuals from the metropole (*e.g*., from England to Australia), but an advanced ETC seems unlikely to undertake interstellar colonization for that purpose. Nor does it seem likely that religious belief would motivate interstellar colonization. More broadly, widespread colonization seems an unlikely endeavor for an old, necessarily stable society [104, pp. 59-60], especially as it would come with security risks. Since the discovery of a single colony would likely disclose the locations of the rest, a colonial empire of 100 planets would be 100 times more observable, hence more vulnerable, than a single planet. Colonization could also create its own “dark forest.” Colonies hundreds or thousands of light years distant from each other would grow apart genetically, culturally and commercially. Intermittent exchanges would not prevent this, as they have not prevented wars among Earth’s human populations. A “Hobbesian predicament in which all actors are perpetually in fear of being destroyed” could easily result [105, p. 74], especially if colonies generated colonies of their own.

*6.3. ETC is likely to be active on or near Earth*.

The *extraterrestrial hypothesis* (“ETH”) asserts that some unidentified anomalous phenomena (“UAP”) are best explained by reference to an extraterrestrial intelligence or civilization. UAP can be broadly defined as “sources of detection of anomalous detections … that are not yet attributable to known actors and that demonstrate behaviors that are not readily understood by sensors or observers” [114, p. 14]. The ETH is a narrower claim than the claim that ETC is active on or near Earth: even if no UAP has an extraterrestrial explanation, ETC could still be here. We will, however, assume their equivalence here.

The dual-goals hypothesis, if sound, would probabilistically support the ETH in this sense: For the latter *not* to follow from the former, one of three unlikely things would need to be true: Either a) we are the first spacefaring civilization ever to be active in our galaxy, or b) our predecessors no longer exist, or c) one or more exist but have not reached us yet. Each of these conditions is improbable because:

a) For reasons described at the outset, it is hard to imagine that we are the only spacefaring civilization *ever* to evolve in this galaxy or to enter it from elsewhere;

b) There is little reason to think a spacefaring civilization would destroy itself or be destroyed by natural events,[[5]](#footnote-5) and if it was a dark forest victim its predator would replace it; and

c) ETC would likely be quite old, so it would have had plenty of time to search.

Objections to the ETH are of two types: *A priori* objections cite the difficulties of interstellar travel and the absence of a reason for ETC to travel to Earth; the objection *a* *posteriori* is the absence of compelling evidence for its presence. The dual-goals hypothesis addresses both.

 *6.3.1.* *A priori objections*.

It has long been known that distance is no barrier to interstellar travel. Even at modest speeds, there has been ample time for civilizations to traverse the galaxy many times over. Several forms of propulsion seem physically possible given our current knowledge [47], and more exotic options may exist [*e.g*., 56, 106]. *A priori* objections are therefore not grounded in physical impossibility but in other questions. How would ETC learn of Earth? Why would it expend resources to come here? Once here, why would it not disclose its presence?

 ETC would learn of Earth because it would search for planets that might harbor a current or future threat and/or provide copious non-strategic information. That search might be costly, but a) human societies have typically been willing to bear very heavy costs when their existence is at stake; b) on a strategy like that laid out in section 4.3, these would be mostly one-time costs; and c) these costs would yield a return in the form of information.

ETC would come to Earth for the reasons discussed above. We are, it is sometimes argued, too insignificant to merit the attention of an advanced civilization. This might be true if human science and technology was locked in its current state or if humans were the only source of information on the planet, but neither is the case. Advances in human knowledge occur almost weekly, and Earth’s biosphere may contain as much information as its noosphere. Most crucially, artificial intelligence is progressing rapidly, and with it our threat capacity. A survey of researchers assigns a 50% or greater probability to the claim that AI will outperform humans in nearly all occupational tasks by 2060 [54]. As these tasks would include computer design and programming, an advanced society might believe humanity would advance “superexponentially” [55] and soon become an imminent threat.

ETC would not disclose its presence because, though we are currently no threat, we would likely interfere with their activities if we became aware of them. We might also use what we learn from their activities to improve our own technology. Perhaps most crucially, if we knew they were here, we could try to discern their origin. Finally, our independent cultural path could be deflected, making the information ETC obtains from us less unique and thus less valuable.

*6.3.2. Objection a posteriori*

The ETH is supported by no dispositive evidence, but that tells us little. If ETC were here, it would understand human languages, history, technology and behavior quite well; indeed, gaining that understanding would be central to its mission. Together with this knowledge, its advanced technology and powers of cognition would make it highly capable of concealment and strategic deception. Deception aims to confuse an opponent’s decision makers about one’s intentions and activities [79]. An expert on the topic writes, “[E]mpirical evidence confirms assumptions drawn from cognitive psychology that deception seldom fails when it exploits a target’s preconceptions. The target’s tendency to assimilate discrepant information to existing mental sets generally negates the risks to deception posed by security leaks and uncontrolled channels of information” [106, p. 294]. ETC would know that, of humanity’s two relevant elites, the scientific community tends to reject “extraordinary” claims unless clear and convincing evidence supports them,[[6]](#footnote-6) and its militaries tend to focus on known threats.[[7]](#footnote-7) Its deception strategy would make use of these attitudes. By contrast, our knowledge of its capacities and practices would be negligible.

Though no dispositive evidence for the ETH exists, anything that tends to make a relevant fact more or less probable counts as evidence in a court of law.[[8]](#footnote-8) By this standard, well-attested witness reports, photographs, radar returns, and other physical effects that i) an alien presence could explain and that ii) after investigation, have no other persuasive explanation count as some evidence for this hypothesis. Evidence of this sort certainly exists [113-116]; the challenge is how to evaluate it. This subsection will address that challenge in the absence of the dual-goals hypothesis.

To understand it clearly, it will help to view it from a Bayesian perspective. Let:

P(**¬**B) = the probability that the *report* of a hypothetical anomalous event is a false positive, *e.g*., a hoax or a mistake.

P(BǀA) = the likelihood function of the ETH on the event, assuming it is not a false positive.

P(A) = the prior probability of the ETH, absent this event.

P(AǀB) = the posterior probability of the ETH, given this event.

In a hypothetical case, set P(**¬**B) at 0.05 (signifying that the event is well-attested) and P(BǀA) at 0.5, an “uninformative” value. On these values, assigning P(A) a value of 0.01 yields a posterior probability – P(AǀB) – of 9.2% while assigning P(A) a value of 0.1 yields a posterior probability of 52.6%. The significant weight the Bayesian calculus gives to prior probability – corroborates the aphorism that “extraordinary claims” – claims taken to be extraordinary because they are unsupported by an accepted theory – are said to require “extraordinary evidence” [112].

What of the other Bayesian factors? Applying normal evidentiary considerations – *e.g*., an event independently reported from different perspectives by technical systems and multiple observers [see 108, 113-115] would, all else being equal, be better attested than one evidenced by a photograph from a single witness – expert judgment can assess the P(**¬**B) of an event, perhaps using the Delphi method [117]. As to its P(BǀA), *some* alternative, non-ETC explanation is likely to be proffered for most reports of anomalous events. In some cases, this explanation will seem persuasive; in others, it will not. Again, the Delphi method could be used to determine the probability of this alternative explanation, call it P(AE). But subtracting P(AE) from 1 will not yield the probability of an extraterrestrial explanation, because the true explanation may be non-extraterrestrial but unknown. Denoting an unknown, non-extraterrestrial explanation as X:

P(BǀA) = 1 – {P(AE) + P(X)}

This would not be a problem if we could independently determine P(X). But the value we assign to it will surely turn on the value we assign to P(A), the prior probability of the extraterrestrial hypothesis. That is, if we assign a low value to P(A), we will likely be willing to assign a high value to P(X). If we assign a high value to P(A), we will be more likely to constraint P(X) to a lower value. But if everything depends on P(A), which is unknown, even low values of P(AE) and P(**¬**B) will tell us nothing. This is the central epistemic challenge to any evaluation of proffered evidence of ETC’s presence. Can the dual-goals hypothesis meet it?

*6.3.3. The epistemic challenge and the dual-goals hypothesis*

Thus far, the arguments for the dual-goals hypothesis have only been its intrinsic logic and the fact that it explains the Fermi paradox. We turn now to the question of whether it might be tested using facts we have on hand.

The hypothesis predicts specific behaviors an ETC operating on Earth would likely exhibit and others it would likely avoid. These include the following:

● Since, if ETC is here at all, it is likely to have been here for some time it would already have collected information created in the past. Its current efforts would focus on newly created information and therefore on locations where that is likely to arise.

● ETC would monitor humanity’s most advanced weapons systems, and would attempt to control them if it could do so without active hostility. Active hostility would pose an immediate threat and could engender a response.

● ETC would expose itself to observation as little as possible, no more than is necessary to pursue its mission. Lights in the night sky, objects that intentionally present themselves to multiple observers, and other widely apparent events are therefore highly unlikely to be ETC-related.

● For the same reason, ETC would not routinely reappear in the same observable locations.

● ETC would not share information or technology with humanity or attempt to affect the course of human history at any point. *Pace* various claims about ancient aliens, it would not show humans how to carve stone.

● Events in which human observers stumble upon or surprise a UAP are more likely to have an extraterrestrial explanation than events in which a UAP enters a space clearly occupied by human observers.

● Insofar as possible consistent with its mission, ETC would avoid urban areas or other environs where human observers are plentiful.

● ETC’s technology would *occasionally* malfunction or fail to operate properly; its technology would only be technology, not magic.

● ETC would *occasionally* engage in radar or visual *spoofing* (transmission of signals intended to deceive).

If we accept these behaviors (or others) as its logical consequences, the dual-goals hypothesis can be tested statistically. Set the P(A) of the ETH at some arbitrary value, then estimate the P(**¬**B) and P(AE) of a representative set of UAP reports selected without reference to these behaviors. As noted above, both should be determinable. Then, setting P(X) to zero, calculate the P(BǀA) and P(AǀB) of each report. These values will be meaningless in themselves but can be ranked from highest to lowest. Next, determine whether each report exhibits behaviors predicted above. A better-than-chance correlation between reports that generate a high P(AǀB) and those in which the UAP exhibits these behaviors would tend to confirm the dual-goals hypothesis and with it the ETH; the lack of any positive correlation would tend to falsify the former and to a degree the latter.

*7. Conclusion*

The dual-goals hypothesis would, if accepted, resolve the Fermi paradox. It also offers a way to test the ETH using data already at hand. But it does these things at a cost: If it is true, it strongly suggests that humanity faces an unsuspected existential risk. Whether this risk exists or not, the possibility is worthy of investigation.

***References***

Kunimoto, Michelle and Jaymie Matthews (2020) Searching the Entirety of Kepler Data: II. Occurrence Rate Estimates for FGK Stars, *The Astronomical Journal* **159**:6. <https://doi.org/10.3847/1538-3881/ab88b0>.

Schaefer, Laura and Dimitar Sasselov (2015) The Persistence of Oceans on Earth-like Planets: Insights from the Deep-Water Cycle, *The Astrophysical Journal* **801**:1, 40. <http://dx.doi.org/10.1088/0004-637X/801/1/40>.

Heller, René and John Armstrong (2014) Superhabitable Worlds, *Astrobiology* **14**:1.

Mahulikar, Shripad and Heinz Herwig (2009) Exact Thermodynamic Principles for Dynamic Order Existence and Evolution in Chaos, *Chaos, Solitons and Fractals* **41**:4, 1939-1948. <https://doi.org/10.1016/j.chaos.2008.07.051>.

Xavier, Joana and Stuart Kauffman (2022) Small-molecule autocatalytic networks are universal metabolic fossils, *Philosophical Transactions A* **380**:2227. <https://doi.org/10.1098/rsta.2021.0244>.

Whitmire, Daniel (2022) Abiogenesis: The Carter Argument Reconsidered, *International Journal of Astrobiology* **22**:2, 94-99. <https://doi.org/10.1017.S1473550422000350>.

Madhusudhan, Nikku *et al*. (2023) Carbon-bearing Molecules in a Possible Hycean Atmosphere, <https://arxiv.org/abs/2309.05566>.

Greenbury, Sam, Ard Lewis and Sebastian Ahnert (2022) The structure of genotype-phenotype maps makes fitness landscapes navigable, *Nature Ecology & Evolution* **6**, 1742-1752.

1. Knoll, Andrew and Richard Bambach (2019) Directionality in the history of life: Diffusion from the left wall or repeated scaling of the right? *Paleobiology* **26**:S4, 1-14. <https://www.jstor.org/stable/1571651>.

Wong, Michael *et al*. (2023) On the roles of function and selection in evolving systems, *PNAS* **120**:43, e2310223120. <https://doi.org/10.1073/pnas.2310223120>.

1. Doolittle, W. Ford (2020) Is the Earth an organism? *Aeon* [The Gaia hypothesis reimagined by one of its key sceptics | Aeon Essays](https://aeon.co/essays/the-gaia-hypothesis-reimagined-by-one-of-its-key-sceptics).

Bourrat, Pierrick (2023) A Pricean Formalization of Gaia, *Philosophy* *of* *Science* **90**:3, 704-720. <https://doi.org/10.1017/psa.2023.44>.

Kasting, James (2010) *How to Find a Habitable Planet.* Princeton.

Schulze-Makuch, Dirk and William Bains (2017) *The Cosmic Zoo*. Springer.

1. Roth, Gerhard (2015) Convergent evolution of complex brains and high intelligence, *Philosophical Transactions B* **370**, 20150049.<https://doi.org/10.1098/rstb.2015.0049>.

Pinker, Steven and Paul Bloom (1990) Natural Language and Natural Selection, *Behavioral and Brain Sciences* **13**:4, 707-727. <https://doi.org/10.1017/S0140525X00081061>.

1. Corballis, Michael (2017) *The Truth about Language*. Chicago.

Labbé, Ivo *et al*. (2023) A population of red candidate massive galaxies ~600 Myr after the Big Bang, *Nature* **616**, 266-269.

Schroeder, Karl (undated) The Deepening Paradox, [The Deepening Paradox — KarlSchroeder.com (kschroeder.com)](https://www.kschroeder.com/weblog/the-deepening-paradox).

Cabrol, Nathalie (2016) Alien Mindscapes – A Perspective on the Search for Extraterrestrial Intelligence, *Astrobiology* **16**:9, 661-675. <https://doi.org/10.3390/life10040042>.

Freeland, Stephen (2022) Undefining life’s biochemistry: Implications for abiogenesis, *Journal of the Royal Society Interface* **19**:20210814. <https://doi.org/10.1098/rsif.2021.0814>.

Bartlett, Stuart and Michael Wong (2020) Defining Life in the Universe: From Three Privileged Functions to Four Pillars, *Life* **10**:4, 42. <https://doi.org/10.3390/life10040042>.

1. Dick, Steven (2003) Cultural evolution, the postbiological universe, and SETI, *International Journal of Astrobiology* **21**:1, 65-74. <https://doi.org/10.1016/j.actaastro.2008.01.025>.
2. Dȍbler, Niklas and Marius Raab (2021) Thinking ET: A discussion of exopsychology, *Acta* *Astronautica* **189**, 699-711. <https://doi.org/10.1016/j.actaastro.2021.09.032>.
3. Lindsey, Jason (2022) SETI from the Perspective of Intercivilizational Politics, *Acta Astronautica* **61**, 101490. <https://doi.org/10.1016/j.spacepol.2022.101490>.

Kipping, David, Adam Frank and Caleb Scharf (2020) Contact Inequality: First Contact Will Likely Be with an Older Civilization, *International Journal of Astrobiology* **19**:6, 430-437. <https://doi.org/10.1017/S1473550420000208>.

Bostrom, Nick (2014) *Superintelligence: Paths, Dangers, Strategies*. Oxford. Kindle ed.

Russell, Stuart and Peter Norvig (2022) *Artificial Intelligence: A Modern Approach*. Pearson India.

1. Benson-Tilsen, Tsvi and N. Soares (2016) Formalizing Convergent Instrumental Goals, *AAAI-16 Workshop: AI, Ethics,* [FormalizingConvergentGoals.pdf (intelligence.org)](https://intelligence.org/files/FormalizingConvergentGoals.pdf).

Real, Leslie (1991) Animal Choice Behavior and the Evolution of Cognitive Architecture, *Science* **253**:5023, 980-986. https:/doi.org/10.1126/science.1887231.

Rudnyk, Adrian (2021) *The Assessment: The Arrival of Extraterrestrials*. Rudnyk.

Schmidt, G.R. *et al.* (2000) Antimatter Requirements and Energy Costs for Near-term

Propulsion Requirements, *Journal of Propulsion and Power* **16**:5.

1. Markovic, Dragan, *et al*. (2012) Impact of nanotechnology advances in ICT on sustainability and energy efficiency, *Renewable and Sustainable Energy Reviews* **16**:5, 2966-2972. <https://doi.org/10.1016/j.rser.2012.02.018>.
2. Kim, Donggeun and Young-Joon Park (2022) Nonlinear causality between energy consumption and economic growth by timescale, *Energy Strategy Reviews* **44**, 100949. <https://doi.org/10.1016/j.esr.2022.100949>.
3. Barrow, John (1998) *Impossibility: The Limits of Science and the Science of Limits*. Oxford.

Garrett, M.A. (2015) Application of the mid-IR radio correlation to the G sample and the search for extraterrestrial civilizations, *Astronomy & Astrophysics* **581** L5. <https://doi.org/10.1051/0004-6361/201526687>.

Pearce, David (2020) Humans and Intelligent Machines: Co-Evolution, Fusion or Replacement? in Steven Gouvela, ed. *The Age of Artificial Intelligence: An Exploration*. Vernon.

1. Huffman, David, Raimund Maurer and Olivia Mitchell (2019) Time Discounting and Economic Decision Making in the Older Population, *Journal of Economics of Aging* **14**, 100121. <https://doi.org/10.1016/j.jeoa.2017.05.001>.

Giaimo, Stefano and Arne Traulsen (2019) Generation Time Measures the Trade-Off between Survival and Reproduction in a Life Cycle, *American Naturalist* **194**:2. <https://doi.org/10.1086/704155>.

Smart, John (2012) The Transcension Hypothesis, *Acta Astronautica* 78, 55-68. <https://doi.org/10.1016/j.actaastro.2011.11.006>.

Armstrong, Stuart and Anders Sandberg (2013) Eternity in six hours: Intergalactic spreading of intelligent life and sharpening the Fermi paradox, *Acta Astronautica* **89**, 1-13. [https://doi.org/10.1016 Andrew /j.actaastro.2013.04.002](https://doi.org/10.1016%20Andrew%20/j.actaastro.2013.04.002).

Baum, Seth *et al*. (2019) Long-Term Trajectories of Human Civilization, *Foresight* **21**:1, 53-83. <https://dx.doi.org/10.1108/FS-04-2018-0037>.

Murphy, Sean and Anthony Atala (2014) 3D bioprinting of tissues and organs, *Nature Biotechnology* **32**, 773-785. <https://doi.org/10.1038/nbt.2958>.

Neumayer, Eric (2000) Scarce or Abundant? The Economics of Natural Resource Availability, *Journal of Economic Surveys* **14**:3, 307-335. <https://doi.org/10.1111/1467-6419.00112>.

Pooley, Gale and Marian Tupy (2020) Luck or Insight? The Simon-Ehrlich Bet Re-examined, *Economic Affairs* **40**:2, 277-280. <https://doi.org/10.1111/ecaf.12398>.

Lampton, Michael (2013) Information-driven societies and Fermi’s paradox, *International Journal of Astrobiology* **12**:4, 312-313. <https://doi.org/1017/S1453550413000153>.

1. Litchford, Robert and Jeffrey Sheehy (2020) Prospects for Interstellar Propulsion, NASA Preprint AAS 20-068 – 200200000759-1.
2. Loeb, Avi (2021) *Interstellar: The Search for Extraterrestrial Life and Our Future in the Stars*. Mariner.

Borgue, Olivia and Andreas Hein (2021) Near-term self-replicating probes – a concept design, *Acta Astronautica* **187**, 546-556. <https://doi.org/10.1016/j.actaastro.2021.03.004>.

Matloff, Gregory (2022) Von Neumann probes: rationale, propulsion, interstellar transfer timing, *International Journal of Astrobiology* **21**:4, 205-211. <https://doi.org/10.1017/S1473550422000027>.

Qian, Chao and Hongsheng Chen (2021) A perspective on the next generation of invisibility cloaks – intelligent cloaks, *Applied Physics Letters* **118**, 180501.

Kipping, David and Alex Teachey (2016) A Cloaking Device for Transiting Planets, *Monthly* *Notices of the Royal Astronomical Society* **459**:2. <https://doi.org/10.1093/mnras/stw672>.

1. Fitzgerald, McKenna, Aaron Boddy and Seth Baum (2020) *2020 Survey of Artificial General Intelligence Projects*, Global Catastrophic Risk Institute.

Grace, Katja *et al.* (2018) Viewpoint: When Will AI Exceed Human Performance? Evidence from AI Experts, *Journal of Artificial Intelligence Research* **62**, 729-754.

Nagy, Béla (2011) Superexponential Long-term Trends in Information Technology, *Technological Forecasting and Social Change* **78**:8, 1356-1364. <https://doi.org/10.1016/j.techfore.2011.07.006>.

Lentz, Erik (2022) Hyper-fast positive energy warp drives, arXiv:2201.00652 (gr-qc).

Abellan, Gabriel, Nelson Bolivar and Ivaylo Vasilev (2023) Alcubierre warp drive in spherical coordinates with some matter configurations, *European Physical Journal C* **83**, 7. <https://doi.org/10.1140/epjc/s10052-022-11091-5>.

National Research Council (U.S.) (2007), *A Strategy for Assessing Science*, I. Feller and P.C. Stern, eds. National Academies Press.

Parfit, Derek (1984) *Reasons and Persons*. Clarendon.

Bostrom, Nick and Milan Ćirković, eds. (2008) *Global Catastrophic Risks*. Oxford.

NASA (2023) NASA’s Planetary Defense Strategy and Action Plan, NASA Releases Agency Strategy for Planetary Defense to Safeguard Earth | NASA.

1. Liu Cixin (2015) *The Dark Forest*. Tor.

Panskepp, Jules and Garet Lahvis (2011) Rodent empathy and affective neuroscience, *Neuroscience and Biobehavioral Reviews* **35**:9, 1864-1875. <https://doi.org/10.1016/j.neubiorev.2011.05.013>.

1. Dothan, Shai (2021) When Immediate Responses Fail, *Vanderbilt Law Review* **51**, 1075.

Kurzweil, Ray (2005) *The Singularity Is Near*. Penguin.

Yasser, Shehab (2020) Aliens, the Fermi Paradox, and the Dark Forest Theory, *Towards Data Science,* [Aliens, The Fermi Paradox, And The Dark Forest Theory: A Game Theoretic View | by Shehab Yasser | Towards Data Science](https://towardsdatascience.com/aliens-the-fermi-paradox-and-the-dark-forest-theory-e288718a808).

Susskind, Ron (2006) The Untold Story of al-Qaeda’s Plot to Attack the Subways (June 26) *Time*.

Yu, Chao (2015) The Dark Forest Rule: One Solution to the Fermi Paradox, *Journal of the British Interplanetary Society* **68**:5-6, 142-144.

Webb, Stephen (2015) *If the Universe Is Teeming with Aliens, Where Is Everybody?* 2nd ed. Springer.

1. Dall, Sasha *et al*. (2005) Information and its use by animals in evolutionary ecology, *TRENDS in Ecology and Evolution* **20**:4. <https://doi.org/10.3390/life10040042>.
2. Cervera, Roberto, Maya Wang and Benjamin Harden (2020) Systems neuroscience of curiosity, *Current Opinion in Behavioral Sciences* **35**, 48-55. <https://doi.org/10.1016/j.cobeha.2020.06.011>.
3. Burke, James (1996) *The Pinball Effect*. Little, Brown.
4. Jaegle, Andrew, Vahid Mehrpour, and Nicole Rust (2019) Visual novelty, curiosity and intrinsic reward, *Current Opinion in Neurobiology* **58**, 167-174. <https://doi.org/10.1016/j.conb.2019.08.004>.

Ekstig, Börje (2015) Complexity, Natural Selection and the Evolution of Life and Humans, *Foundations of Science* **20**:2, 175-187. [https://doi.org/10.1007%2Fs10699-014-9358-y](https://doi.org/10.1007/s10699-014-9358-y).

Wiseman, Charles (1988) *Strategic Information Systems*. Richard Irwin.

1. Hillier, Brian (1997) *The Economics of Asymmetric Information*. Macmillan.

Andrew, Christopher (2019) *The Secret World: A History of Intelligence.* Yale.

Omand, Dennis (2015) *Securing the State*. Oxford.

Sun Tzu (2022) *The Art of War*, trans. Lionel Giles. Union Square.

Gordon, Sue (June 28, 2023) If we cannot trust our former presidents, whom can we trust? *Washington Post*.

Solan, Eilon and Leeat Yariv (2004) Games with Espionage, *Games and Economic Behavior* **44**:1, 172-199. [https://doi.org/10.1016/S0899-8256(03)00177-5](https://doi.org/10.1016/S0899-8256%2803%2900177-5).

Gertz, John (2016) ET Probes: Looking here as well as there, <https://doi.org/10.48550/arXiv.1609.04635>.

Benford, James (2019) Looking for Lurkers: Co-Orbiters as SETI Observables, *The Astronomical Journal* **158**:4, <https://doi.org/10.3847/1538-3881/ab3e35>.

Hystad, Grethe *et al*. (2015) Statistical analysis of mineral diversity and distribution: Earth’s mineralogy is unique, *Earth and Planetary Science Letters* **426**, 154-157. <https://doi.org/10.1016.j.epsi.2015.06.028>.

1. Jessen, Tyler, *et al*. (2022) Contributions of Indigenous Knowledge to ecological and evolutionary understanding, *Frontiers in Ecology and Environment* **20**:2, 93-101. <https://doi.org/10.1002/fee.2435>.
2. Johnson, Lydia *et al.* (2023) Weaving indigenous and western ways of knowing in ecotoxicology and wildlife health, *Environmental Reviews* <https://doi.org/10.1139/er-2022-0087>.

Brooks, Dorothy (1956) The Influence of African Art on Contemporary European Art, *African Affairs* **55**:218, 51-59.

Diamond, Jared (1997) *Guns, Germs and Steel*. W.W. Norton.

1. Turchin, Peter (2018) *Historical Dynamics: Why States Rise and* *Fall*. Princeton.
2. Bohm, David (1989) Meaning and Information, in *The Search for Meaning: The New Spirit in Science and Philosophy*, Paavo Pylkkanen, ed. Crucible, 43-62.
3. Ball, John (1973) The zoo hypothesis, *Icarus* **19**:3, 347-349.

Fogg, Martyn (1986) Temporal Aspects of the Interaction among the First Galactic Civilizations: The “Interdict Hypothesis,” *Icarus* **69**, 370-384.

Forgan, Duncan (2011) Spatio-temporal constraints on the zoo hypothesis, and the breakdown of total hegemony, *International Journal of Astrobiology* **10**:4, 341-347.

Gowanlock, M.G., D.R. Patton and S.M. McConnell (2011) A Model of Habitability within the Milky Way Galaxy, *Astrobiology* **11**:9, 855-73. <https://doi.org/10.1089/ast.2010.0555>.

Loeb, Abraham and Sean Kirkpatrick (2023) Physical Constraints on Unidentified Aerial Phenomena (Under Review). [LK1.pdf (mentealternativa.com)](https://www.mentealternativa.com/ma_media/2023/03/LK1.pdf).

1. Lockley, Andrew and Daniele Visioni (2020) Detection of pre-industrial societies on exoplanets, *International Journal of Astrobiology* **20**:1, 73-80. <https://doi.org/10.1017/S1473550420000361>.

Boucher, Anne *et al.* (2021) Characterizing Exoplanetary Atmospheres at High Resolution with SPIRou, *The Astronomical Journal* **162**:6, 233.

1. Lingam, Manasvi *et al*. (2023) Technosignatures: Frameworks for their Assessment, *The Astrophysical Journal* 943:1, 27.
2. Hein, Andreas and Stephen Baxter (2018) Artificial Intelligence for Interstellar Travel,

[1811.06526.pdf (arxiv.org)](https://arxiv.org/pdf/1811.06526.pdf).

Cocconi, Giuseppe and Philip Morrison (1959) Searching for Interstellar Communications, *Nature* **184**, 844-846.

1. Hippke, Michael (2021) Searching for interstellar quantum communications, *The Astronomical Journal* **162**:1. <https://doi.org/10.3847/1538-3881/abf7b7>.
2. Heuer, Richards (2019) *The Psychology of Intelligence Analysis*. Martino.
3. Hanson, Robin, *et al*. (2021) If Loud Aliens Explain Human Earliness, Quiet Aliens Are Also Rare, *The Astrophysical Journal* **922**:182. <https://doi.org/10.3847/1538-4357/ac2369>.

Bainbridge, William (2018) *Computer Simulations of Space Societies*. Springer.

Torres, Phil (2018) Space colonization and suffering risks: Reassessing the “maxipok rule,” *Futures* **100**, 74-85. <https://doi.org/10.1016/j.futures.2018.04.008>.

1. Heuer, Richards (1981) Strategic Deception and Counterdeception: A Cognitive Process Approach, *International Studies Journal* **25**:2, 294-327. <https://doi.org/10.2307/2600359>.

Kharif, Christian *et al.* (2009) *Rogue Waves in the Ocean.* Springer.

1. Gounelle, Matthieu (2006) The meteorite fall at L'Aigle and the Biot report: exploring the cradle of meteoritics, in G.J.H. McCall, *et al*., eds. *The History of Meteoritics and Key Meteorites Collections: Fireballs, Falls, and Finds*. Geological Society of London, 73-89. <https://doi.org/10.1144/GSL.SP.2006.256.01.03>.

Seligman, Lara (2023, November 8) Pentagon boss to step down next month, *Politico*.

Achinstein, Peter (2003) *The Book of Evidence*. Oxford.

McMahon, Sean (2020) Do Extraordinary Claims Require Extraordinary Evidence? in *Conceptual Issues in Astrobiology*, Kelly Smith and Carlos Mariscal, eds. Oxford, 117-129.

McCulloch, M.E. (2018) Propellant-less Drive from Quantum Inertia, *Journal of Space Exploration* 7:3, 151.

Office of the Director of National Intelligence (“ODNI”) (2021) *Preliminary Assessment of Unidentified Aerial Phenomena*. [Prelimary-Assessment-UAP-20210625.pdf (dni.gov)](https://www.dni.gov/files/ODNI/documents/assessments/Prelimary-Assessment-UAP-20210625.pdf).

 ----- (2023) FY 2023 Consolidated Report on Unidentified Anomalous Phenomena.

Knuth, Kevin, Robert Powell, Peter Reali (2019) Evaluating Flight Characteristics of Anomalous Unidentified Aerial Vehicles in the 2004 Nimitz Encounter, *Proceedings (MDPI)* **2019**:33, 26. <https://doi.org/10.3390/e21100939>.

Graff, Garrett (2023) *UFO*. Avid Reader.

Okoli, Chitu and Suzanne Pawlowski (2004) The Delphi method as a research tool: an example, design considerations and applications, *Information & Management* **42**:3, 15-29. <https://doi.org/10.1016/j.im.2003.11.002>.

1. It is unlikely that a post-biological ETC would reproduce at all. [↑](#footnote-ref-1)
2. This does not mean an ETC would engage in widespread colonization. See section 7.2. [↑](#footnote-ref-2)
3. “Probe” refers to any uncrewed system that might be delivered to a location and operate effectively there. A von Neumann probe reproduces itself using local materials; a super von Neuman probe is capable of producing a range of useful probes and instrumentalities as well as reproducing itself. [↑](#footnote-ref-3)
4. Given an initial launch of 50 probes, each capable of producing and launching ten more, the galactic disc could be occupied by probes, each at the center of a non-overlapping sphere 100 light years in radius, in three replications. As an alternative strategy, motherships could drop off first-order probes at many locations [95]. [↑](#footnote-ref-4)
5. Of the “global catastrophic risks” discussed in reference 57, only four seem possibly existential. Of these, humans have already ameliorated the risk from planet-killing astronomical objects, and the risk from superintelligent AGI, while potentially existential for its biological creators, would leave the galactic threat environment unchanged. Only nuclear war and molecular manufacturing seem capable of taking a civilization off the threat board entirely, and neither seems likely to do that. Other astronomical risks (*e.g*., a nearby supernova) would be avoidable once a civilization attains a capacity for interstellar colonization. [↑](#footnote-ref-5)
6. One example of the scientific attitude regarding “extraordinary” claims concerns “rogue” ocean waves. These were reported for centuries, but science largely denied their existence before 1995 [107]. Meteors [108] and continental drift followed the same pattern. [↑](#footnote-ref-6)
7. The Director of the All-domain Anomaly Resolution Office says this about its findings: “If we don't prove it’s aliens, then what we're finding is evidence of other people doing stuff in our backyard. And that's not good” [109]. It may be worse if it *is* aliens. [↑](#footnote-ref-7)
8. See Rule 401, Federal Rules of Evidence. Achinstein’s “potential evidence” [110] is evidence by this standard. [↑](#footnote-ref-8)