Universalist Ethics in Extraterrestrial Encounter

Seth D. Baum, http://sethbaum.com
Department of Geography & Rock Ethics Institute, Pennsylvania State University

Published in: *Acta Astronautica*, 2010, 66(3-4): 617-623

Abstract

If humanity encounters an extraterrestrial civilization, or if two extraterrestrial civilizations encounter each other, then the outcome may depend not only on the civilizations' relative strength to destroy each other but also on what ethics are held by one or both civilizations. This paper explores outcomes of encounter scenarios in which one or both civilizations hold a universalist ethical framework. Several outcomes are possible in such scenarios, ranging from one civilization destroying the other to both civilizations racing to be the first to commit suicide. Thus, attention to the ethics of both humanity and extraterrestrials is warranted in human planning for such an encounter. Additionally, the possibility of such an encounter raises profound questions for contemporary human ethics, even if such an encounter never occurs.

Keywords: extraterrestrials, ethics, universalism

1. Introduction

To date, humanity has never encountered extraterrestrial life, let alone an extraterrestrial civilization. However, we can also not rule out the possibility that such an encounter will occur. Indeed, insights from the Drake equation (see e.g. [1]) suggest that such an encounter may be likely. As human exploration of space progresses, such an encounter may become increasingly likely. Thus analysis of what would happen in the event of an extraterrestrial encounter is of considerable significance. This analysis is particularly important for the astronautics community to consider given that it is on the leading edge of space exploration.

There is extensive debate on what would be the outcome of an encounter between humanity and an extraterrestrial civilization. This debate can also be extended to consider encounters between two extraterrestrial civilizations. Much of the debate centers on the moral character of the extraterrestrials and the significance of this for how humanity would fare in such an encounter. Several commentators have speculated that the extraterrestrials would be benevolent and thus safe to humans [2,3] while others have speculated that the extraterrestrials would be malicious and thus dangerous to humans [4,5]. For broad reviews of the debate, see [1,6].

This paper considers an important set of scenarios, largely overlooked by the existing literature, in which either humanity or the extraterrestrial civilization or both act according to a universalist ethical framework. Universalist ethics roughly refers to ethics

where the two civilizations value specific aspects of each other equally, regardless of which civilization these aspects occur in. Universalist ethics is defined more precisely and elaborated in greater detail in Section 2. Meanwhile, for purposes of this article, civilization can be defined as a system of individuals working towards some common objective. Heterogeneity within a civilization, though undoubtedly important, is beyond the scope of this article.

Ethics in general, and universalist ethics in particular, are important in humanity-extraterrestrial encounters because the outcome of such an encounter will depend not only on the relative strengths of the civilizations (i.e. who would destroy the other in an inter-civilizational war) but also on some specifics of the ethics held by the civilizations. Encounters in which one or both civilizations act according to a universalist framework hold particularly interesting properties. For example, if each civilization acts according to a different universalist framework, then an encounter might lead to a race between the civilizations to be the first one to commit suicide. Section 3 discusses a broad range of encounter scenarios involving universalism.

The particular specifics of universalist ethics possibly held by humanity or extraterrestrials have important implications both for human civilization strategy and for contemporary ethics. The implications for civilizational strategy, discussed in Section 4, are important for humanity's planning for extraterrestrial encounters and its response should such an encounter occur. The basic message is that humanity would be wise to consider extraterrestrials' ethics in addition to their war-fighting strengths, because the ethics can be as an important factor in the outcome of an encounter. The implications for contemporary ethics, discussed in Section 5, hold even if no encounter occurs. In particular, the possibility of extraterrestrial encounter challenges certain forms of anthropocentrism commonly found in contemporary human ethics because extraterrestrials might be superior to humans on the same grounds that humans consider ourselves to be superior to other Earth species.

2. Universalist Ethics

The term *universalist ethical framework* comes from the term *universalism* as used in the philosophy and psychology literatures on human values. The terms *values* and *ethics* can mean different things, although for the purposes of this paper both terms will be taken to mean views about right and wrong and about what should be done. As discussed in the psychology literature, universalism is a type of ethical framework humans might support in which there is great equality. For example, Schwartz and Boehnke define universalism as "Understanding, appreciation, tolerance, and protection for the welfare of all people and for nature (equality, social justice, wisdom, broadminded, protecting the environment, unity with nature, a world of beauty)" [7, p. 239].

For this paper I will employ a slightly different definition of universalism. First, I must review the concept of *intrinsic value*. Intrinsic value is that which is valuable for its own sake, independent of anything else [8]. Intrinsic value is contrasted with *extrinsic value*,

which is anything that is valuable but is not intrinsic value [9]. For example, we might consider human welfare to hold intrinsic value (such as in anthropocentric variations of the utilitarianism ethical framework). In this case, phenomena such as food, clothing, and shelter would hold a form of extrinsic value called *instrumental value*, which is valuable because it causes other value [9]- in this case the intrinsic value of human welfare.

There has been much philosophical debate over the question of whether intrinsic value actually exists or if it is instead only considered to exist by individuals with sufficient cognitive capacity to form such a consideration (e.g. humans) [8]. This question is at the heart of *meta-ethics*, i.e. the study of the nature of ethics and ethical knowledge. Possible answers to this question will not be discussed here because this paper focuses on what different civilizations consider to hold intrinsic value, which is a topic that can be examined independent of any knowledge of what might or might not actually hold intrinsic value. The reason for this focus is to explore what civilizations might do in an encounter. No attempt is made at assessing whether the civilizations might be actually right or wrong in the ethics that they support and in the actions they perform. While they are beyond the scope of this paper, such assessments could be readily made given knowledge of what ethics actually are correct if such knowledge could somehow be achieved.

For the purposes of this paper, a *universalist ethical framework* is an ethical framework in which the phenomena considered to hold intrinsic value hold the same intrinsic value regardless of where or when the intrinsic value occurs. For example, a universalist form of anthropocentric utilitarianism would place the same amount of intrinsic value on all human welfare. Likewise, a universalist form of non-anthropocentric utilitarianism would place the same amount of intrinsic value on all welfare, regardless of what species (or non-species) the welfare occurred in. It should be noted that the types of ethical frameworks considered here are all *consequentialist*, meaning that they only place intrinsic value the consequences of actions. No consideration is given to whether certain actions are fundamentally right or wrong (as in deontological ethics) or to whether what is important is not what actions we perform but is instead the character of who we are (as in virtue ethics). While such ethics are important and have enough support among contemporary humans to merit attention, they require a somewhat different analysis and are beyond the scope of this paper.

Human philosophers have extensively debated the extent to which humans should be universalist. Some argue that we have special relations to ourselves and those near us which justifies non-universalism (see [10] for examples). Others argue that universalism is too demanding and thus while being a universalist may be commendable, it is not morally required [11]. Meanwhile, still others argue that non-universalism is immoral and that we should strive for universalism [12]. A prominent argument for universalism stems from a thought experiment in which we select our ethics as if we don't know which

member of society we are, thereby removing any incentive for non-universalist favoritism [13].[1]

Much of the debate on universalism has existed within anthropocentric ethical frameworks. These frameworks only place intrinsic value on human phenomena (welfare, health, etc.) and thus only debate how intrinsic value should be distributed among humans. However, such anthropocentrism is not unanimously supported. For example, several prominent philosophers have advocated non-anthropocentric forms of utilitarianism, placing equal intrinsic value on the welfare of human and non-human animals [15-17]. Others have called for ecocentric ethics, also placing intrinsic value on non-sentient nature [18,19]. The definition of universalism provided by Schwartz and Boehnke includes ecocentrism, as does other psychology research on environmental values [20].

For the present paper it is crucial for universalism to extend beyond humanity. Specifically, this paper explores universalism with respect to civilizations: whether humanity places the same intrinsic value on phenomena that occur in an extraterrestrial civilization we encounter, and whether the extraterrestrials do the same for us. Universalism in this context can be contrasted with *civilizationism*, in which civilizations place more intrinsic value on what happens to themselves than on what happens to other civilizations. Define *pure civilizationism* as the view that what happens to other civilizations holds zero intrinsic value. This definition allows the more general term civilizationism to refer to views in which what happens to other civilizations holds less intrinsic value but not necessarily zero intrinsic value. There thus exists a continuous scale from pure civilizationism to pure universalism. However, for ease of exposition, in this paper the terms universalism and pure universalism will be used interchangeably.

Available evidence suggests that human and extraterrestrial civilizations could be universalist with each other, but would not necessarily be so. The historic record of encounters among human civilizations – often mined for insights on extraterrestrial encounters [21, 22] – shows a wide range of ethics from civilizationism to universalism. For example, the English Pilgrims and Wampanoag Native Americans coexisted in New England with mutual cooperation and assistance [23], suggesting at least some degree of universalism among the two civilizations. Many other cases of human encounters are marked by a civilizationist fight-to-win mentality. Thus if the historic record is to be any guide, universalism in extraterrestrial encounter may be possible but hardly inevitable.

The existing psychological literature has only begun to explore the extent to which humans may be universalist with respect to extraterrestrial civilizations [24]; initial results are inconclusive. There has been some discussion of this issue in the ethics literature. For example, Lupisella [25] argues for respecting and protecting life on other planets. Singer [26] expresses a similar view. Meanwhile, some have speculated that intelligent civilizations evolve universalist tendencies [3, 27]. Thus, if it is possible for

[1] This thought experiment is very similar to the "original position" thought experiment developed by Rawls [14]. Rawls uses the original position thought experiment to assess how idealized members of society might form societal rules instead of to assess what ethical framework might be selected by idealized ethical agents.

humanity to encounter an extraterrestrial civilization, or for two extraterrestrial civilizations to encounter each other, it appears possible for at least one of these civilizations to be universalist with respect to each other. I now explore some of the possible outcomes of such an encounter, given that one or both civilizations are universalist.

3. Universalism In Extraterrestrial Encounter

If two civilizations (human or otherwise) encounter each other, how would the outcome be affected if one or both civilizations are universalist? Several factors are important here: (1) whether the civilizations place intrinsic value on the same set of phenomena; (2) whether the civilizations are equally efficient at producing the intrinsic value; (3) whether the civilizations are both universalist; (4) whether one civilization is capable of destroying the other (i.e. there is a *stronger* and a *weaker* civilization; note that by *destruction* I mean to include both annihilation and enslavement); and (5) whether the two civilizations could utilize the same resources. This section discusses several scenarios that result from combinations of these factors. An overview of these scenarios is provided in Figure 1.

(A) If the two civilizations cannot utilize the same resources then universalism is a non-issue.
Otherwise:
(B) If there is a stronger civilization then that civilization decides the outcome based on which civilization it thinks produces intrinsic value more efficiently. This holds regardless of whether either civilization is universalist.
Otherwise:
(C) If the civilizations place intrinsic value on the same set of phenomena then the less efficient civilization sacrifices itself.
Otherwise:
(D) If there is agreement on which civilization is more efficient at producing the intrinsic value then the less efficient civilization sacrifices itself.
Otherwise:
(E) If each civilization thinks itself more efficient then a fight-to-win conflict ensues.
(F) If each civilization thinks the other more efficient then fight-to-lose conflict ensues.

Figure 1: Overview of the scenarios considered in this section. Letters in parentheses correspond to the paragraph labels throughout this section.

It should be noted that the ethics relevant to the present discussion are the ethics held by the civilizations during their interaction. Codignola [22] raises the possibility that, due to an encounter, one or both civilizations may change their ethical views. In the context of the present article, this means that the civilization(s) may change what they place intrinsic value on. Codignola speculates that humanity may find the extraterrestrials' ethics to be superior and may thus feel compelled to adopt the extraterrestrials' ethics. This would be

an important effect. Indeed, any shifts in ethical views due to the encounter would be important to account for. If there are any such shifts, then the ethics held during the civilizations' interaction are the ethics being referred to here.

(A) If the two civilizations did not utilize the same resources, then universalism may be a non-issue. This would be the case because neither civilization would have any reason to destroy the other. To see this, consider a scenario in which humans encounter a civilization consisting of intelligent photosynthetic organisms. Here, the two civilizations may find substantial opportunity for symbiotic relations that both civilizations would consider optimal regardless of whether either happened to be universalist. Alternatively, if no symbiosis existed, then instead the two civilizations may simply have no effect on each other's capacity to produce intrinsic value. Here, both civilizations would carry on maximizing intrinsic value however they could, as if the other civilization wasn't there.

(B) If the stronger civilization is not universalist, then it may not matter that the weaker civilization is universalist. Specifically, if the stronger civilization is non-universalist in such a way that it would desire the destruction of the weaker civilization, then the stronger civilization would simply destroy the weaker civilization, regardless of what the weaker civilization's ethics are. (It may be the case that the stronger civilization was not universalist but still did not desire destroying the weaker civilization. For example, the stronger civilization may be indifferent to the existence of the weaker civilization. Alternatively, a symbiotic relationship could develop.) Thus, a non-universalist humanity could find itself destroying a weaker universalist civilization. Alternatively, a universalist humanity could find itself being destroyed by a stronger non-universalist civilization despite humanity being universalist.

If the stronger civilization is universalist, then this civilization may decide how to handle the encounter based on which civilization more efficiently produces intrinsic value. If the two civilizations define intrinsic value differently, then the intrinsic value that decides which civilization survives would be the intrinsic value as defined by the stronger civilization. If the stronger civilization finds itself to be more efficient (for example due to superior physiology and technology) then it may destroy the weaker civilization in order to produce more intrinsic value. This would be the case even though the stronger civilization has no bias against intrinsic value produced by the weaker civilization. Likewise, if the stronger civilization finds itself to be less efficient, then it may instead opt to destroy itself. Again, humanity could either be the destroyer or the destroyed.

If there is no stronger civilization, i.e. if the two civilizations are not capable of destroying each other, then decision making may be made according to efficiency. Here it would matter whether the two civilizations had the same conception of intrinsic value.

(C) First, consider the case in which the two civilizations do have the same conception of intrinsic value. Suppose here that one civilization is more efficient than the other at producing intrinsic value. If this is the case then the less efficient civilization may agree to sacrifice itself so that the more efficient civilization could produce more intrinsic value. This scenario would occur if both civilizations use the same resources, or, to put it

more precisely, if whatever resources were used by the less efficient civilization could be more efficiently used by the other civilization. That is, the more efficient civilization may also be able to use some resources that the less efficient civilization cannot use. By definition, there are no resources which the less efficient civilization can use but the more efficient civilization cannot use. Otherwise, the less efficient civilization would be more efficient at using these resources and would remain in existence in order to convert these resources into intrinsic value. In such instances of heterogeneous capacity to convert resources into intrinsic value, the civilizations would expand, contract, and/or shift so as to maximize the production of intrinsic value.

(D) Second, consider the case in which the two civilizations do not have the same conception of intrinsic value. In this case, conflict can occur. Conflict would occur if the two civilizations disagreed over which civilization more efficiently produced intrinsic value. Conflict would not occur if they instead disagreed on what held intrinsic value but agreed on whom more efficiently produced that intrinsic value. There would be no conflict here because they would agree on which civilization was more worthy of continuing to exist. This scenario is similar to the scenario in which they agreed on intrinsic value, as discussed above. In this scenario, the less efficient civilization sacrifices itself for the sake of producing more intrinsic value.

(E) If each civilization thought itself to be more efficient, then a fight-to-win conflict could ensue. This fight would resemble a traditional fight-to-win conflict in which the civilizations are pure (or close to pure) civilizationists. However, the pure civilizationist fight-to-win phenomenon is not identical to the universalist fight-to-win phenomenon. This is because there exists a scenario in which a pure civilizationist fight-to-win conflict would occur but a universalist fight-to-win conflict would not occur. This scenario is the scenario in which the would-be winning civilization would suffer so much damage that it would be left less efficient after the conflict than the other civilization would be without a conflict. In this scenario, the would-be winning universalist civilization might sacrifice itself, in order for there to be more total intrinsic value. A pure civilizationist civilization would make no such sacrifice, because it would place no intrinsic value on anything that happened to the other civilization.

(F) If each civilization thought the other to be more efficient, then a fight-to-lose conflict would ensue. In this remarkable situation, the two civilizations might race to be the first to commit suicide. The civilizations would not want to be the second to commit suicide, because then no civilization would remain to produce any intrinsic value, however measured. In other words, even if a civilization preferred that the other civilization exist instead of itself existing, the civilization would still prefer itself existing over no civilizations existing. (If the civilization preferred no civilizations existing over itself existing, then it would have committed suicide already, before any encounter with other civilizations. There actually are some human philosophers that call for the end of human civilization on these grounds. One such philosopher is Benatar [28]; see [29] for further discussion.) Thus, if one civilization successfully committed suicide while the other was still intact, then the other could exploit the resources that otherwise would have been

consumed by the suicide civilization. This would result in more intrinsic value as defined by the suicide civilization.

4. Implications for Human Civilization Strategy

If humans never encounter an extraterrestrial civilization, then the discussion here is of no strategic significance. In other words, humanity's universalism towards other civilizations and other civilizations' universalism towards humanity only affect human strategy if humanity may encounter one of these civilizations. There are other, non-strategic implications to the possibility of universalism; some of these implications are discussed below.

If humans do encounter an extraterrestrial civilization, then the outcome of the encounter will depend not only on the civilizations' relative strength, i.e. their relative capacity to destroy the other civilization, but also on their respective ethical frameworks. Specifically, if one or both civilizations are universalist with respect to the other civilization, then the stronger civilization might not be the survivor. Indeed, it is possible that one or even both of the civilizations would attempt sacrificing itself so that the other could produce more of whatever phenomena were considered to hold intrinsic value. Thus, if humanity is interested in surviving an encounter with an extraterrestrial civilization, it would be wise to pay attention to the extraterrestrial civilization's ethics. Likewise, if humanity is interested in facilitating the *best* possible outcome of such an encounter (i.e. the outcome with the most intrinsic value), then it would be wise for humanity to also reflect on its own ethics, because its strategy will depend on how it defines the good. Such reflection includes reflecting on the question of whether humanity even defines the good in terms of an intrinsic value to be maximized instead of, for example, supporting a deontological or virtue ethics framework.

The exact strategic implications of universalism in encounter scenarios depend on the particular circumstances of the encounter. Analysis such as that presented above is an appropriate starting point.

The importance of ethics in extraterrestrial encounter implies the importance of detecting and diagnosing the ethics of extraterrestrials. Here detection refers to the process of identifying information about the extraterrestrials' ethics. Diagnosis refers to the process of interpreting this information to learn what the extraterrestrials' ethics actually are. Such detection and diagnosis are part of the broader technical and intellectual challenge of identifying and interpreting whatever evidence of extraterrestrial civilization we may find. Of all the aspects of extraterrestrial civilization we can detect and diagnose, its ethics may be particularly important. This importance is suggested by the analysis presented throughout this article. Further research is necessary to establish protocols for the detection and diagnosis of extraterrestrial ethics.

One important consideration in the detection and diagnosis of extraterrestrial ethics is that any information sent to us from the extraterrestrials may not be honest. Extraterrestrials

could send us an 'ethics Trojan horse', i.e. a message designed to trick humanity into misinterpreting the extraterrestrials' true ethics. The extraterrestrials' motivation for such trickery would be to influence humanity's course of action to the extraterrestrials' advantage, based on however they defined "advantage". Given this possibility, humanity should be prepared to be very careful in handling any messages about extraterrestrials' ethics in order to avoid any harmful misinterpretation.

The 'ethics Trojan horse' game can of course be played both ways. Just as extraterrestrials can try to trick humanity, humanity can try to trick extraterrestrials. Humanity could gain strategic advantage from such a tactic. Indeed, humanity already has extensive experience with this sort of tactic, given our long and ongoing experience with military deception [30, 31]. Thus it is recommended that studies of ethics in extraterrestrial encounter include consideration of military and other deception so that humanity can most successfully handle the possibility of an 'ethics Trojan horse'. This recommendation holds no matter what it is that humanity considers to be 'successful', i.e. no matter what our ethics are.

5. Implications for Contemporary Ethics

The civilizational encounters discussed in this paper raise some important ethical issues. These issues are important even if no extraterrestrial encounter ever occurs, because how we assess these issues says something about the nature of our character as moral beings. Furthermore, these issues raise profound questions about how we should value and treat other species here on Earth.

One ethical issue posed by the possibility of extraterrestrial encounter is the justification of ethical anthropocentrism. By *ethical anthropocentrism*, I mean the view that humans are more morally important due to some inherent human trait not found or not found as strongly in other species.[2] For example, humans may be considered more cognitively, spiritually or intellectually advanced, and thus meriting of special consideration. This justification of ethical anthropocentrism is challenged by the possibility of extraterrestrial encounter because the extraterrestrials may turn out to be substantially more advanced than humans in any of these regards. If we believe we have encountered a more advanced civilization, then would we continue advocating anthropocentrism, or would we instead consider members of this other civilization to be more morally important? (It is the belief in the civilization being more advanced that is important here, not whether the civilization actually is more advanced, because it is the belief that drives the thoughts and the actions of the believers.) How we answer this question reveals whether we truly believe in the ethics we state or if instead we are simply using them as an excuse to prioritize ourselves. If we would not prioritize members of other, more advanced

[2] This usage of the term *anthropocentrism* is distinct from a similar usage of the term in the ethics literature which appears in discussions of whether human valuations are inevitably anthropocentric as long as it is humans that are conducting the evaluation [32]. There is of course some truth to this, but humans can nonetheless change their valuations and behaviors in response to the view that nonhuman phenomena hold intrinsic value.

civilizations, then on what grounds other than simple selfishness can we justify anthropocentrism in our encounters with other species on Earth?

The issue of ethical anthropocentrism is particularly vivid in our food choice decisions. Humans readily eat members of other Earth species, an act that may be justified by ethical anthropocentrism. Would we thus condone being eaten by extraterrestrials, if the extraterrestrials proved to be more advanced? If we would not condone this, then on what grounds could we justify eating other Earth species? Of course, our survival depends on eating other species (that is, in the absence of synthetic alternatives, which may soon include synthetic meats [33]), but human survival does not depend on eating, for example, sentient species, which are sometimes considered more morally significant. Likewise, the extraterrestrials might dine on humans even if they have culinary options they consider to be less morally significant.

Closely related to food choice decisions is our evaluation of the utility monster scenario. Whereas our food choice decisions concern what we should eat, the utility monster scenario concerns what we should feed ourselves to. The scenario comes from the famous ethical thought experiment posed by philosopher Robert Nozick [34]. Nozick's utility monster is a hypothetical creature that produces more intrinsic value by eating humans than humans can produce on their own. In Nozick's original version, intrinsic value is defined as utility, although the scenario generalizes to other forms of intrinsic value. In this scenario, universalist humans would be morally obligated to feed themselves to the utility monster. Nozick perceived this obligation to be absurd and took it as an argument against universalist utilitarianism. However, if we are not to feed ourselves to the utility monster, then it must be the case either that other species should not feed themselves to us or that humanity is morally more important than all other species. If an extraterrestrial species could be more advanced than us in all of the regards we consider to be morally significant, then it becomes difficult to justify other species feeding themselves to us but us not feeding ourselves to other species.

A separate counter to the utility monster objection to utilitarianism is that the utility monster is a hypothetical thought experiment with no real-world significance. However, in the event of an extraterrestrial encounter, humanity may face a real utility monster. Given that an extraterrestrial encounter is, as far as anyone currently knows, a possibility (however remote), the utility monster should not be considered strictly a hypothetical thought experiment. Thus, the possibility of extraterrestrial encounter defeats this objection. To be sure, it is possible that there could be an extraterrestrial encounter in which humans would not be edible to the extraterrestrials (or vice versa) [35]. In this case, humanity's positions on eating or being eaten by members of other civilizations would not be of strategic significance but may still be of ethical significance. However, the opposite is also possible, i.e. that humans may be edible to the extraterrestrials. The existence of this possibility is sufficient to defeat the "thought experiment" objection to the utility monster issue.

6. Summary, Conclusion, and Future Work

If humans encounter an extraterrestrial civilization, or if two extraterrestrial civilizations encounter each other, then the outcome of the encounter will depend not only on the civilizations' relative strength to destroy each other, but also on their respective ethical frameworks. If one or both civilizations are universalist with respect to the other civilization, then the stronger civilization might not be the survivor. Indeed, it is possible that one or even both of the civilizations would attempt sacrificing itself so that the other could produce more of whatever phenomena were considered to hold intrinsic value. Thus, if humanity is interested in surviving an encounter with an extraterrestrial civilization, it would be wise to pay attention to the extraterrestrial civilization's ethics.

Meanwhile, the possibility of civilizational encounter raises several profound ethical issues. These issues primarily concern whether anthropocentricism can be justified on ethical grounds given the possibility of there existing more advanced extraterrestrials. It should be noted that for those frameworks (such as anthropocentric utilitarianism) that face these issues, the issues are, for contemporary humans, relatively minor and unimportant details. That is, there are much bigger issues at stake, ones where the recommendations are far less controversial. Chief among these is the issue of ensuring long-term survival by reducing the risk of global catastrophes such as nuclear warfare, pandemic outbreaks, environmental destruction, and large asteroid impact [36]. Avoiding such catastrophe enables humanity to produce much more intrinsic value, both here on Earth and, in particular, beyond Earth [37]. (For more on opportunities for expansion into space, see [38].) Of course, one means of helping ensure long-term survival is to colonize space, which makes humanity resistant to catastrophes that only involve planetary destruction [39]. But colonizing space also increases humanity's chance of an extraterrestrial encounter, in which the issues in this paper become important again. So, while consideration of universalism with respect to civilization encounter may not be humanity's most pressing need, it is a need nonetheless.

One important topic not considered in the present paper is when there may exist a diversity of ethical views within a civilization. Indeed, it is the case that humanity features such a diversity of views, ranging in, among other factors, the extent to which humans are universalist. Likewise, extraterrestrial civilizations may also have such diversity of attitudes in their populations. In this case, the outcome of an encounter may also depend on the roles of specific civilization members in the encounter. Analysis of such situations might consider these roles, building on analyses of human response to extraterrestrial contact [40, 41].

The present analysis is also limited by the range of universalist ethical frameworks considered. These frameworks are consequentialist in that they argue that what we should do is a function of the intrinsic value of the consequences of our actions. Other types of frameworks may behave different. Some important non-consequentialist types of frameworks are deontological ethics frameworks which claim that there are duties to perform certain acts regardless of the consequences, and virtue ethics frameworks which emphasize what we should be instead of what we should do. Within consequentialism,

an important type of framework not considered here is social choice frameworks (such as democracy), which recommend doing some function of what society in aggregate wants. Social choice frameworks are of interest in the extraterrestrial encounter context because they raise the question of how extraterrestrials are represented and counted in a social choice scheme.

Another important topic not considered here is the extent to which ethics may be affected by whether the civilizations are biological or computational. Ćirković [42] suggests that humanity might become more universalist if it evolves into a post-biological state. On the other hand, Yudkowsky [43] cautions that an artificial intelligence not pre-programmed to be 'Friendly' may be highly destructive. It is at least plausible that a biological civilization may develop empathy towards a civilization it newly encounters whereas a computational civilization might not. Given the possibility that humans may encounter a computational civilization or themselves become a computational civilization [44], this possibility may be worth exploring further.

Acknowledgments

Jacob Haqq-Misra, Luke Haqq, and two anonymous reviewers provided helpful comments on earlier drafts of this paper.

References

[1] M.A.G. Michaud, Contact with Alien Civilizations: Our Hopes and Fears About Encountering Extraterrestrials, Copernicus Books, New York, 2007.
[2] A.C. Clarke, Voices from the Sky, Harper & Row, New York, 1965.
[3] C. Sagan, W.I. Newman, The solipsist approach to extraterrestrial intelligence, Quarterly Journal of the Royal Astronomical Society 24 (1983) 113-121.
[4] T. Ferris, The Mind's Sky: Human Intelligence in a Cosmic Context. Bantam Books, New York, 1993.
[5] J. Diamond, To whom it may concern, New York Times Magazine 5 December (1999) 68-71.
[6] J.L. Heilhron, J. Conway, D.K. Cullers, S. Dick, B. Finey, K.S. Guthke, K. Keniston, SETI and history, in: J. Billingham, R. Heyns, D. Milne, S. Doyle, M. Klein, J. Heilbron, M. Ashkenazi, M. Michaud, J. Lutz, S. Shostak (Eds.), Social Implications of the Detection of an Extraterrestrial Civilization: A Report of the Workshops on the Cultural Aspects of SETI. SETI Press, Mountain View, CA, 1994, pp. 33-60.
[7] S.H. Schwartz, K. Boehnke, Evaluating the structure of human values with confirmatory factor analysis, Journal of Research in Personality 38 (2004) 230-255.
[8] T. Rønnow-Rasmussen, M.J. Zimmerman (Eds.), Recent Work on Intrinsic Value, Springer, Dordrecht, 2005.
[9] B. Bradley, Extrinsic value, Philosophical Studies 91 (1998) 109-126.

[10] D.M. Smith, How far should we care? On the spatial scope of beneficience, Progress in Human Geography 22 (1998) 15-38.
[11] M. Slote, Beyond Optimizing: A Study in Rational Choice, Harvard University Press, Cambridge, MA, 1989.
[12] P. Singer, Peter, Famine, affluence, and morality, Philosophy and Public Affairs 1 (1972) 229-243.
[13] J.C. Harsanyi, Cardinal utility in welfare economics and in the theory of risk-taking, Journal of Political Economy 61 (1953) 434-435.
[14] J. Rawls, A Theory of Justice, Belknap Press, Cambridge, MA, 1971.
[15] J. Bentham, An Introduction to the Principles of Morals and Legislation, 1789.
[16] P. Singer, Animal Liberation: A New Ethics for Our Treatment of Animals, Random House, New York, 1975.
[17] Y.K. Ng, Towards welfare biology: Evolutionary economics of animal consciousness and suffering. Biology and Philosophy 10 (1995) 255-285.
[18] A. Næss, The shallow and the deep, long-range ecology movement. Inquiry 16 (1973) 95-100.
[19] L. Vilkka, The Intrinsic Value of Nature, Rodopi, Amsterdam, 1997.
[20] T. Dietz, A. Fitzgerald, R. Shwom, Environmental values, Annual Review of Environment and Resources 30 (2005) 335-372.
[21] A.W. Crosby, Micro-organisms and extraterrestrial travel, in: L. Codignola, K.-U. Schrogl (Eds.), Humans in Outer Space: Interdisciplinary Odysseys. SpringerWienNewYork, Vienna, 2009, pp. 6-13.
[22] L. Codignola, Future encounters: Learning from the past?, in: L. Codignola, K.-U. Schrogl (Eds.), Humans in Outer Space: Interdisciplinary Odysseys. SpringerWienNewYork, Vienna, 2009, pp. 14-21.
[23] J. Hakim, A History of US, Book 2: Making Thirteen Colonies: 1600-1740, Oxford University Press, New York, 2002.
[24] D.A. Vakoch, Y.S. Lee, Reactions to receipt of a message from extraterrestrial intelligence: a cross-cultural empirical study, Acta Astronautica 46 (2000) 737-744.
[25] M. Lupisella, The rights of Martians, Space Policy 13 (1997) 89-94.
[26] P. Singer, The great ape debate, Project Syndicate, May, 2006. at: <http://www.project-syndicate.org/commentary/singer11> [accessed January 2009].
[27] J.F. Galloway, An international relations perspective on the consequences of SETI, Space Policy 12 (1996) 135-137.
[28] D. Benatar, Better Never to Have Been: The Harm of Coming Into Existence, Clarendon Press, Oxford, 2006.
[29] S.D. Baum, Better to exist: A reply to Benatar, Journal of Medical Ethics 34 (2008) 875-876.
[30] Sun Tzu, The Art of War, Dover Publications, Mineola, NY, 2002.
[31] Military Deception, United States Joint Chiefs of Staff Joint Publication 3-13.4, 13 July 2006. at <http://www.dtic.mil/doctrine/jel/new_pubs/jp3_13_4.pdf> [accessed May 2009].
[32] T. Hayward, Anthropocentrism: A misunderstood problem, Environmental Values 6 (1997) 49-63.
[33] P.D. Edelman, D.C. McFarland, V.A. Mironov, J.G. Matheny, In vitro-cultured meat production, Tissue Engineering 11 (2005) 659-662.

[34] R. Nozick, Anarchy, State, and Utopia, Basic Books, New York, 1974.
[35] C.S. Cockell, M. Lee, Interstellar predation, Journal of the British Interplanetary Society 55 (2002) 8-20.
[36] N. Bostrom, M. Ćirković (Eds.), Global Catastrophic Risks, Oxford University Press, Oxford, 2008.
[37] N. Bostrom, Astronomical waste: The opportunity cost of delayed technological development, Utilitas 15 (2003) 308-314.
[38] C.M. Hempsell, A history of space and limits to growth, Journal of the British Interplanetary Society 51 (1998) 323-336.
[39] A. Sandberg, J.G. Matheny, M.M. Ćirković, How can we reduce the risk of human extinction?, Bulletin of the Atomic Scientists, September (2008).
[40] P. Schenkel, Legal frameworks for two contact scenarios, Journal of the British Interplanetary Society 50 (1997), 258-262.
[41] M.A.G. Michaud, Broadening and simplifying the first SETI protocol, Journal of the British Interplanetary Society 58 (2005), 40-42.
[42] M. Ćirković, Observation selection effects and global catastrophic risks, in: N. Bostrom, M. Ćirković (Eds.), Global Catastrophic Risks, Oxford University Press, Oxford, 2008, pp.120-145.
[43] E. Yudkowsky, Artificial intelligence as a positive and negative factor in global risk, in: N. Bostrom, M. Ćirković (Eds.), Global Catastrophic Risks, Oxford University Press, Oxford, 2008, pp. 308-345.
[44] W.I. McLaughlin, Pathways of evolution for man and machine, Journal of the British Interplanetary Society 36 (1983), 215-222.