[Essay]

Catastrophe Ethics
Towards a Trans/Posthuman Philosophy of Life?

Kyle Takaki[*]

Abstract

Catastrophes shift the conditions of behavior and existence, potentially displacing whole systems of human culture, technology, possibility, and so forth. Can there be an adaptively robust ethics under these conditions? A speculative framework is proposed to guide a "monastic" transhumanist community, drawing ideas from Michael Polanyi and C.S. Peirce.

1. A Brief Historical Preface

Did money enable the rise of what Karl Jaspers calls the *Axial Age*, from which major philosophical schools of thought emerged? Rebecca Goldstein (2014) considers this possibility offered by David Graeber in his book *Debt*—what he terms the *military-coinage-slavery complex*—as well as the scenario put forth by evolutionary psychologists Pascal Boyer and Nicolas Baumard, where a surplus of available energy reserves (enough food, shelter, etc.) permeating the Axial Age created space for philosophical investigation to grow. Whether one is the cause of the other or mutually influencing is a difficult matter to discern, but as Goldstein observes, we can still glean key philosophical points of interest.

Foremost for Goldstein is the Greek preoccupation with excellence—what she calls the *Ethos of the Extraordinary*—and its implications for quintessential qualities that have come to characterize Western philosophy. Another intriguing facet of the philosophical leisure space opened up via the military-coinage-slavery complex and/or surplus of energy reserves is Goldstein's claim that the fragility of life is rendered in clearly non-transcendent terms (Goldstein 2014, 133-35). This is one way to read Plato's Homeric tradition, and the Greek founding of Western philosophy more generally. However, there is another reading of the spiritual and religious character of Plato and the Homeric Ethos of the

---

[*] Independent Scholar. Email: ktakaki[a]hawaii.edu. Thanks to Professor Morioka for his patience and generosity.

Extraordinary that Simone Weil directs us to in her incisive insights on the "intimations" of Christianity in ancient Greece. Underlying Greek fragility is a deeper recognition of cosmic "might" (Weil 1957, Chapter 4) that all eventually confront, whose depths of suffering pave the way for "transcendent" insights. For Weil, the mystic dimensions of Greek (and Christian) thought are not only present in Plato (compare Kennedy 2014), but are at the heart of Greek notions like fragility and measure. It might also be said more broadly that mystic dimensions run through the Axial Age, embodying core wisdom across differing cultures.

But whether we opt for Goldstein's mainstream reading of Plato, or Weil's iconoclastic rendering of the Greek tradition, is something we need not adjudicate here. The point of bringing out these differing readings at the founding of philosophy writ globally—*contra* Heidegger and Derrida, who assert that philosophy properly speaking is only attributable to the Western tradition starting with Plato[1]—is to disclose the close relation that the rise of spiritual-religious-philosophical schools of thought have with the availability of energy reserves/military-coinage-slavery complex. In brief, reflection on excellence, suffering, the nature of the cosmos, etc. are parasitic upon having enough space-for-leisure, where one is not engulfed by the brutalities of mere survivalism. This raises the question: What happens to philosophical projects when these background enabling conditions deteriorate?[2]

What I am calling *catastrophe ethics* takes as its starting point inquiry in the face of significant curtailing of space-for-leisure. I proceed from the assumption that standard ethical categories or approaches (deontological, utilitarian, etc.) cannot here be suitably deployed, or may even be entirely irrelevant. This differs from critiques seeking alternatives to apparently dated frameworks, as with the claim that Rawlsian justice is no longer appropriate in the face of liberalism's global crisis.[3] These projects adopt a more parochial time frame in seeking a response to current problems. Catastrophe ethics takes a longer and wider view of future generational challenges (but short of species-level extinction events) in attempting to articulate a framework that is fundamentally rooted in living—short of mere survivalism—with enough, though scarce, space-for-leisure.[4]

---

[1] See https://aeon.co/essays/why-the-western-philosophical-canon-is-xenophobic-and-racist
[2] Compare https://www.truthdig.com/articles/noam-chomsky-america-has-built-a-global-dystopia/ and also Berman 2000, 2006, and especially 2012.
[3] See https://bostonreview.net/philosophy-religion/katrina-forrester-future-political-philosophy and https://newrepublic.com/article/155294/john-rawls-missed-create-just-society
[4] Compared to this catastrophe viewpoint, what is called *collapsology* affords more leisure space and

2. What Are Catastrophes?

   In terms of riskiness, catastrophes fall short of existential risks, which deal with extinction-level threats. Mathematically speaking, catastrophes mark points of emergence where some dramatic shift occurs (a simple example would be adding grains of sand to a pile, where past a certain point a landslide occurs—but understanding how and why this happens isn't merely a quantitative matter, from the standpoint of continuous aggregation, but rather concerns nonlinear bifurcations that issue in abrupt change; compare Renfrew and Cooke 1979, section VI). What both senses share is some rupture (relative to a timescale and context in which change is qualified) that issues in new behavior for a novel landscape.[5] Catastrophes shift the parameters of the previous state of affairs in a way that systemically alters the subsequent space of behavior.[6] As this applies to our current global crises, climate change being the most critical, there probably won't be one catastrophe, but a perfect storm of contributing catastrophes (and/or

less systemic chaos; see https://www.nybooks.com/daily/2020/01/21/collapsologie-constructing-an-idea-of-how-things-fall-apart/ and Servigne and Stevens (2020).

[5] Catastrophes are situated along the continuum between extinction-level events and localized collapses. Joseph Tainter writes, regarding collapse:

> 'Collapse' is a broad term that can cover many kinds of processes … Collapse, as viewed in the present work, is a *political* process … A *society has collapsed when it displays a rapid, significant loss of an established level of sociopolitical complexity*. The term 'established level' is important. To qualify as an instance of collapse a society must have been at, or developing toward, a level of complexity for more than one or two generations … Collapse is manifest in such things as:
>> a lower degree of stratification and social differentiation;
>> less economic and occupational specialization, of individuals, groups and territories;
>> less centralized control … ;
>> less behavioral control and regimentation;
>> less investment in … those elements that define the concept of 'civilization' … ;
>> less flow of information between individuals, between political and economic groups … ;
>> less sharing, trading, and redistribution of resources;
>> less overall coordination and organization of individuals and groups;
>> a smaller territory integrated within a single political unit.
> Not all collapsing societies, to be sure, will be equally characterized by each item on this list, and the list is by no means complete … This list, however, provides a fairly concise description of what happened in most of the better known cases of collapse. (Tainter 1988, 4)

Note that catastrophes differ from Tainter's narrow use of the term (Tainter 1988, 52), which concerns mainly one-off disasters, since here they possess a greater level of multivariate and systemic threat, damage, and/or decay.

[6] For a historical examination of localized collapse, see Diamond (2011); for a geological view of catastrophes and extinction-level/existential events, see Kolbert (2014); for a post-catastrophic/existential imagining of the world without us, see Weisman (2007).

collapses) leading overall to a new constrained space for what leisure is available. Plausible candidates might be the rise of AI, the collapse of liberal democracies, the emergence of climate-change related pandemics, various conflicts over resources that scarcity will induce (disputes over water being perhaps most prominent), and so on.[7] These ingredients would simmer in the pot of climate change, enacting major catastrophe—a new space of existence similar to a (hypothetical) state of nature, but actually inaugurating an unprecedented geological epoch.[8]

Insofar as there are rights in this regime, they would tend to be tribal, or exclusive. Insofar as there are advanced technologies, they would tend to be hoarded, or developed by the most elite—who in turn would be consumed with the urgent pursuit of transformation (if there ever were a transhumanist ivory tower, a "monastic"[9] crucible for shifting the very nature of humanity, this would be it). And insofar as there is space for religious and spiritual pursuit, it would occur not amidst surplus akin to the Axial Age, but environmental knackeredness centering on scarcity of a fundamentally different sort than at any time in human history. Catastrophe here would radically shift the conditions of behavior and existence, displacing whole systems of human culture, technology, possibility, and so forth.[10]

Both deep pessimism and optimism find expression in the midst of chaos. What can we glean philosophically under such a regime, which shifts the (Axial) boundary conditions that enabled philosophy in the first place? Is catastrophe philosophy even a conception of possibility? For living is now grounded in survivability, where leisure-space for philosophical reflection is shackled to this expediency. And while not a state of nature, these conditions induce similar behavior, though tempered by long-wrought historical "ratchets" (technological,

---

[7] For a data-driven view short of collapse but projecting, to 2052, tapered coping amidst increasing global problems, see Randers (2012; although see pages 76 and 306). Taking a longer view of things, an imaging of collapse-and-catastrophe enmired in increasing "cascading violence" is offered by Wallace-Wells (2019).

[8] For overviews of past extinctions and dire projections for the future as related to climate change, see Ward (2007) and Brannen (2018). In terms of Deep Time, compounded catastrophes that issue in mass extinctions (existential-level events) mark geological transitions—shifts which, past certain tipping points, create network collapses that kill nearly all life on the planet. The sort of catastrophe imagined here would be short of a planetary extinction-level event, but major enough to mark the end of the Anthropocene. See also https://www.nature.com/articles/d41586-019-03595-0

[9] Compare Berman 2000.

[10] Such a shift would induce a wholly new Being-*and*-ground-of-Being, effacing once affordable distinctions between ontic and ontological.

economic, sociological, etc.) that guide survivability. In other words, catastrophe is somewhere between a raw state of nature (which has a surplus of energy reserves, but where mere survivability dominates) and a state of scarcity (which, tempered by historical ratchets, affords a bit of leisure-space).

Can there be any ethics here, given ubiquitous free-rider predicaments inclining towards annihilation? The only reason for ethics is to prevent this slide, with some streamlined sense of community salvaged. This isn't a scaled-down version of standard ethical views, but rather an eclectic mix of ethical "kludges" that contribute to basic survivability. I speculate the biggest difference lies in the cultivation of a "hive mind," that current surveillance capitalism grows and guides (Zuboff 2019), requiring a different kind of patchwork ethics, largely crowdsourced, but grounded in pragmatic considerations.

If these sketches appear plausible, they largely pertain to groups at the "ground level" of catastrophic living. The question "What is a catastrophe?" then amounts to "A catastrophe for whom?" since the other type of group for which a related yet markedly different catastrophe ethics holds are the transhumanists seeking to move to posthumanism, perhaps especially in the face of scarcity funding their extraordinary and highly risky transition.[11] Typical transhuman ethical codifications—most especially the Transhumanist Declaration[12]—only apply to this most monastic community, segregated from the rest of the world. To see how this declaration fares in the face of a catastrophic landscape, let's address each of its claims.

3. Transhumanist Declaration and Catastrophe

The first claim is that "[h]umanity stands to be profoundly affected by science and technology in the future. We envision the possibility of broadening human potential by overcoming aging, cognitive shortcomings, involuntary suffering, and our confinement to planet Earth."[13] There is a lurking inequality at the core of these ambitions exposed through catastrophe—for resources simply wouldn't be available for all except those monastic few. The second claim fares no better: "We believe that humanity's potential is still mostly unrealized. There are possible

---

[11] Groups at the "ground level" need to be resilient to persist as a collective, which their brand of catastrophe ethics reflects; by contrast, a transhumanist catastrophe ethics needs more—it has to embody antifragile qualities.

[12] See https://humanityplus.org/philosophy/transhumanist-declaration/

[13] All quotes are from the above link to the Transhumanist Declaration.

scenarios that lead to wonderful and exceedingly worthwhile enhanced human conditions." But these benefits would only accrue to a circumscribed transhumanist community.

What holds true for the third claim is that with scarcity, change may not indicate progress: "We recognize that humanity faces serious risks, especially from the misuse of new technologies. There are possible realistic scenarios that lead to the loss of most, or even all, of what we hold valuable … Although all progress is change, not all change is progress." The question then becomes: Under drastic conditions, would what we hold valuable be jettisoned in view of "realistic" (and thus justifiable) "misuse"? Or justifiable "change"?

The fourth proposes that "[r]esearch effort needs to be invested into understanding these prospects. We need to carefully deliberate how best to reduce risks and expedite beneficial applications. We also need forums where people can constructively discuss what should be done, and a social order where responsible decisions can be implemented." This presupposes capability and time for careful deliberation, constructive discussion, and responsible decision-making—in other words, leisure-space. When curtailed, it isn't clear to what extent they can robustly occur.

Fifthly, "[r]eduction of existential risks, and development of means for the preservation of life and health, the alleviation of grave suffering, and the improvement of human foresight and wisdom should be pursued as urgent priorities, and heavily funded." A horrific wisdom might be memorialized here, given the profound sacrifices required to avoid existential collapse; in any event, the concrete ways in which this would be carried out probably beggar our contemporary, and comparatively pacific, imaginations.

Sixthly, "[p]olicy making ought to be guided by responsible and inclusive moral vision … respecting autonomy and individual rights, and showing solidarity with and concern for the interests and dignity of all people around the globe. We must also consider our moral responsibilities towards generations that will exist in the future." These hopes can only be projected to a posthuman future, not a catastrophe-saturated world; enlightenment values like "autonomy" and "rights" whither in poisoned, catastrophic soil. And "moral responsibilities" become survivabilities in the guise of a re-engineered humanity. The same amendments apply to the seventh claim: "We advocate the well-being of all sentience, including humans, non-human animals, and any future artificial intellects, modified life forms, or other intelligences to which technological and

scientific advance may give rise."

Eighthly, "[w]e favour allowing individuals wide personal choice over how they enable their lives. This includes use of … [many] possible human modification and enhancement technologies." This would be the provenance and privilege of transhumans on the way to posthumanism; select variants of catastrophe ethics would likewise coevolve in the transition to posthumanism.

What all of this suggests is that just as the Axial Age arose with the tradeoff of the availability of energy reserves/military-coinage-slavery complex, so a similar transitional age emerging out of catastrophe will come at the price of passing through a very severe bottleneck. Any ethics picked up along the way will be, of necessity, an adaptive patchwork of philosophical-political-religious-spiritual (and so forth) kludges that enable pragmatic survivability for the two types of communities (the ordinary hive-minds and the transhuman monastics).

I suspect this narrowest of bottlenecks implies that any catastrophe ethics which can persevere and potentially prosper in an unknown "human" future will eventually need to be tied to a transhumanist-posthumanist program. As alluded to above, since Enlightenment values won't suffice for catastrophic regimes, a differing value system needs to be developed that can navigate the razor's edge between, on the one hand, mere survivalism and avoiding the seductions of defecting to a state of nature, and, on the other hand, a rejiggered transhumanist declaration making maximal use of scarce leisure-space. In the remainder of the paper I offer a speculative framework, drawing from Michael Polanyi and C.S. Peirce, which hopefully can ride the fine line between realpolitik and post-Enlightenment techno-values—values that fund the rise of transhumanism amidst entrance into a post-Anthropocene epoch.

4. Polanyi's Society of Explorers

Michael Polanyi's Society of Explorers (SoE) forms an ideal community embodying the best practices of inquiry, broadly conceived and properly accredited (he lays out a narrower conception for science termed the *Republic of Science*). Briefly, its lowest levels concern material practices upholding the society's endeavors—e.g., economic practices that are mutually enforced, adjusted, and adjudicated through laws, officers of the court, etc. Intermediate levels focus on institutions of civic responsibility enculturing and articulating loyalty to the SoE's values and practices. And the highest levels regard the

pluralistic ideals at the core of the SoE: pursuing wisdom, truth, excellence, and so forth. These are not static levels but rather express a dynamic, heterarchically-conceived hierarchy of mutually reticulated, spontaneous orders[14] of emergence (see Cordner 2019 and Takaki 2011, 2013).

Polanyi describes four *coefficients of society* in *Personal Knowledge*, which have the following characteristics:

> In their most general form, these four sets of institutions [i.e., the forms that these coefficients tend to take] consist in (i) institutions of *culture*, which foster shared convictions (e.g., churches, museums, universities, theatres, etc.); (ii) institutions fostering *group loyalty* (e.g., social intercourse, rituals, and common defence); (iii) an *economic system* which fosters cooperation for the purposes of achieving a joint material advantage, and (iv) *public power* to shelter and control the other institutions of society, through the use of [properly accredited] authority and coercion [which is phronetically embodied and deployed]. (Cordner 2019, 46-7)

As Colin Cordner notes, the reason why these features matter—features extracted from an exemplary community of scientific practice to a broader image of a SoE—is that they give us a sense of what communal excellence might look like:

> What this largely amounts to in a healthy free society is the ongoing cultivation of individual excellence, together with a restriction of individual selfishness—particularly moral inversion. The children of a free society are persuaded to overcome the tendency towards selfishness through their enculturation, through which they find expression for the higher passions. On a personal level this is accomplished by initiation into convictions transcending the baser appetites and drives, by the creation of an atmosphere wherein genuine fellowship and conviviality can flourish, and, when necessary, by the imposition of standards of morality through the authority or the coercion of the institutions of public power. Individual

---

[14] What makes for a spontaneous order, at least in a Republic of Science, is the "cooperation and the exercise of authority among scientists … [whose following] three practices or institutions … constitute the Republic as a *spontaneous order*: i) *self-coordination through mutual adjustment*, ii) *discipline under mutual authority*, and the iii) *principle of overlapping neighbourhoods*" (Cordner 2019, 49).

> excellence will therefore be bound and channeled by the standards and practices of the prevailing institutions of culture. (Cordner 2019, 47)

Given the mutually reticulated nature of the heterarchical-hierarchy that guides this development and health of community, what it in effect does is minimize free-rider problems while simultaneously promoting inquiry and the highest aspirations for human potentiality.

Polanyi's SoE has clear bearing on transhumanist ambitions, which I think could benefit from Polanyi's profound framework. For perhaps the most important unaddressed background problem in transhumanism is the unacknowledged presumption that power/might gains one legitimate access to the elite company of the "catastrophe ivory tower." It should not, and would actually exacerbate free-rider problems that obstruct transhuman evolution.[15] Polanyi's SoE isn't immune to freeriderism and related ilk (e.g., the alluring charisma of an Alcibiades-like wielder of power), but it does contain resources that can, when coupled with accredited transhumanist technologies, make possible maximal use of scarce leisure-space[16]—the crucible that could fuel the leap to posthumanism.

In brief, Polanyi's SoE harnesses the emergent complexity of human potentiality.[17] Its bearing on catastrophe ethics is the opportunity to use the crucible of scarcity for bootstrapping transhumanism to a posthuman space. But this can be achieving only by minimizing various sorts of freeriderisms; the transition must navigate a razor's edge indeed.[18] The next section revises the

---

[15] The problem is more general than this, if the great biologist Ernst Mayr is right in claiming that intelligence is a lethal mutation for our (or any evolved) species. It might even be said that the levels of lethality opened up by significant intelligence is the core tangle expressed by the problem of freeriderism. See also https://chomsky.info/20100930/

[16] Leisure here would be closer to Aristotelian *scholê* rather than the leisure class that Thorstein Veblen skewers.

[17] On a much smaller scale, for an approach taking complexity and uncertainty seriously in developing "post-normal science," see Benessia et al. (2016).

[18] According to Tainter, two key features of complex societies are heterogeneity—"the number of distinctive parts or components to a society, and … the ways in which a population is distributed among these parts" (Tainter 1988, 23)—and inequality—"vertical differentiation, ranking, or unequal access to material and social resources" (ibid.). A SoE promotes maximal productive heterogeneity and minimal inequality (due to the internal minimization of free-rider problems). For in contrast to Tainter's contrast between smaller social arrangements and larger complex societies, a SoE lies not just between these extremes, but also harnesses the best from both. The former is characterized by "small, internally homogeneous, minimally differentiated groups characterized by equal access to resources, shifting, ephemeral leadership, and unstable political formations"; the latter by "large, heterogeneous, internally differentiated, class structured, controlled societies in which the resources that sustain life are not equally available to all" (Tainter 1988, 38). The four coefficients of a SoE cut across both of these characterizations, thus potentially enabling it to avoid social collapse. Tainter

Transhumanist Declaration (TD) by offering amendments aligning with the spirit of a SoE, thus providing a framework for thinking about an adaptively robust catastrophe ethics.

5. Catastrophe Ethics and the TD

Recall the first claim of the TD: "We envision the possibility of broadening human potential by overcoming aging, cognitive shortcomings, involuntary suffering, and our confinement to planet Earth." Applied to a catastrophe ivory tower, these techno-possibilities should accord with Polanyi's four coefficients when considering what technologies prudently expand human potential given the scarcity of leisure-space. The second claim—"[t]here are possible scenarios that lead to wonderful and exceedingly worthwhile enhanced human conditions"— requires similar accordance.

The third and fourth claims—"[t]here are possible realistic scenarios that lead to the loss of most, or even all, of what we hold valuable … We need to carefully deliberate how best to reduce risks and expedite beneficial applications"—express the need for a framework, like the SoE, which puts us on a path to phronetic deliberation about what amounts to, in essence, the future survival, prospects, and potential flourishing of a (trans-post) human species.

The fifth and sixth claims—"[r]eduction of existential risks, and the improvement of human foresight and wisdom should be pursued as urgent priorities … Policy making ought to be guided by responsible and inclusive moral vision"—express the same phronetic exigency, which the SoE fills. For the SoE places the pursuit of wisdom writ large at its core (see below), and articulates how responsibility, training, and accredited authority enable the spontaneous orders of emergence that are critical elements of robust techno-social complexity.

The seventh and eighth claims are where things get interestingly different; to recap: "We advocate the well-being of all sentience … and any future artificial intellects … to which technological and scientific advance may give rise … We

---

characterizes collapse as "declining marginal returns on investment in complexity" (Tainter 1988, 192; see Ch.5); a SoE buffers against decline by harnessing "complexity's complexity," as it were, since it investigates (and invests in) resiliency and its enabling conditions. More specifically, Tainter's three underpinnings for his characterization of collapse—that "human societies are problem-solving organizations; [that] sociopolitical systems require energy for their maintenance; [and that] increased complexity carries with it increased costs per capita" (Tainter 1988, 194)—are streamlined in a SoE. This also applies to Cook (2000).

favour allowing … possible human modification and enhancement technologies." Most likely, the SoE will require a kind of techno-oracle to (mutually) develop its co-evolutionary trajectory, making for a "dialectical crowdsourcing" that synergizes group and individual excellence (where there is enough freedom afforded to enable robust complexity). An oracular AI[19] plausibly on the path to wisdom would do well to heed the following Socratic insight:

> [The] SoE can be thought of as the post-critical philosopher writ large (to borrow a Socratic maxim): it represents a balance in the understanding between critique, faith, and comprehension, moved by the all-pervading desire to know. Its dynamism is an image, so to speak, of the comfortable embodiment of the post-critical mind, generally accepting an inheritance of articulate culture as its happy dwelling place, while encouraging the emergence of meaningful dissent and reform in response to its principles. (Cordner 2019, 55-6)

The SoE as a set of coefficients/institutions that dynamically enable inquiry is one structural aspect of its being a community; it is also a collection of individuals (AI, transhuman, etc.) that dynamically enable-and-are-enabled-by these institutions.

In brief, a framework for catastrophe ethics, if properly enacted, is actually geared toward transfiguring catastrophe into its very opposite. It isn't a political ethics for escaping out of a state of nature, nor an ethics of mere survivalism. It is an ethics to transform the species into something both recognizable and yet not— to heed the angels of our better nature while re-engineering aspects of our being, using a pragmatic (in the best sense of the word) toolkit aimed at inquiry and tethered by the pursuit of wisdom.

It might then be asked: But what about this overall sketch makes it an ethics as such? For in drawing from political economy, technology, science, values, and so forth, what specifically makes it a branch of ethics? This good question actually reveals the entrenched nature of modern inquiries that segregate ethics from its broader roots in philosophy as a way of life, whose key import of living-via-inquiry constitutes philosophical activity as a core ethical practice. Its Greek root in *ethos* conveys the "habit," "custom," or "character" that Aristotle links to *ethike*

---

[19] See, for example, Copeland (2002), and https://www.nickbostrom.com/papers/oracle.pdf. More generally, a friendly AI on the way to superintelligence would need to co-evolve with the SoE.

(*arete*). But even with this expanded sense, ethics still doesn't appear to be able to carry the workload that the above framework requires.

The sense in which catastrophe ethics is an ethics is best interpreted via C.S. Peirce's notion of *Pure Ethics* as the category articulating the guiding principles of various modalities of inquiry. It is a category distinct from (though related to) morality-as-practiced; it forms a theoretical space from which to address these practical considerations—for it focuses on the architecture of the norms that govern practice and inquiry (at the level of what Peirce calls *thirdness*). So the sense in which, say, symbolic logic is a sub-branch of Pure Ethics regards the manner in which disciplined inference discerns valid from invalid arguments. More generally, the study of what disciplined inference amounts to, in producing (semiotically) fruitful lines of inquiry, is the field of Pure Ethics. Despite Peirce's emphasis on the theoretical nature of Pure Ethics, there is significant resonance with the Greek view that ethics and excellence of character go hand in hand— both are concerned with the nature of disciplined habits aimed at certain goods or values. The difference lies in scope; as Rachel Herdy writes, Peirce "sees an element of generality in ideals of conduct because he does not take them to be a fragment of a continuous process [as do those who restrict ethics to being a subject of study], but the continuous process itself; i.e. ideals of conduct are seen as relations of ideals of conduct."[20]

Just as the SoE must deliberate about ideals of conduct regulating the various coefficients of society, so the more difficult task of deliberating about relations among ideals remains, which must address perhaps the most difficult question: Which values and ideals are affirmed, which revised, and which jettisoned? In Peircean fashion, this process will require elements of logic (structural reasoning and justification), elements of controlled inference (ethics), and elements of educated-and-educable preference (aesthetics). A catastrophe ethics that is able to avoid succumbing to the temptations of scarcity while navigating the bottleneck to posthumanism will manifest the best of what a Peircean pragmatism—and a Polanyian SoE—envisions.

---

[20] Herdy (2014) offers the very interesting insight that the level of Thirdness is constituted by the regulative ideals of Pure Ethics, Logic, and Aesthetics: "Peirce did not equate value with esthetical goodness alone; but with a complex triadic structure that elicits not only the idea of beauty, but also the notions of right and truth. As I understand, Peirce argued for a threefold notion of the normative ideal, in which logic, ethics, and esthetics are equally identified with the category of Thirdness." In short, semiotic inquiry, which centrally characterizes Peirce's pragmatism, can be significantly identified with this threefold structure of regulative ideals that delineate the breadth and depth of inquiry's consequential dimensions. See also Feibelman (1943).

6. Dark Side of the Transition

But catastrophic scarcity will incur steep tradeoffs entangled with the cultivation of exemplary efficiencies, casting a tragic shadow on the trajectory of the SoE and its Peircean ideals, for which extenuating justifications ("the ends justifies the means," and so forth) offer little consolation. There are at least three darknesses that scarcity will induce.

The first concerns who remains in the community and who is either "expired" or banished. The severe nature of the meritocracy that catastrophe engenders presents a twofold difficulty. On the one hand, inquiry requires space for creativity, a degree of freedom, and access to resources in order to explore in fruitful fashion, which will be compressed given the scarcity of leisure-space. On the other hand, those who are able to achieve less yet within budget (a kind of localized maximin) may have a higher perceived merit. This tension threatens a race to the bottom at worst, and in its middling variations might not be enough to accord with the SoE and Peircean pragmatics. Middling efforts probably won't suffice to shift the community past the threshold to posthumanism; additionally, meritocracy here requires robust wisdom to determine what tradeoffs are worth risking and what resources to allocate. There will be rather dark compromises/sacrifices made in this regard, with concomitant rationales; the seductions of moral inversion remain throughout.[21]

A second consideration is the modification of at least three evolutionary factors: our genetic inheritance, our cultural heritability, and gene-culture co-evolutionary dynamics. For the first, perhaps an oracular AI could run simulations forecasting ways to modify—if not completely overhaul—our genetic heritage. Putting this into practice would be highly risky and potentially dark indeed. For the second, cultural heritability could likewise be simulated concerning how to tweak the coefficients of the SoE to best maximize use of scarce resources. Putting this into practice would also be highly risky with consequences possibly rivaling humanity's darkest hours. Risk only increases with the last, given the complex

---

[21] For a contemporary version of these seductions, see N. Klein (2007). Polanyi's warnings about moral inversion apply to two types of advanced social groups—collectives that curtail freedom enough to be characterized as totalitarian, and communities that are largely free (note that a SoE aims to avoid both the corruption of moral inversion and the unaccredited denial of freedom). These warnings can also be applied to competing transhuman groups, like Socratic SoEs and their darker Machiavellian brethren.

ways in which gene-culture dynamics may enable "unknown unknowns" (or true black swans), making for profound moral-political-epistemological darknesses.

The third consideration might be the most disturbing, as it suggests that the very nature of posthumanism is one where we not only transcend humanity, but in order to do so we must reject central features of humanism. What if, for example, pursuing justice is irretrievably tied to the sorts of embodied creatures we are, ensnared in fallenness, for which this ideal makes sense? But given a being/collective where perfectibility in various domains may no longer be an issue, what work would an ideal like this really do? Furthermore, what if ideals like pursuit of (pure) ethics, aesthetics, and so on are merely expressions of limitations for not just finite but profoundly constrained beings like us—limitations that no longer hold for those on the cusp of transition to posthumanity (AI, transhuman, or otherwise)? Similar considerations hold for practices embodying the ideals and aspirations of our particular style of finitude. The darkness here is metaphysical: we no longer have any idea what a truly posthuman future would look like, what posthuman beings would actually be and how they think, and so on. This might be a matrix altogether alien to us.

7. Towards a Posthuman Future

But would it be a matrix so radical as to wholly jettison its founding humanistic roots? While it may appear so to prior communities of inquirers, given the magnitude of the transformation (akin to A.C. Clarke's dictum that any sufficiently advanced technology is indistinguishable from magic), I suspect it would still be compatible with the spirit of the SoE and trajectory of Peircean pragmatism. Peirce writes that inquiry

> [m]ust not stop at our own fate, but must embrace the whole community. This community, again, must not be limited, but must extend to all races of beings with whom we can come into immediate or mediate intellectual relation. It must reach, however vaguely, beyond this geological epoch, beyond all bounds. (Peirce 1957, 67)

Peircean inquiry embraces transhuman stages of development and types of advanced civilization (expressed by the Kardashev scale, for example), and links human and posthuman development through a kind of cosmic social principle. As

James Feibelman writes, reflecting on Peirce:

> The command to love one's neighbors more than one's self, and their neighbors more than them, and so on in an ascending hierarchy of love which eventually must embrace the entire universe of being, makes love the over-all deontological requirement. Hence hatred and evil, in this scheme, must become, as they do, "mere imperfect stages of love and loveliness" (VI.287), and through the struggle against evil, which "it is man's duty to fight" (VI.479), we are enabled to increase the amount of love in the actual world. Evil is thus the adaptation of means to ends, and it is, after all, as great a thing as the law of growth which imposes fighting upon man. He will not despair to see the things for which he fights perish, since, according to the doctrine of the unlimited community, he must expect it, "accepting his little futility as his entire treasure" and understanding that "though his desperate struggles should issue in the horrors of his rout, and he should see the innocents who are dearest to his heart exposed to torments, frenzy and despair, destined to be smirched with filth, and stunted in their intelligence, still he may hope that it be best for them." Thus in Peirce's system, just as logic leads to [pure] ethics, so ethics in turn leads to the discernment of reasons and values which lie beyond human comprehension but which demand human allegiance and even sacrifice, and hence to the province of religion. (Feibelman 1943, 109)

So while posthumans may abandon sundry relics of humanism, the project still retains humanistic vestiges in its indefinite community of inquirers. For whatever form of techno-philosophy this community embodies, its deeper and wider values are still entwined with a religious community expressing a quintessential spiritual search.

Firstly, a transhuman/posthuman SoE (or its evolved descendants) qualifies as "religious" in Durkheim's well-worn sense:

> *A religion is a unified system of beliefs and practices relative to sacred things, that is to say, things set apart and forbidden—beliefs and practices which unite into one single moral community called a Church, all those who adhere to them.* The second element thus holds a place in my

definition that is no less essential than the first: In showing that the idea of religion is inseparable from the idea of a Church, it conveys the notion that religion must be an eminently collective thing. (Durkheim 1995, 44)[22]

A SoE can be viewed as a transhuman "Church"[23] whose system of beliefs and practices is bound by an amended TD, where what is held "sacred" concerns commitment to a Peircean cosmic social principle.

Secondly, this principle expresses a spiritual search, for which Peirce also provides three guiding sentiments mapping to the regulative ideals of charity, faith, and hope:

> I … put forward three sentiments, namely, interest in an indefinite community, recognition of the possibility of this interest being made supreme, and hope in the unlimited continuance of intellectual activity, as indispensable requirements of logic … It interests me to notice that these three sentiments seem to be pretty much the same as that famous trio of Charity, Faith, and Hope, which, in the estimation of St. Paul, are the finest and greatest of spiritual gifts. Neither Old nor New Testament is a textbook of the logic of science, but the latter is certainly the highest existing authority in regard to the dispositions of heart which a man ought to have. (Peirce 1957, 68-9)

To which we might add: and the highest dispositions of spiritual value for which any community can strive. For they would embody, for any sort of posthumanism, a key remnant of humanity worth affirming. While humanity may be surpassed, its long-wrought philosophies of spirit/being and their illuminative insights will conceivably not only stand the test of time, but enact indefinite communities in unlimited inquiry, aimed at cosmic love. And so, perhaps, through the darkness emerges light.

---

[22] Compare S. Klein (2014) and Wade (2009).
[23] See also Durkheim 1995, 446-47, and Peirce 1957, 192-94.

References

Benessia, Alice, Silvio Funtowicz, Mario Giampietro, Andrea Saltelli, Ângela Guimarães Pereira, Jerome R. Ravetz, Andrea Saltelli, Roger Strand, and Jeroen P. van der Sluijs. 2016. *The Rightful Place of Science: Science on the Verge*. Tempe, AZ: Consortium for Science, Policy, & Outcomes.

Berman, Morris. 2000. *The Twilight of American Culture*. New York: Norton.

———. 2006. *Dark Ages America: The Final Phase of Empire*. New York: Norton.

———. 2012. *Why America Failed: The Roots of Imperial Decline*. New Jersey: John Wiley and Sons.

Brannen, Peter. 2018. *The Ends of the World: Volcanic Apocalypses, Lethal Oceans, and Our Quest to Understand Earth's Past Mass Extinctions*. New York: Ecco.

Cook, Richard I. 2000. "How Complex Systems Fail." https://web.mit.edu/2.75/resources/random/How%20Complex%20Systems%20Fail.pdf

Copeland, B.J. 2002. "Hypercomputation." *Minds and Machines*, 12: 461-502.

Cordner, Colin. 2019. "Michael Polanyi and the Society of Explorers." *Tradition and Discovery*, 45(3): 45-60.

Diamond, Jared. 2011. *Collapse: How Societies Choose to Fail or Succeed*. New York: Penguin Books.

Durkheim, Émile. 1995. *The Elementary Forms of Religious Life*, translated by Karen E. Fields. New York: The Free Press.

Feibelman, James. 1943. "A Systematic Presentation of Peirce's Ethics." *Ethics*, 53(2): 98-109.

Goldstein, Rebecca. 2014. *Plato at the Googleplex: Why Philosophy Won't Go Away*. New York: Pantheon Books.

Herdy, Rachel. 2014. "The Origin and Growth of Peirce's Ethics: A Categorical Analysis." *European Journal of Pragmatism and American Philosophy*, 6(2).

Kennedy, J.B. 2014. *The Musical Structure of Plato's Dialogues*. Hoboken: Taylor and Francis.

Klein, Naomi. 2007. *The Shock Doctrine: The Rise of Disaster Capitalism*. New York: Metropolitan Books.

Klein, Stefan. 2014. *Survival of the Nicest: How Altruism Made Us Human and Why It Pays to Get Along*, translated by David Dollenmayer. New York: The

Experiment.

Kolbert, Elizabeth. 2014. *The Sixth Extinction: An Unnatural History*. New York: Henry Holt and Company.

Peirce, Charles S. 1957. "The Doctrine of Chances" in *Charles S. Peirce: Essays in the Philosophy of Science*, edited by Vincent Tomas. New York: The Liberal Arts Press.

Randers, Jorgen. 2012. *2052: A Global Forecast for the Next Forty Years*. Vermont: Chelsea Green Publishing Renfrew, Colin and Kenneth Cooke, eds. 1979. *Transformations: Mathematical Approaches to Cultural Change*. New York: Academic Press.

Servigne, Pablo and Raphaël Stevens. 2020. *How Everything Can Collapse: A Manual for our Times*, translated by Andrew Brown. Medford, MA: Polity Press.

Tainter, Joseph. 1988. *The Collapse of Complex Societies*. New York: Cambridge University Press.

Takaki, Kyle. 2011. "Enactive Realism." *Tradition and Discovery* 38(1): 34-59.

———. 2013. "Juarrero, Polanyi, and Complexity." *Tradition and Discovery* 40(1): 5-15.

Wade, Nicholas. 2009. *The Faith Instinct: How Religion Evolved and Why It Endures*. New York: The Penguin Press.

Wallace-Wells, David. 2019. *The Uninhabitable Earth: Life After Warming*. New York: Tim Duggan Books.

Ward, Peter D. 2007. *Under a Green Sky: Global Warming, the Mass Extinction of the Past, and What They Can Tell Us About Our Future*. New York: Smithsonian Books.

Weil, Simone. 1957. *Intimations of Christianity Among the Ancient Greeks*. New York: Routledge.

Weisman, Alan. 2007. *The World Without Us*. New York: Thomas Dunne Books/St. Martin's Press.

Zuboff, Shoshana. 2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York: PublicAffairs.