On the Intrinsic Value of Diversity

Seth D. Baum¹ and Andrea Owe¹,²

¹ Global Catastrophic Risk Institute
² Civilization Research Institute

Forthcoming in Inquiry, DOI 10.1080/0020174X.2024.2367247.
This version 17 June 2024.

Abstract: Diversity is an important ethical concept, but it is almost exclusively studied within two domains: biodiversity and diversity of sociological attributes such as race and gender. We provide a general study of the intrinsic value of diversity. We survey prior literature on the intrinsic value of biodiversity and sociological diversity in search of insights relevant to the intrinsic value of all types of diversity. We then present three thought experiments designed to clarify intuitions about the intrinsic value of small amounts of diversity, large amounts of diversity, and diversity as compared to other intrinsic values. We find that many types of diversity are intrinsically valuable at both small and large amounts, but that diversity may be a weak intrinsic value in comparison to others. Noting that diversity can be defined in many ways, we propose that, for purposes of moral evaluation, diversity should be defined to include both a diversity of individual elements within a group and a diversity of the overall patterns or structures of the group. Some of our findings about the intrinsic value of diversity are tentative, indicating that moral intuitions about diversity are sometimes ambiguous and would benefit from further study.

Keywords: diversity, intrinsic value, biodiversity, sociodiversity, ethics

1. Introduction

Is diversity intrinsically valuable? Prior research has debated the idea that certain kinds of diversity may be intrinsically valuable, especially biodiversity and various forms of sociological diversity (henceforth sociodiversity), such as diversity of race, gender, culture, or language. But, despite the great societal prominence of diversity issues, it is rare to consider the intrinsic value of diversity itself.¹ In contrast, utilitarians, for example, do not only argue that certain kinds of pleasure are intrinsically valuable; instead, they commonly argue that all pleasure is intrinsically valuable. Similarly, should all diversity be intrinsically valued? That is the question we will pursue in this paper.

A simple moral intuition is the idea that, all else equal, it is better to have two types of thing than one. This idea dates to at least Aquinas, who said ‘Just because an angel is better than a stone, it does not follow that two angels are better than one angel and one stone’.² This intuition has prompted study of non-additive population ethical theory, in which the intrinsic value of the n⁰ member of the species may be different, and generally greater, than the intrinsic value of the (n+1)⁰ member.³

---

¹ A recent wide-ranging survey of non-anthropocentric conceptions of intrinsic value concludes that the intrinsic value of diversity is “inadequately explored” (Owe et al. 2022: 21). We present detailed discussion of the prior literature below.
² The quotation is cited in Hurka (1983).
³ Again see e.g. Hurka (1983).
The concept of diversity builds on the idea that categories can be significant. As a starting point, a good definition of diversity is that of Stirling (2007, 708): diversity is ‘an attribute of any system whose elements may be apportioned into categories’, and whose value is a function of three parameters: (1) variety, being the number of categories, (2) balance, being the closeness of the numbers of elements in different categories, and (3) disparity, being the magnitude of the differences between categories. Per this definition, increases in any of these three parameters increase the amount of diversity in a system. Below, we also discuss definitions of diversity in which increases in balance reduce overall diversity because it reduces the diversity of the number of elements in different categories. Some work has quantified diversity in terms of aggregate disparity (Weitzman 1992) or the number of attributes of all elements in a system (Nehring & Puppe 2002). Other work includes interconnections between system elements, especially symbiotic relations between members of different species in the context of biodiversity (Næss 1989).

Common across all definitions is that diversity is an attribute of systems or groups, specifically an attribute that is based on differences between elements of the group. Therefore, evaluating the intrinsic value of diversity likewise requires a holistic moral perspective, bringing systems analysis to our moral intuitions. For example, it is not just that two angels may be less good than one angel and one stone. An angel and a stone may additionally be better than an angel and a deity, because angels and stones are more different from each other than angels and deities. Or, in more familiar terms, all else equal, it is better to preserve a species of rodent and a species of frog than it is to preserve a species of rodent and a species of marsupial, because marsupials are more similar to rodents, both being mammals.

After clarifying the concept of intrinsic value, the paper explores the intrinsic value of diversity through two approaches. First, we survey prior literature on biodiversity and sociodiversity in search of insights relevant to the study of all diversity. The survey is not intended to be comprehensive, but we nonetheless believe it covers many notable arguments that have been made for and against the intrinsic value of biodiversity and sociodiversity. It is not a large literature, despite the importance of diversity as an ethical concept. Furthermore, in seeking to extend these arguments to all diversity, we do not mean to assert that the authors of these arguments would agree with how we extend them. Perhaps some authors intended their argument to only apply to biodiversity (or, alternatively, sociodiversity), but we may nonetheless find their arguments insightful for the study of all diversity. Our approach is to mine the prior literature for ideas we find to be helpful, not to put words into other authors’ mouths. This part of the paper (Sections 3 and 4) does not have a single overarching narrative, but instead moves from topic to topic as these topics appear in the prior literature.

Second, we present three thought experiments designed to clarify intuitions about the intrinsic value of diversity. The first thought experiment adapts the isolation test of Moore (1903) via a space capsule that survives the destruction of the universe; it helps to clarify intuitions about small amounts of diversity. The second considers a box that maximizes diversity, intended to clarify intuitions about large amounts of diversity and on how diversity should be defined. The third considers a genie with the power to turn the entire cosmos into intrinsic value; it helps to clarify intuitions about the importance of diversity as compared to other moral values. These three thought experiments are presented in Sections 5-7.

Taken together, the two parts of the paper paint a clearer picture of the intrinsic value of diversity than was previously available. However, the paper does not provide a definitive argument about the intrinsic value of diversity. The issue is too ambiguous
and too understudied to be fully resolved in one paper. Instead, the paper serves as a
general resource for the study of the topic, compiling the small prior literature and
presenting new tools in the form of thought experiments. We also present our own
views and moral intuitions about the intrinsic value of diversity, though we caution that
our intuitions do not always point clearly in a certain direction and at times diverge
from each other. We do conclude strongly in favor of many types of diversity having
some intrinsic value, but the details are harder to resolve.

2. Intrinsic Value
Something is of intrinsic value when it has value in itself. Some thing X is intrinsic to
some other thing Y if X is an essential attribute of Y, such that X cannot be separated
from Y. Likewise, a thing Y has intrinsic value X if X is valuable and X derives from
some essential attribute of Y and not from anything extrinsic to Y. If Y has intrinsic
value, then it is good for Y to exist even if Y has no relation to anything else. For
example, a utilitarian might say that a person is intrinsically valuable if the person is
happy; in this case, the person is Y and the person’s happiness is X.

Diversity is commonly valued for its relations to other things. For example, a
diverse set of tools can be valuable because it is useful for completing complex tasks. In
this case, the value derives from the completion of the task and is therefore extrinsic to
the diversity of the tools. This type of extrinsic value is called instrumental value
because diversity is an instrument for achieving some other value (Bradley 1998).
Instrumental value is sometimes contrasted with intrinsic value, but it may be more
precisely contrasted with final value, meaning something valued as a final end, in
contrast with instrumental value as something valued as a means to some other end
(Korsgaard 1983). The distinction between intrinsic and final value is worth unpacking
further.

Rabinowicz and Rønnow-Rasmussen (2000) argue that when an object’s value
depends on its relation to something else, the object can have final value but not
intrinsic value. For example, the object Napoleon’s hat is considered valuable because
of its relation to Napoleon. This value is not intrinsic to the hat—absent its relation to
Napoleon, it is just another hat. Thus, Napoleon’s hat is of final value but not intrinsic
value. Meanwhile, a hat that is of intrinsic value (for whatever reason) would also be of
final value: the value that is intrinsic to the hat makes the hat worth valuing as an end.
In general, something of intrinsic value will also be of final value, but not necessarily
the reverse.

The evaluation of diversity poses a complication because diversity is a relational
category. An individual object does not have its own diversity in the way that it has its
own height or weight. Instead, diversity derives from the relations between multiple
individual objects. It would not make sense to assess the diversity of a frog or a human,
but it would for a frog and a human. Or rather, it only makes sense to assess the
diversity of a frog in terms of the collection of ‘sub-objects’ inside of it, such as the
diversity of microorganisms in its gut. Analogously, a collection of objects such as a
frog and a human can be considered as a type of object, which we will call a ‘group
object’. Diversity is intrinsic to group objects in the same way that height and weight
are intrinsic to individual objects.

---

4 Napoleon wore black bicorne hats made of felt and silk. Approximately 25 of his hats
remain; one was recently auctioned by Sotheby’s for $1.4 million (Dafoe 2021).
Finally, intrinsic value literature sometimes distinguishes between the intrinsic value of objects and the intrinsic value of states of affairs, such as the state of being diverse. This is the difference between the group object itself and the diversity of the group object. Is the group object intrinsically valuable because it is diverse, or is it the diversity itself that is intrinsically valuable? In the Introduction, we asked, should all diversity be intrinsically valued? This phrasing implies that it is states of affairs that are intrinsically valuable. Alternatively, we could ask, should all diverse group objects be intrinsically valued? That would imply that it is (group) objects that are intrinsically valuable. For purposes of this paper, we are not concerned with the distinction between objects and states of affairs. Our analysis seeks to inform either of these two perspectives on intrinsic value.

3. Biodiversity

‘Conscious of the intrinsic value of biological diversity’. So begins the preamble to the Convention on Biodiversity, which has been ratified by all UN member states except the United States. This is a notable data point showing broad support for the idea that biodiversity is intrinsically valuable. Similar support can be found in other sources, such as the Dasgupta Review on the Economics of Biodiversity (Dasgupta 2021), moral psychology research (Berry et al. 2018), and commentaries by scientists working in preservation biology (Soulé 1985; Ghilarov 2000). The intrinsic value of biodiversity is not universally embraced, but it is common enough to merit consideration. What follows surveys some literature on the intrinsic value of biodiversity, though a comprehensive review is beyond the scope of this paper.

One significant line of research on the intrinsic value of biodiversity is from holistic environmental philosophy. These studies do not focus on biodiversity as a separate category of intrinsic value. Instead, biodiversity is treated as one aspect of some broader conception of intrinsic value, such as the realization of nature’s potential (Næss 1989; Mathews 1991) the richness of ecological systems (Miller 1982; Mikkelson 2011, 2014), or the preservation and continuation of the evolutionary story of Earth-life (Tonn 2002). These holistic philosophies emphasize the interconnectedness of nature, such that it would be inappropriate to evaluate the intrinsic value of biodiversity separately from other aspects of nature.

These holistic philosophies can generalize to other types of diversity, but little insight is obtained. The arguments above are that biodiversity is intrinsically valuable because it is an essential part of an intrinsically valuable natural system. In that case, perhaps other types of diversity are valuable if they are essential parts of other intrinsically valuable systems. For example, perhaps sociodiversity is an essential part of society, and perhaps society is intrinsically valuable, and therefore sociodiversity is also intrinsically valuable. In contrast, perhaps a diversity of silverware is an essential part of a formal dining room, but perhaps formal dining rooms are not intrinsically valuable, and therefore a diversity of silverware is not intrinsically valuable. Per this reasoning, diversity is not intrinsically valuable just by the fact of it being an essential part of something else. To evaluate the intrinsic value of diversity using this approach, one would need to laboriously consider each type of diversity on a case-by-case basis. Alternative approaches would be helpful. Finally, if diversity is intrinsically valuable due to being an essential part of an intrinsically valuable system, then it stands to reason that other essential parts of an intrinsically valuable system would also be intrinsically valuable. Indeed, the arguments above also recognize other things besides biodiversity

---

5 See, for example, Rabinowicz and Ronnow-Rasmussen (2000).
as being intrinsically valuable aspects of natural systems, such as life itself. It follows that this perspective is fundamentally about essential parts of intrinsically valuable systems and not about diversity per se.

Outside of holistic environmental philosophy, a few other studies have considered the intrinsic value of biodiversity. Sober (1986) argues that the intrinsic value of biodiversity is insufficient to explain some environmentalists’ motivation to avoid species extinction. If protecting biodiversity was the only goal, then, in a world with many species, losing a few species would only be a small loss. Some environmentalists, however, worry a great deal about any species extinction. This is a reasonable point, but it does not imply that biodiversity is not intrinsically valuable. Instead, it only implies that, for some environmentalists, other things must be intrinsically valuable, instead of or in addition to biodiversity, such as species (Rolston 2020). This paper is not interested in the claim that diversity is the only thing that is intrinsically valuable, so this point can be set aside.

Boldt (2013) presents two ideas for why biodiversity may be intrinsically valuable: because it inspires awe and because there is a kinship relation among different forms of life. Boldt argues that awe is a reason to protect existing biodiversity but not a reason to create new biodiversity via advanced biotechnology, whereas kinship could be a reason to do either. However, neither awe nor kinship are reasons to intrinsically value diversity in general. Awe is not intrinsic to the diversity: it is rooted in the relation between the diversity and a (generally human) observer who experiences the awe. Kinship is only applicable to forms of diversity where a kinship relation exists and is therefore inapplicable to many forms of diversity.

McShane (2016: 161) argues that large amounts of biodiversity can be bad. McShane imagines ‘a world with too much biodiversity... in which there are so many different kinds of things so different from one another that there is hardly any unity, cohesion, or integrity possible within the natural world’. This makes sense: one would generally not want to keep adding different types of species to an ecosystem just to increase its biodiversity. However, the obvious reason to not do this is instrumental: excessive biodiversity harms the functionality of the ecosystem. Here, the intrinsic value can be taken to be something like ecosystem flourishing. For any given ecosystem, there is some instrumentally optimal amount (or amounts, or range(s) of amounts) that maximizes the ecosystem’s flourishing. Anything above that amount would be instrumentally harmful for the sorts of reasons McShane describes: unity, cohesion, etc. To assess the intrinsic value of large amounts of biodiversity, any instrumental effects should be set aside. We pursue this matter further in Section 6.

Finally, some debate about the intrinsic value of biodiversity focuses on its rhetorical value for achieving biodiversity conservation objects. This perspective takes as its starting point the idea that biodiversity conservation is an important goal. It then considers which rhetorical arguments are most effective at achieving this goal. For example, it has been said that ‘both intrinsic and instrumental values... are important arguments in stemming the tide of biodiversity loss’ (Reyers et al. 2012: 506). This debate is not about whether biodiversity is intrinsically valuable—instead, it is about whether the rhetoric of intrinsic value is instrumentally valuable toward the goal of biodiversity conservation. The debate may be vital for biodiversity conservation efforts, but it does not inform this paper’s project of analyzing the intrinsic value of biodiversity.

---

6 Sober writes in terms of environmentalists in general, but environmentalists hold a variety of views on these issues.
4. Sociodiversity

When studying the intrinsic value of sociodiversity, the striking thing that stands out is how little prior attention it has gotten. There has been plenty of research on other aspects of the ethics of sociodiversity. The instrumental value of sociodiversity is commonly recognized, such as in the financial benefits businesses may get from having diverse staff (Van Dijk et al. 2012) or the pedagogical benefits schools may get from having diverse student bodies (Yee 2014). Programs to support sociodiversity, such as affirmative action, have been defended in a variety of ways for the value they provide to disadvantaged populations and to society as a whole (Sher 1999; Van Dijk et al. 2012; Yee 2014). Political philosophy has studied the procedural implications of moral diversity within citizen populations (Muldoon 2017). These are all worthy lines of inquiry. Nonetheless, amidst all this research and general interest in sociodiversity, it is remarkable that the intrinsic value of sociodiversity has been seldom considered. What follows is a large portion of the literature we identified. Of course we cannot prove the absence of additional literature, but we nonetheless believe the intrinsic value of sociodiversity to be a remarkably understudied topic, especially given how important of a concept sociodiversity is.

Yee (2014:87) presents the clearest articulation we were able to find of moral intuitions underlying the intrinsic value of sociodiversity. Yee states that sociodiversity ‘In itself it is so beautiful. It is so glorious. It is so enriching. It is so special. We just like it. It is part of us. It is constitutive of society’. This passage is phrased in terms of ‘diversity’, not any specific type of diversity, but it can be inferred from the surrounding context, including the reference to ‘society’, that it is focused on sociodiversity. Regardless, the ideas here could apply to diversity in general. Other types of diversity might not be ‘constitutive of society’, but perhaps they are beautiful, glorious, and special, and ‘we just like them’.

Shin (2009) argues against the intrinsic value of racial diversity. Shin considers the case of a group of violent, hostage-taking bank robbers and claims that it would be absurd to care about the racial diversity of either the robbers or the hostages. Shin additionally claims that racial diversity would not matter for a group of people who collaborate to philanthropically support a disaster relief effort. Instead, Shin argues, racial diversity only matters in certain circumstances, such as education, and therefore its value is not intrinsic. It is correct that if the value of something is context-dependent, then its value depends on its relation to the context and therefore is not intrinsic to that thing. This reasoning applies to anything, including all types of diversity. However, we are not persuaded that there is no value in the racial diversity of robbers, hostages, or philanthropists. If, as Yee (2014) proposes, diversity is intrinsically beautiful, glorious, and special, why would this not apply to the racial diversity of robbers, hostages, and philanthropists? The arguments of Yee and Shin would appear to be incompatible with each other.

Bouville (2008) makes a similar argument as Shin. Bouville observes that women are underrepresented in prison populations and says therefore, to improve prison gender diversity, one would need to imprison more women. Bouville takes this to imply that sociodiversity is not the only intrinsic value and may not even be a particularly important one. This strikes us as a reasonable conclusion: at least in some situations, the intrinsic value of diversity may be outweighed by other factors. However, this does not resolve the question of whether diversity is of any intrinsic value.

Levy (2002) applies the biodiversity argument of Sober (1986) to cultural diversity. Levy proposes an ‘assimilation machine’ thought experiment in which members of a disadvantaged minority culture are painlessly and willingly transformed into members
of the dominant majority culture, such that the minority culture disappears. Examples given include Inuits or Australian Aboriginals transformed to Western Anglophone liberals. Levy posits that the transformation may actually benefit the transformed individuals because they would gain the advantages of belonging to the dominant culture. However, Levy observes—and we agree—that there is a moral intuition that finds the loss of a culture to be a bad thing even if its former members would benefit from the process. Following Sober, Levy argues that, in a world with many cultures, the intrinsic value of cultural diversity is insufficient to explain our aversion to the loss of a culture. This argument seems plausible, in which case cultural diversity cannot be the only reason to value a culture. However, as with Sober’s argument about biodiversity, Levy’s argument does not inform this paper because this paper is not concerned with the idea that diversity is the only intrinsic value.

Bauböck (2001) interprets Parekh (2000) as implying that cultural diversity is intrinsically valuable. The interpretation is based on cultural diversity offering people a wider range of perspectives and opportunities, including the means of reflecting on their own culture, and because intercultural dialog is easier in a society that embraces cultural diversity. However, none of these are intrinsic reasons to value cultural diversity. Instead, they are instrumental reasons, amounting to the idea that the existence and embrace of cultural diversity enables individuals to have better lives and societies to function more successfully.

Bouville (2008) and Sarkar (2010) present examples in which diversity appears to be bad. Bouville observes that, in a population in which most people are smart and healthy, diversity is increased by having more stupid and sick people. Sarkar does the same for wealth: in a wealthy population, increasing diversity requires more poverty. These examples are fundamentally similar and can be analyzed together. Essentially, they argue that when more of something (intelligence, health, wealth) is good, a diversity of it will often be bad.

Some complications with this argument are apparent. First, wealth is primarily valuable for instrumental reasons: it is useful for those who have it. Intelligence and health are also instrumentally valuable, though they could conceivably also be intrinsically valuable, so for simplicity we focus here on wealth. It is hard to disentangle the instrumental value of wealth from any potential intrinsic value of wealth diversity. Doing so may benefit from supposing that wealth had no instrumental value. Consider two different economic policies. Policy EP1 contains two wealthy families. Policy EP2 contains a wealthy family and a poor family. Suppose that all four of these families are equally happy. 

This is perhaps unusual but not implausible: it is possible to be poor but happy or wealthy but unhappy. In that case, which policy should be considered to be better? EP1 results in more wealth; EP2 results in more diversity of wealth; both have equal happiness. It is not implausible to suppose that EP2 is better. Perhaps it is good to have a world in which people achieve happiness in a diversity of ways. In that case, wealth diversity is good. However, it is not intrinsically good: it is instrumentally good to the extent that it promotes a diversity of ways of achieving happiness. It is less obvious that EP2 may be intrinsically better, though it is similarly non-obvious that EP2 may be intrinsically worse. Removing the instrumental value of wealth weakens Sarkar’s argument against wealth diversity. The same also applies to the argument against intelligence and health diversity, though this is complicated by the potential intrinsic value of intelligence and health.

Second, the argument pertains to diversity of amounts of wealth. In contrast, intuitions for the intrinsic value of biodiversity and sociodiversity are typically focused

---

7 Or equally flourishing, or living equally worthwhile lives, or any other such metric.
on the *types* of life and humans. Here is an illustration of this point. The aim of biodiversity preservation is generally not to have a diversity in the number of living beings. Consider two different biodiversity preservation projects. Project BP1 would result in 2,000 living beings in one ecosystem and 2,000 living beings in another ecosystem. Project BP2 would result in 2,000 living beings in one ecosystem and 1,000 living beings in another ecosystem. BP2 has greater diversity in the number of living beings in different ecosystems, which is a certain type of biodiversity. Likewise, the choice between BP1 and BP2 resembles the choice between EP1 and EP2 discussed above, which involves diversity in amounts of wealth. However, the choice between BP1 and BP2 is fundamentally different from the choices considered in standard discussions of biodiversity. In standard discussions, diversity in the types of living beings is central. Indeed, we are unaware of any discussions of biodiversity in which there is only diversity in the number of living beings, not also in the types of living beings; ditto for discussions of sociodiversity. Amounts can factor, as in the Stirling (2007) concept of balance, but here the amounts are related to different types, so there is still a diversity of types. In contrast, the argument against intelligence/health/wealth diversity involves only a diversity of amounts. Existing moral intuitions for the intrinsic value of biodiversity and sociodiversity, being rooted in type, may not apply to the intrinsic value of diversity of amounts of wealth.

A challenge here is that the distinction between amount and type is sometimes blurry. For example, carbon monoxide and carbon dioxide are different types of molecules that differ in terms of their amounts of oxygen atoms. Unicellular and multicellular organisms are different types of organisms that differ in terms of their amounts of cells. However, this blurriness does not apply to the arguments of Bouville and Sarkar, which are clearly about amounts, or to standard treatments of sociodiversity, which are largely about types (and the amounts of people of different types).

Third, the argument is sensitive to how diversity is defined. Under some definitions, diversity of amounts reduces total diversity. For example, Stirling (2007: 709) states, ‘All else being equal, the more even is the balance, the greater the diversity’. This means that diversity of amounts results in a smaller amount of diversity. According to this, EP1 and BP1 are actually more diverse than EP2 and BP2. It follows that the argument is not a rejection of the intrinsic value of diversity per se, but instead is a rejection of the intrinsic value of certain conceptions of diversity. When diversity is conceived in a way that increases when balance increases, then the argument aligns with the idea that diversity is intrinsically valuable. It is fair to question whether diversity should be defined such that total diversity increases when diversity of amounts is reduced. However, it is clear that the arguments of Bouville and Sarkar do not apply to all conceptions of diversity.

Could there be intrinsic value in diversity of *types* of intelligence, health, and wealth? Intelligence can come in different types: literacy, numeracy, creativity, and so on. Ditto for health: cardiovascular health, muscle strength, mental health, etc. And for wealth: bank accounts, stocks, bonds, real estate holdings, and so on. Consider wealth. All else equal, is it better to have a diversity of types of wealth? Compare wealth portfolio WP1 containing $2,000,000 in stocks and wealth portfolio WP2 containing $1,000,000 in stocks and $1,000,000 in bonds. Once again, the instrumental value of wealth complicates the analysis: there can be instrumental value to having a diversified financial portfolio. Suppose, then, that this does not apply to WP1 and WP2: both are of equal instrumental value by virtue of having identical financial attributes (such as risk and liquidity) and by bringing identical outcomes (such as financial security and happiness) to their owners. Is there a case for favoring WP2 over WP1 deriving from
the intrinsic value of wealth diversity? Similarly, could there be a case for intrinsically valuing a diversity of types of intelligence or health? Perhaps. Bouville’s and Sarkar’s arguments do not apply to these questions because their arguments are about amounts. At a minimum, it is not obvious that there is not intrinsic value in diversity of types of intelligence, health, and wealth.

One last point about the arguments of Bouville and Sarkar: they are about stupidity, sickness, and poverty, which are all things that are generally taken to be bad. The diversity of bad things poses distinct challenges, which we revisit below.

Another compelling sociodiversity topic is moral diversity, meaning the diversity of moral views held across a human population. Moral diversity has been a subject of compelling research in political philosophy (Muldoon 2017) and moral psychology (Haidt et al. 2003), though to our knowledge the intrinsic value of moral diversity has not been considered. It does pose an interesting quandary. Moral views commonly involve views about intrinsic value. A diversity of moral views can mean a diversity of views about intrinsic value. The intrinsic value of moral diversity can therefore mean that there is intrinsic value in the existence of a diversity of views about intrinsic value. This is not implausible. One could see moral diversity as another aspect of the beautiful tapestry of human society, something glorious and special in the sense of Yee (2014). Alternatively, one could posit, as Haidt et al. (2003) find, that moral diversity is different. For example, if one is confident in one’s own moral views, then moral diversity would imply that other members of society have inferior views and are pursuing immoral or at least suboptimal ends.

Finally, Sher (1999) dismisses the intrinsic value of sociodiversity as an unsound concept. Sher sees intrinsic value arguments as mere personal opinion based on aesthetic taste and not something suitable for moral evaluation. Surely this is mistaken. The same argument could be made to any notion of intrinsic value, in which case a rather large portion of moral philosophy would be lost. Indeed, the entirety of moral philosophy may be lost if one cannot appeal to one’s own moral intuitions—though whether this is true is not a matter we will entertain. For our purposes, the important part is this: yes, belief in the intrinsic value of diversity may ultimately derive from moral intuitions, and no, that is not a problem.

5. Thought Experiment 1: Space Capsule Isolation Test

Having considered a range of prior arguments about the intrinsic value of diversity, we now turn to the first of three thought experiments to clarify intuitions about it. Moore (1903: Section 112) proposed an isolation test for gauging whether something is intrinsically valuable. To our knowledge, the isolation test has not previously been applied to diversity.⁸

To help bring the isolation test to life, we present it in terms of a space capsule thought experiment. Suppose that the universe is about to be destroyed. It will be pulverized into a fine dust that then evaporates into nothingness. The entire planet Earth will be destroyed, along with everything and everyone on it: humans, other species, the continents, the oceans, the planetary core, everything. The moon and the other planets and the Sun and the stars and black holes and galaxies and everything else will no longer exist. It will be an empty universe. Before the universe is destroyed, humanity can take one last action. Specifically, it can send a capsule into outer space. The capsule is built out of a special material that will survive the destruction of the universe.

⁸ Bouville (2008) briefly notes the isolation test for gauging the intrinsic value of diversity but does not explore the test in any detail.
preserving itself and its contents in perpetuity. The question is what to put into the capsule.

This space capsule is intended as a thought experiment and the details probably violate multiple laws of physics. However, there are some actual or plausible potential situations that resemble the thought experiment. Mass destruction is threatened by a variety of forces such as nuclear warfare and large asteroid collision. The most destructive threat may be gamma ray bursts caused by explosions of large stars. These events may destroy large portions of the galaxy, and there may further be some exotic technology that can withstand them (Ćirković & Vukotić 2016). Additionally, the Voyager spacecrafts, launched by NASA in 1977, both carry golden records containing sounds and images about Earth. The content of the records was selected to convey information about Earth, not to materialize philosophical thought experiments. Nonetheless, the space capsule thought experiment is at least vaguely in the vicinity of decisions that humanity could actually face. Of course, as a thought experiment, it does not need to be realistic.

Now, we can consider various objects to put into the space capsule to help us evaluate the intrinsic value of diversity.

Let us start with something banal. Suppose the only options for objects to put into the space capsule are ordinary household objects such as a cup, a ball, and a shoe. There is diversity in having a cup, a ball, and a shoe instead of three cups, three balls, or three shoes. Under ordinary circumstances, it seems hard to argue that this diversity is intrinsically valuable. It’s still just a cup, a ball, and a shoe. However, the space capsule is an extraordinary circumstance.

What should humanity do? We believe humanity should put the cup, the ball, and the shoe in the space capsule instead of three cups, three balls, or three shoes. If the space capsule and its contents are to be the only surviving artifacts of human civilization, and indeed of the entire universe, then we believe it is better for there to be a more diverse collection of artifacts. We believe this despite recognizing that no one will ever be affected by these artifacts. They will just exist. Furthermore, we find a diminishing marginal value of artifacts in the space capsule: the second surviving cup contributes less value than the first, and the third still less than the second. Our intuition is based on the fact that each successive cup adds successively less to the diversity of the contents of the space capsule. We likewise see value in the diversity of the three different types of objects: cups, balls, and shoes.

Is this value intrinsic? Recall Rabinowicz and Rønnow-Rasmussen’s (2000) argument that the value of Napoleon’s hat is not intrinsic because it derives from its relation to Napoleon, even though Napoleon no longer exists. Similarly, perhaps we only value the diversity of the cup, the hat, and the shoe due to its connection to human civilization. They are not just a cup, a hat, and a shoe; they are the cup, hat, and shoe that constitute the last remaining artifacts of human civilization. Perhaps their value derives not from their intrinsic properties, but instead due to their relation to human civilization. One might observe our preference for cup/ball/shoe over three cups, three balls, or three shoes as an indication of value that is not rooted in the relation to human civilization and may therefore be intrinsic, but it is still the case that this relation is a complicating factor.

Let us then consider something more exotic. Suppose the only options for the space capsule are types of objects that never previously existed and indeed can only exist inside the space capsule. Again, there are three options. Humanity can choose one blargh, one criftula, and one dombit, or it could choose two of one and one of another, or three of one. No information is available about blarghs, criftulas, and dombits except
that they are different from each other. Again, we believe humanity should choose one of each. We have no idea what these objects are. However, our intuition is that if objects will continue to exist in the universe, all else equal, it seems better for there to be a diversity of objects.

It stands to reason that this diversity is an intrinsic value. The blargh, criftula, and dombit have no relation to anyone or anything else except for the act of humanity selecting them for inclusion in the space capsule. That act of selection is inherent to moral agency. If the value of this diversity is extrinsic, then it would seem that there can be no intrinsic value in anything involving moral agency. That seems incorrect. It seems that the value of the diversity is intrinsic to the group object consisting of one blargh, one criftula, and one dombit, whatever those things are.

One can go one step further and remove the moral agency, just to be sure. Suppose that, instead of humanity selecting among blarghs, criftulas, and dombits for the space capsule, the selection is made via a random process that humans have no control over. There is no opportunity to decide what is put in the space capsule, but one can still have preferences about what happens to end up in it. In this case, we continue to favor the space capsule containing one blargh, one criftula, and one dombit due to our valuation of the diversity. This value would seem to be definitively intrinsic. Therefore, we conclude with confidence that yes, according to our own moral intuitions, in at least some cases, diversity is intrinsically valuable. We expect that our moral intuitions are not unusual in this regard.

What about more familiar cases? Suppose the options were three biological organisms and the decision was between more or less biodiversity. Or, suppose the options were three human beings and the decision was between more or less sociodiversity. Assume that these various individuals would somehow continue to live out their lives in the space capsule, and that their quality of life would be unaffected by whatever else was in the space capsule. We would favor the options with greater diversity.

We can revisit scenarios from earlier in the paper. Suppose the three people to enter the space capsule were all violent bank robbers and the decision was to select a set of bank robbers that is more racially diverse or a set that is less racially diverse. One might lament that the only people who get to survive the destruction of the universe are violent bank robbers, but if that is the only option, and their lives in the space capsule would be the same either way, then it does seem better to favor the racially diverse group of bank robbers. Ditto for hostages or philanthropists, though without the lament. We would likewise favor people with diverse types of intelligence, health, and wealth, though it is hard to imagine the meaning of wealth in a space capsule. We would even favor people with diverse amounts of intelligence, health, and wealth, as long as it was assured that they would somehow have lives of equal quality. We would also favor people with diverse moral views, though moral diversity may be of limited significance in the space capsule. Indeed, as long as all else is equal, we find ourselves favoring diversity across a wide range of cases.

The one type of case that we struggle with involves negative intrinsic value. To clarify, by negative intrinsic value, we mean something so bad that it would be better if it did not exist. Negative intrinsic value is worse than nothingness. An example could be excruciating pain—the sort of pain that doctors would anesthetize so the patient did not have to endure it, the sort of pain that, were it all someone would ever experience, it would make their life worth not living, to the point that euthanasia may be appropriate.\footnote{See negative utilitarianism and related theories (Smart 1958; Benatar 2006).} Suppose the only options for the space capsule were three people experiencing...
excruciating pain. Suppose further that there must be three people on the space capsule: one cannot choose to have zero people on the capsule such that the people could be put out of their misery. Is it better to select three people experiencing different types of excruciating pain or three people all experiencing the same type? We struggle to answer this question. We tentatively think one type of pain would be better. Perhaps there is a disvalue to the existence of a type of pain, making it better to have fewer types. But the issues are less familiar. We are unaccustomed to thinking about the diversity of bad things, and our intuitions feel less reliable. Of course, we hope to never have to make such a decision.

6. Thought Experiment 2: Maximization Box

The space capsule has been fruitful for exploring the intrinsic value of small amounts of diversity. However, there is reason to believe that it may be different at larger amounts. Recall the McShane (2016) argument that too much biodiversity makes things worse, which could be due to either intrinsic or instrumental disvalue of large amounts of biodiversity. Here we present a thought experiment for assessing moral intuitions about large amounts of diversity.

Imagine a box that maximizes the diversity of its contents. The box is used by putting some group object inside of it. The box automatically changes its size to accommodate any arbitrarily large group object. One could put into it the entire global human population, or the Amazon rainforest, or the Milky Way galaxy. Whatever group object is put inside, the box immediately transforms it. Furthermore, the box only transforms the diversity of its contents. It does not affect the contents in any other way. The box also does not change the extrinsic value of the diversity of its contents. Finally, putting group objects inside the box is trivially easy, requiring no effort other than the decision to do so. The question is whether it is good to put group objects into the box. Because the box maximizes the diversity of its contents and does not change its contents in any other way, the decision of whether to put group objects in the box raises the question: all else equal, is it good to maximize diversity?

We find two countervailing moral intuitions about the maximization box. First is the intuition that more is better. If diversity is intrinsically good, then more diversity should be intrinsically better. In general, intrinsic value seems like a thing one should want to have more of in the world. Perhaps there is not always a duty to maximize intrinsic value, but if maximizing intrinsic value is as easy as deciding to put something in the maximization box, then that seems like a good thing to do. Perhaps there are some types of intrinsic value that should not be maximized, but we find it reasonable to suppose that maximizing diversity is intrinsically good, when all else remains equal.

Second is the intuition for a ‘happy medium’ in which there is neither too little nor too much of something. For diversity, too little may be too simple, too monotonous, too lacking in richness; too much may be too disjointed, too incohesive, just ‘too much’. Recall the Yee (2014) intuition that diversity is beautiful, glorious, and special. Perhaps this only applies at moderate amounts of diversity. At high amounts, diversity may lose its beauty. One generally wants, for example, art that is not too plain or too messy: there is an intermediate ‘just right’ amount of variation within a work that produces pleasing patterns, symbols, etc. It is important to distinguish between aesthetic and moral valuation. However, if the moral intuition in favor of diversity is rooted in something akin to beauty, then the analogy to art may indicate something important about the intrinsic moral value of diversity.
In evaluating these two intuitions, we find ourselves favoring ‘more is better’. Though we do not entirely dismiss ‘happy medium’, we do have some concerns with it. First, in the real world, a happy medium of diversity is of clear instrumental value. Large amounts of diversity can reduce the functionality of ecosystems, social systems, and other types of systems. It is therefore intuitive to assume that large amounts of diversity are bad. In the maximization box, large amounts of diversity are instrumentally neutral, but this is counterintuitive. Upon recognizing this instrumental neutrality, we are less inclined to support the ‘happy medium’ intuition about intrinsic value. Second, we are concerned that the ‘happy medium’ intuition may be primarily aesthetic. It clearly has an aesthetic component. This aesthetic component diminishes the ethical significance we attribute to the ‘happy medium’ intuition, though it may not eliminate it entirely. Third, we are concerned that the extreme complexity of highly diverse systems confuses our intuitions. It is difficult for us to wrap our minds around such high complexity. To the extent that we are able to do so, we find ourselves favoring more diversity. In other words, the ‘happy medium’ intuition seems to derive from limitations of our minds, not from something intrinsically inferior about large amounts of diversity. We can even imagine that if we had more capable minds, we would be more confident in ‘more is better’. Meanwhile, we find that the general case for ‘more is better’ is quite strong, for the same reasons it is in other corners of moral theory. In practice, maximizing diversity may not always be better due to the effort required and due to conflicts with other moral values. However, with the maximization box, which avoids these issues, the case for maximizing diversity seems compelling. That said, we find this to be a difficult question and that our intuitions do not provide a definitive answer.

Some more specific issues can be seen through an example. What would it mean to put in, say, the entire global human population? Consider linguistic diversity. Suppose the maximization box transforms everyone into bilinguals who all speak one common language (to ensure continued functionality of society) plus a second language. The number of languages that humans can speak is virtually unlimited, as long as one includes languages that have never previously existed but could be learned by human minds. So, the maximization box could have everyone speak a different second language, with each being a language isolate. Or, it could maximize the interconnectivity between the second languages via a rich language family. Or, it could maximize variation in the amounts of people who speak various second languages. All of these would be compatible with different definitions of diversity.

One takeaway is that, at large amounts of diversity, the definition of diversity can matter a lot. Parameters include variety, balance, and disparity (Section 1, Stirling 2007), interconnectivity between elements (Section 1, Næss 1989), and variation in types vs. amounts (Section 4).

Here is our preliminary intuition on how diversity should be defined for purposes of moral evaluation. In other words, this is the type of diversity that we tentatively think should be maximized, at least in the context of the maximization box, where other factors are not applicable. We stress the tentative nature of these intuitions. Refining and critiquing them would be a worthy subject of future research.

Our overall orientation is toward reconciling diversity of individual system elements with diversity of the overall system pattern or structure. Maximizing individual diversity can result in a homogeneous system pattern. For example, maximizing variety

---

10 Human diversity exists across many dimensions, of which language is just one. Here we focus on linguistic diversity to explore a specific issue related to diversity maximization. If the entire global human population could be put into the maximization box, the full range of dimensions of diversity would be pertinent.
could entail each system element being of a different type, while maximizing balance could entail the same number of system elements for each type. Both of these produce a homogeneous pattern with zero diversity of amounts. Similarly, maximizing disparity could entail a homogeneous pattern of each type being completely unrelated to each other type, whereas a diversity of disparities yields a more complex pattern. We believe that when maximizing diversity, it is important to have a diversity of both individual elements and overall system patterns. This holds especially for large systems, such as those considered above for the maximization box.

To illustrate this, here is a more specific example, concerning the tradeoff between the number of types (i.e., variety) and the number of amounts. In the space capsule thought experiment, the number of types is maximized by choosing one blargh, one criftula, and one dombit, whereas the number of amounts is maximized by choosing two of one (e.g., two blarghs) and one of another (e.g., one criftula). As discussed in Section 5, we favor one blargh, one criftula, and one dombit over two of one and one of another. For such a small number of system elements, the value of variety outweighs the value of the more complex pattern from a diversity of amounts. However, at larger numbers of system elements, we would sacrifice some variety for a more complex pattern. For example, we would not choose one of each of 100 different types. Instead, we would choose something in the range of 80 to 90 types, with 10 to 20 types sacrificed in order to have a diversity of amounts. The 80 to 90 range is not precisely derived but nonetheless conveys our underlying intuition, which is for a high degree of emphasis on diversity of types while retaining some emphasis on diversity of amounts so as to produce a more complex and diverse pattern.

Here is another example, concerning degrees of interconnectivity. Interconnectivity is maximized by having all system elements connected to all other elements. However, that results in a homogeneous pattern in which each element has the same number of interconnections. For sufficiently small systems, we favor connecting all elements, for example with three elements: \{A-B, A-C, B-C\}. However, for larger systems, we favor removing some connections to increase the diversity of degrees of interconnectivity.

7. Thought Experiment 3: Cosmic Genie

The space capsule and maximization box have served to refine our intuitions about the intrinsic value of diversity on its own. This leaves open the question of how the intrinsic value of diversity compares to other intrinsic values or other moral goals. Our last thought experiment aims to refine intuitions on these comparisons.

Imagine a genie that will grant a single wish. Specifically, it will convert the entire cosmos into some configuration that optimizes for moral value. In other words, whatever one takes to be the morally best way to configure the universe, the genie will arrange the cosmos accordingly. In doing so, the genie can rearrange all the atoms and molecules in the universe. Laws of physics still apply, such as gravity and conservation of mass-energy. The genie otherwise has total control over the entirety of the universe. The question is what to ask the genie to do.

The cosmic genie is hypothetical, but humanity may someday face similar questions. The cosmic genie thought experiment is based on debate about very powerful future technology, especially advanced artificial intelligence (AI). Some research has proposed that advanced AI could not only take over the world, but could further reconfigure the accessible portions of the cosmos in ways that are similar to the cosmic genie. The research has mainly focused on mistakes that humans could make in designing such AI, resulting in pathological outcomes such as tiling the universe with
smiley faces instead of with proper happiness or wellbeing (Loosemore 2014). The smiley face example pertains specifically to utilitarianism, but the underlying idea also applies to other moral theories. For purposes of designing powerful technology, avoidance of mistakes is of clear importance. However, for our purposes, we wish to set aside the mistakes and focus on what goals would be pursued if there would not be any mistakes. Therefore, we assume that the cosmic genie would correctly implement whatever moral theory is requested of it.

As a starting point, consider the prospect of tiling the universe with value—not a false value, like smiley faces, but a true value, like wellbeing. The word ‘tiling’ is important. Tiles are elements that repeat over and over again to fill some space. To maximize some notion of intrinsic value, such as wellbeing, it may be the case that certain patterns of atoms and molecules would repeat over and over again across the cosmos. Perhaps there is some configuration of atoms and molecules that maximizes wellbeing locally, and that to maximize wellbeing across the entire cosmos, that configuration would be repeated over and over again. The result would be extremely repetitive, but it may result in the highest possible amount of that notion of intrinsic value.

The tiling pattern could result from a variety of notions of intrinsic value, such as in various forms of utilitarianism or ecocentrism (in which ecosystem flourishing is intrinsically valued). Whether a tiling pattern would result depends on the details of the framework as it relates to the space of possible configurations of atoms and molecules. For example, some forms of ecocentrism include an accounting of biodiversity and therefore may not tile. However, tiling would rate low in diversity. If diversity is intrinsically valuable, then it may be appropriate to ask the cosmic genie to do something other than tile the universe with value. However, doing something else may result in less of other notion(s) of intrinsic value: less wellbeing, less ecosystem flourishing, etc. In that case, the question of what to ask of the cosmic genie entails making a tradeoff between diversity and other intrinsic value(s).

Our own moral intuitions diverge on how to make this tradeoff. One of us would make the tradeoff so as to maximize the other intrinsic value(s). In this respect, other intrinsic values are favored lexicographically, with diversity serving only as a tiebreaker. In other words, the genie is instructed to first maximize the other intrinsic value(s) and then only maximize diversity if doing so would not reduce the other intrinsic value(s). Separately, one of us would make the tradeoff to maximize some weighted combination of diversity and the other intrinsic value(s), with diversity getting some nonzero weight. However, even then, the weight to be placed on diversity would be small. The intuition here is a willingness to sacrifice some of the other intrinsic value(s) to increase diversity, but not much. We both agree that when other intrinsic value(s) are at stake, the intrinsic value of diversity should not be a factor of high importance.

Given the variation in our own intuitions, it does seem reasonable to suspect that others may have a range of views on this. In other words, there may exist moral diversity for intuitions about how to make tradeoffs between diversity and other intrinsic value(s).

---

11 Alternatively, it may be the case that intrinsic value is maximized via other patterns, such as a small number of very large objects. For example, it has been proposed that advanced civilizations could create massive, planet-sized ‘Jupiter brains’ (Sandberg 1999).

12 On cosmic-scale ecocentrism, see Owe (2023).
8. Conclusion

Here are our primary findings. First, whereas prior literature has focused on the intrinsic value of specific types of diversity, especially biodiversity and (to a lesser extent) sociodiversity, we find that it is meaningful to study the intrinsic value of all types of diversity. This can be done by generalizing insights developed for specific types of diversity and by considering certain general-purpose thought experiments. Second, many types of diversity are intrinsically valuable, including biodiversity, sociodiversity, the diversity of banal objects, and the diversity of objects of unknown composition. Third, diversity of disvalues, such as excruciating pain, may be intrinsically disvaluable. Fourth, if diversity is intrinsically valuable, then more diversity is more intrinsically valuable when all else remains equal, even at very large amounts of diversity. Fifth, for purposes of moral evaluation, it may be best to define diversity in terms of a balance between the diversity of individual system elements and the diversity in the overall patterns and structures of the system. Sixth, diversity is of low importance compared to other intrinsic values. Seven, findings three, five, and six are less certain and more in need of further study.

The study of the intrinsic value of diversity is compelling for both practical and intellectual reasons. In practical terms, it is compelling due to the societal importance of diversity issues. Biodiversity and sociodiversity get almost all the attention, but perhaps there is a case for attention to other types of diversity as well. Furthermore, if advanced technologies ever permit anything akin to the cosmic genie, it goes without saying that a firm understanding of the intrinsic value of diversity could be of paramount importance. In intellectual terms, diversity stands out as a distinctive type of intrinsic value. Most conceptions of intrinsic value are centered on some type of object: a happy individual, a flourishing ecosystem, etc. In contrast, diversity is about patterns of objects within a group. Diversity is likewise applicable to other conceptions of intrinsic value: diverse types of happiness, diverse types of flourishing ecosystems, etc. The intrinsic value of moral diversity raises further puzzles.

This paper is unusual in its focus on the intrinsic value of all forms of diversity. As such, it may raise more questions than it answers. These questions are worthy ones to pursue in future research.

Acknowledgements:
We thank McKenna Fitzgerald, Jessica Hodgson, Matthew Rendall, and Mark Coeckelbergh for helpful feedback on an earlier version of this manuscript. Any remaining errors are the authors’ alone.

References


