# Two notes on Axiological Futurism: The importance of disagreement and methodological implications for value theory

Björn Lundgren [a,b,c,*,1]

[a] Department of Philosophy and Religious Studies, Utrecht University, Utrecht, the Netherlands
[b] Institute for Futures Studies, Stockholm, Sweden
[c] Department of Philosophy, Stockholm University, Stockholm, Sweden

ARTICLE INFO

ABSTRACT

Recently, John Danaher defended a research agenda he calls "Axiological Futurism". In this commentary, I propose two revisions for how that research agenda is defined. I argue that it ought to focus on disagreements and comparison of different axiologies.

## 1. Introduction

Recently, John Danaher (2021) defended a research agenda he called "Axiological Futurism", which he defined as follows:

> The systematic and explicit inquiry into the axiological possibility space for future human (and post-human) civilisations, with a specific focus on how civilisation might shift or change location within that axiological possibility space. (p. 4)

The focus of this paper is on two things. First, I will discuss what I take to be missing from the Axiological Futurism research agenda (i.e., granted that we should engage in the study of Axiological Futurism). This is strictly connected to how Danaher defines Axiological Futurism, in which the focus is on the axiological possibility space for civilisations. I will provide a re-conceptualisation of Axiological Futurism. Second, I will discuss some implications of what we must do if we pursue the Axiological Futurism agenda, which has further implications for how we define the project of Axiological Futurism.

Before we turn to my comments, I need to briefly describe the Axiological Futurism project in a bit more detail. Danaher's article lays out the foundation and motivation for a very ambitious project. Simply put, the motivation behind Axiological Futurism is the historical changes in what people, in the past and now, value. Because of this historical divide, it seems reasonable to expect—or so Danaher argues—that there may be similar changes in the future. Predicting and preparing for such changes, therefore, may be important. For example, there has arguably been a global change in the view of issues such as slavery and in some parts of the world the view on power dynamics in society and families has changed radically over time.

Danaher envisions four different but complementing routes by which the methods and methodology for the Axiological Futurism project can proceed. First, we can focus on mapping the logical space of different axiologies. Second, we can focus on investigating the causes of change. Third, we can do interdisciplinary work that combines the first and second approaches. Fourth, we can turn to pitfalls

---

* Corresponding author at: Department of Philosophy and Religious Studies, Utrecht University, Utrecht, the Netherlands.
  E-mail addresses: b.a.lundgren@uu.nl, bjorn.lundgren@iffs.se, bjorn.lundgren@philosophy.su.se.
[1] ORCID: https://orcid.org/0000-0001-5830-3432

that we ought to avoid. Danaher gives a very brief description of the pitfall approach (giving examples such as group thinking and "being too wedded to one set of values", p. 9). Broadly understood, the importance of potential pitfalls for future axiologies will be central to both of my comments.

Lastly, I do not believe that anything I say in this article will be in tension with Danaher's aim. Indeed, Danaher can fully endorse all of what I say. The important focus here is on the improvement of the concept of Axiological Futurism.

## 2. Re-conceptualising Axiological Futurism

In this section, I will focus on how Danaher has defined the Axiological Futurism project. In particular, I will suggest that Axiological Futurism would benefit from a shift from the focus on civilisations to that of individuals and groups of people. To understand my reasons for this, it would be best if we could determine what Danaher means by "axiological possibility space for future human (and post-human) civilisations" (p. 4). Unfortunately, Danaher does not unpack what it means for a civilisation to have a specific axiology. On the one hand, Danaher might be using "civilisation" as a broad label that includes the subsets that we tend to think are part of a civilisation (e.g., subcultures, institutions, and people). But if so, the definition of Axiological Futurism would benefit from being clarified, so that it is clear that these subparts are the object of study.

On the other hand, Danaher might have the whole civilisation in mind (i.e., ignoring the subsets), but if so then I think the definition needs to be altered. A potential problem of Axiological Futurism, under this interpretation, is that it is not clear that talk of axiology on the level of a civilisation is sensible. Indeed, normative theories often include an axiology, and people can agree or disagree about some axiology, but it is not clear that there is a sensible set of facts that corresponds to an axiology of civilisation, simply because civilisations tend to include individuals with a divergence of ideas. This may, of course, speak in favour of the former interpretation, but if that is what Danaher had in mind, then a reformulation is warranted irrespectively. An alternative interpretation is that what Danaher has in mind is that there may be social norms on the level of a civilisation, which—in turn—imply certain possible axiologies. That would be a sensible way of understanding of Axiological Futurism, but as I will argue, it would be more fruitful to focus on social norms on the level of individuals. Following Bicchieri (2017), a social norm x is a shared second-order belief such that other people think that x should be the norm. This implies, as empirics shows, that we can have a social norm that an absolute majority disagrees with. As I will argue, such individual or group-level disagreements seem to be more important than what holds on the level of a civilisation. As Danaher argues "we cannot run civilisation-wide experiments on possible future axiologies" (p. 7). As I will argue, this shouldn't even be the aim; rather we should explicate or re-conceptualise Axiological Futurism to shift the focus from the level of human or post-human civilisation to a more detailed description of the axiological possibility space on the individual level or axiologically coherent groups. Indeed, below I will present four arguments for why disagreement on interpersonal levels is important and should be an explicit part of the project of Axiological Futurism.

First, axiological disagreement can be the foundation of a lot of serious controversies and may increase the risk of war between nations, distrust among people, and so forth. One obvious example is religious disagreement, which, accompanied by different axiologies, has been a causal contributor to many conflicts throughout history. In one sense, many wars can be viewed as a conflict of disagreement. For example, consider the second world war, in which Nazi Germany gave value priority to what they called "Aryan" people, heavily mistreating non-Aryans and going to war to create "Lebensraum" (living space) for Aryans. This axiology was, of course, in disagreement with the axiology of most other nations and people.

Of course, the previous example is focused on the level of civilisation; however, it is also clear that a focus on the level of civilisation can miss central aspects of contention. The political polarisation today in the western part of the world clearly illustrates the importance of focusing on disagreements within a civilisation (and within nations). If we want the Axiological Futurism agenda to be fruitful, it would be better to focus on understanding the axiological tensions *within* a civilisation, amongst people, rather than focusing on the level of civilisations. (Again, this may be what Danaher already has in mind, but then an explication is further warranted.)

Second, disagreement can also serve as a safeguard against incorrect and harmful axiologies. Indeed, although it is logically possible to share an axiology that devalues oneself, there tends to be disagreement and resistance amongst any set of people that are mistreated by a specific axiology (e.g., women tend to disagree with misogyny, minorities tend to disagree with racism and so forth).[2] Thus, we can expect disagreement about the axiology when it is biased against a particular set of people (i.e., who they are, their lifestyle, etcetera).

Third, disagreement is a route through which we can change and improve. For example, it was by disagreeing with racist norms that Rosa Parks' actions sparked a change in the norms of segregation in the USA. Arguably, this reason is in line with Danaher's focus on the investigation of the possible mechanism of changing axiologies.[3]

Fourth, disagreement can also tear a society apart through political polarisation. We can envision a future in which axiologies (on average) improve, but the disagreements are so extreme that society, civilisations, humans, and mankind are worse off.

Thus, based on the above reasons, I suggest reformulating the Axiological Futurism research agenda to capture the importance of disagreement. I propose the following formulation:

---

[2] Of course, resistance is not a necessary disposition. Sometimes groups of people adapt their preference to what is expected from them (see, e.g., Nussbaum, 2000, pp. 111–166 for some interesting examples).

[3] See Baker (2019) for further examples. According to Baker, *disagreement* (even if he doesn't use that term) is the root of moral change. Indeed, "A *moral revolution* occurs if dissidents displace established moral paradigms with incompatible alternative paradigms." (Baker, 2019, p. 18)

**Axiological Futurism**: The systematic and explicit inquiry into the axiological possibility space for future humans (and post-humans), with a specific focus on (1) how different individuals or groups of people might shift or change location within that axiological possibility space, and (2) the future agreement and disagreement between different individuals and groups of people about different axiologies.

I believe that this formulation would refocus the project of Axiological Futurism in a way that recognizes the importance of potential axiological pitfalls and the role that disagreement may play in creating such pitfalls. In the above definition, I focus on disagreement both between individuals and amongst groups of people. While the latter is arguably closer to Danaher's original project, I think the former is also important since we know that individuals can sometimes be a driver of change (e.g., consider, again, the example of Rosa Parks).

## 3. Methodological implications for value theory

In this section, I will turn to a part of the practical purpose of the Axiological Futurism project and its implications. As I noted in the Introduction, Danaher proposes that we investigate how axiologies change. We might worry about such an investigation for several reasons. For example, some might worry about paternalism. However, the proposed research is not necessarily dedicated to the hierarchical teaching of axiologies (even if many parents would probably appreciate methods to teach their children basic values more efficiently). Indeed, understanding how axiologies change can allow us to create, design, or foster a state of affairs in a way that benefits individuals' axiology, or we might learn how people can self-improve their axiology.[4]

The worry I am concerned with here is that any understanding of how axiologies can be changed is a form of knowledge that can be abused and misused (i.e., people can be misled based on both bad and good, but erroneous intentions, respectively).[5] I take this to motivate a need to do *more* ethics, but beyond more ethics, the Axiological Futurism research agenda also implies a need for new approaches, methods, and questions in the domain of ethics. Currently, work in value theory (axiology) is focused on finding the correct axiology (as well as philosophical questions about values). If such an endeavour were successful, and if we were in a position to know that it had been successful, then we could arguably resolve my worries because we would be in a position to know what axiology everyone should adapt. However, most would agree that even if such results were possible, we could not expect such results within any reasonable timeframe.

Since we cannot expect to know what axiology we ought to accept, Axiological Futurism requires us to approach this part of value theory somewhat differently. While there is (great) uncertainty about what the correct axiology is (i.e., if there is any correct axiology), there is arguably less uncertainty about examples of incorrect axiologies. For example, an axiology that claims that the only intrinsically good is goat cheese is clearly incorrect. While such an example is not worrisome, because few, if any, humans would adopt such an axiology, there are arguably more pressing problems.[6]

Perhaps one of the more pressing problems is that many people today hold axiologies that are clearly wrong and harmful to others, for example, axiologies according to which a person's worth is determined in part by their gender, the colour of their skin, their culture, their religion, their position, or their capabilities, etcetera. We know that these axiologies are incorrect, but they are held by a large part of the global population.[7] If we were to pursue the Axiological Futurism agenda, then the focus of the work done in value theory would need to change. While value theory today is arguably focused on what is good and how we compare it, we would need more work to be done on how to explain why an axiology is incorrect. That would imply a change in focus relative to the object of study (from what the good is, to what is faulty about certain axiologies) and work that is less normative and more empirical (i.e., the focus on explanations).

Some might argue that the examples I have given are fairly easy to explain and hence that there is only a limited need for further investigations into understanding incorrect axiologies. However, there is arguably a related difficulty that also needs to be dealt with. If we investigate the possibility-space for axiologies and we know how people's axiologies can change, then we need to understand how to compare different axiologies. We need tools to understand what makes one axiology better than another and tools to allow some sort of quantification of the incorrectness of various theories of axiology. That is, we need to know whether a change in axiology implies progress or regress and which axiologies should be preferred and why. I believe this is in line with Danaher's prescriptive aim of Axiological Futurism, but the focus should not only be on the way forward; we must also understand the nature of varying degrees of

---

[4] I'm loosely paraphrasing the literature on nudging and boosting (see, e.g., Hertwig & Grüne-Yanoff, 2017; Thaler & Sunstein, 2021). A reviewer pointed out that we can question to what degree a person can self-improve their axiology. While there may be challenges in doing so, we know that there are already ways to affect one's axiology. Consider that people can change their axiology (e.g., a racist becoming an anti-racist or vice versa). While many such situations may be best qualified as situations in which change *happens to an individual* (rather than situations in which an individual *actively attempts to improve their axiology and succeeds*), it is clear that whatever pathway that leads to such change can be engaged with by anyone (e.g., be it reading, discussing, or thinking). One may ask if such engagement is likely; in response to that I would say that it is at least somewhat plausible for those who presume that their axiology is imperfect and seek improvement.

[5] This is not only a worry for Axiological Futurism, but for any kind of work that looks at how to influence people.

[6] Of course, the Axiological Futurism project potentially goes beyond human civilisations so it might merit concern in the long term. Moreover, some would worry about such axiologies being set as goals for a super intelligent AI (see, e.g. Bostrom, 2014). However, that is arguably a different kind of worry.

[7] For an example of people's preference for saving people's lives depending on properties such as their gender, age, or bodily size, see Awad et al. (2018).

*incorrectness* when it comes to different possible axiologies.

This warrants a further change in the definition:

> **Axiological Futurism**: The systematic and explicit inquiry into the axiological possibility space for future humans (and post-humans), with a specific focus on (1) how different individuals or groups of people might shift or change location within that axiological possibility space, (2) the future agreement and disagreement between different individuals and groups of people about different axiologies, and (3) how we evaluate or compare different axiologies.

In the above final definition of Axiological Futurism, I have added a third aim to the project (i.e., the focus on evaluation and comparison), which I have formulated in terms of an *evaluation or comparison of different axiologies*. This is an intentionally ambiguous formulation that ranges from *full evaluation and comparison* to *partial evaluation or comparison.* While it would be excellent if we could fully evaluate and compare any of the different axiologies (i.e., so that we could say for any axiologies x or y, whether x is better than y, y is better than x, or x and y are equally good), it would often be sufficient to make simpler evaluations and comparisons (e.g., an important focus should, as I say, be to qualify when an axiology is unacceptable). Indeed, given that the latter approach is more likely to reach fruition, it would arguably be better to focus on the latter than the former. Nevertheless, I think the definition of Axiological Futurism should include a broader set of activities granted that more ambitious agendas will hardly be detrimental to its aim since it will likely yield at least partial results.

In sum, I believe that the final definition of Axiological Futurism is better suited to satisfy even Danaher's original aims. As I have argued, disagreement on an individual and group-level, as well as the ability to evaluate or compare (partly, or fully), will be central to the endeavours of the project.

## Acknowledgement

## References

Awad, E., Dsouza, S., Kim, R., et al. (2018). The moral machine experiment. *Nature, 563*, 59–64. https://doi.org/10.1038/s41586-018-0637-6

Baker, R. (2019). *The structure of moral revolutions: Studies of changes in the morality of abortion, death, and the bioethics revolution*. Cambridge, MA: The MIT Press.

Bicchieri, C. (2017). *Norms in the wild: How to diagnose, measure, and change social norms*. New York, NY: Oxford University Press.

Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. New York, NY: Oxford University Press.

Danaher, J. (2021). Axiological futurism: The systematic study of the future of values, 102780: 1–14 *Futures, 132*. https://doi.org/10.1016/j.futures.2021.102780

Hertwig, R., & Grüne-Yanoff, T. (2017). Nudging and boosting: Steering or empowering good decisions. *Perspectives on Psychological Science, 12*(6), 973–986. https://doi.org/10.1177/1745691617702496

Nussbaum, M. (2000). *Women and human development: The capabilities approach*. Cambridge: Cambridge University Press.

Thaler, R. H., & Sunstein, C. R. (2021). *Nudge: The final edition (Updated edition)*. New York: Penguin Books.