Dialogue (2024), **63**, 467–474
doi:10.1017/S001221732400026X

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Special Issue: Canadian Philosophical Association 2024 Prize Winning Papers

This Paper Won a Congress Graduate Merit Award (CGMA) Prize at the 2024 Canadian Philosophical Association Conference

Conceptual Dynamite: An Argument for Thoughtful Language in Conceptual Engineering

Bell Luan

Department of Philosophy, York University, Toronto, ON, Canada
Email: bluan@yorku.ca

Abstract

Given the rising trend of politicizing formerly mundane topics, philosophers should guard against the concepts we develop being misunderstood or misconstrued to support arguments we never intended. To demonstrate my assertion, I develop the case study of Alex Kacelnik's notion of "biological rationality" (B-rationality) and show ways in which I believe it is vulnerable to misuse and ways I would suggest re-engineering it to be more resistant. I conclude by urging scholars — especially those working on normatively laden topics — to design our concepts such that they are less likely to become associated with harmful discourse or even oppressive policies.

Résumé

Compte tenu de la tendance croissante à politiser des sujets autrefois banals, nous les philosophes devrions nous garder de ce que les concepts que nous développons soient mal compris ou mal interprétés pour étayer des arguments que nous n'avions jamais envisagés. Pour soutenir cette assertion, je développe une étude de cas sur la notion de « rationalité biologique » d'Alex Kacelnik ; je montre en quoi je pense qu'elle est vulnérable aux abus et je suggère des façons de la repenser pour qu'elle soit plus résistante. Je conclus en exhortant les universitaires — tout particulièrement ceux qui travaillent sur des sujets normativement chargés, dont je fais partie — à concevoir leurs concepts de manière à ce qu'ils soient moins susceptibles d'être associés à des discours dommageables ou même à des politiques oppressives.

Keywords: meta-philosophy; conceptual engineering; applied ethics; social and political philosophy; feminist philosophy

1. Introduction

In today's climate in which science and scholarship so easily get politicized, we philosophers should all be more careful to consider the optics of our writings and safeguard, to some reasonable extent, against the prospect of their inadvertent or intentional misuse. In the worst-case scenario, our ideas can be misused to promote moral verdicts and political positions which we might not have foreseen and with which we might not agree. It is, of course, impossible to fully insure against this, but we owe it to the community (and, perhaps, to our own legacies) to do some work towards mitigating the risk.

In this conceptual engineering article, I explain some ways in which an innocent, apolitical concept might find itself perverted into political ammunition without the original author's agreement. I then provide three examples of real-world concepts that I consider to be vulnerable in this way. The first, and the one I develop in the most detail, is Alex Kacelnik's concept of biological rationality (B-rationality), which I believe could be distorted to *seemingly* ground an argument against reproductive autonomy, certain types of healthcare for people who are pregnant or might become pregnant, and trans rights. The second is the popular culture concept of the 'friend zone,' which has developed over time from a relatively innocent joke to a dangerous part of the broader cultural memeplex surrounding men's entitlement to women's bodies and attention. The third is Nietzche's concept of self-transcendence, commonly known as the 'super-man' idea, which famously found itself co-opted by Adolf Hitler's Nazis and is now culturally inseparable from its misuse in that context.

I develop three tiered levels of response to vulnerabilities like this and conclude by urging that we attend to the harm our concepts might do in the wrong hands.

2. Conceptual Engineering

I adopt the view described by David Chalmers: defining concepts is much less important than thinking about what they *ought* to be. As such, I advocate for conceptual engineering — which Chalmers argues should be thought of as "the design, implementation, and evaluation of concepts" (Chalmers, 2020, p. 16).

By including this normative 'ought' component, I believe that conceptual scholars — including, but not limited to, philosophers — can introduce more useful and less harmful ideas into society. Focusing exclusively on accurately describing concepts would omit these goals, which I think are important to keep in mind. I think we do in some way owe it to society to consider the social good when designing our concepts. Even if you disagree with that, we must at least owe it to our fellow humans not to participate in harming them by equipping bad actors with dangerous concepts. And even if you disagree with *that*, I trust that you care enough about your own legacy and reputation that you will not want to see your ideas bastardized and attached to movements with which you disagree.

Good societal work can be done just by refining concepts — even simple, everyday concepts. My favourite example of this — stemming from my days as a divorce lawyer — is the concept of a parent. We could have distinguished between, say, 'parents' with no adjective (meaning physical progenitors) and then step-parents, foster parents, etcetera. However, this would privilege progenitors by framing them as the only

https://doi.org/10.1017/S001221732400026X Published online by Cambridge University Press

pure, no-adjective parents. Alternatively, we could coin different nouns for every parenting role — but this would deny all but progenitors the valued title 'parent.' So instead, we assign an adjective even to progenitors — we call them 'biological' parents — to indicate through our thoughtful language that while all parents are real parents, no parents are 'realer' than the rest. In this way, our concept of parenthood avoids demeaning parents in 'non-traditional' family structures and so avoids doing harm by accidentally implying that their parenthood is abnormal or invalid.

Significant damage can result from the spread of a toxic or misleading concept through popular culture. For example, consider how the originally jokey, playful idea of the friend zone has turned into a conceptual monster over time. It began as a term for the phenomenon in which two people with shared, mutual romantic interest end up remaining friends because each thinks the attraction is unrequited; however, over time, certain communities have corrupted it into the idea that man-woman friendships are romantic by default unless the woman (it is thought) wrongs the man by sorting him into the 'friend zone,' which is seen by these communities as a grave and emasculating insult. While the cultural patterns associated with dating under patriarchy pre-date this concept's inception in the 1990s, the ever-expanding memeplex surrounding the friend zone has made things worse by "contribut[ing] to a set of attitudes and behaviors which ultimately feed into a culture of sexual violence against women" (Shields, 2017, p. 57).

Worse yet is when harmful concepts enter politics and governance. The most notorious example here is likely that of the *Übermensch* — an idea which famously found itself misinterpreted by Hitler and misused to underwrite atrocities. Here, Friedrich Nietzsche's argument for the idea of an enlightened, spiritual thinker came to be twisted into an endorsement of militarism, fascism, and genocide — in large part, perhaps, because of his penchant for easily sloganized terms like "master morality," the "will to power," and even the mere concept of a "super-man" (Golomb & Wistrich, 2001, p. 309). Note that Nietzsche's main misstep (in this context) was phrasing his ideas in ways that made them dangerously easy to either misunderstand in good faith or intentionally distort — a misstep that led to his cultural legacy being that of the "insane philosopher" behind Nazism and the Holocaust (Golomb & Wistrich, 2001, p. 308).

Poorly crafted concepts can contribute to social harms in many ways; the above examples show only three. A poorly crafted concept of parenthood might *demean certain populations*, namely parents other than biological progenitors. The friend zone idea has been used to *reinforce oppressive structures* in society. And Nietzsche's basically virtue-ethical ideas of self-transcendence were co-opted to *underwrite harmful policies*.

3. Case Study: B-Rationality

In a book titled *Rational Animals?*, Kacelnik contributes a chapter in which he taxonomizes three types of rationality:

- E-rationality, for the type of utility maximization discussed by economists;
- PP-rationality, for the type of reason-sensitive, belief-adopting process discussed by psychologists and philosophers; and

https://doi.org/10.1017/S001221732400026X Published online by Cambridge University Press

- B-rationality, under which an individual is rational if and only if their actions maximize their "inclusive fitness" — which in turn is a complex function roughly equivalent to the spread of the genetic alleles which that individual carries. (Kacelnik, 2006, p. 95)

Note that while personal reproductive success is one way to achieve inclusive fitness, it is not necessary. Rather, an organism is B-rational if it acts such that its genetic makeup makes up the highest possible proportion of the population; thus, an organism which does not itself reproduce can still be inclusively fit if its relatives flourish. For example, imagine a childless adult saving the life of their biological sibling or helping that sibling find a mate. This kind of act would, *prima facie*, score well under B-rationality because preserving/sustaining a sibling's life also helps maximize the spread of the genetic alleles one carries.

Kacelnik allows that evolutionary biologists do not tend to focus on rationality as such. He does not claim that they use the term. However, he himself uses B-rationality to argue for ways in which non-human organisms (his interest is in birds) can be understood as rational even if we deny them E-rationality or PP-rationality (though he argues that they might have these as well). In pursuit of this argument, he extends rationality to include inclusive fitness maximization. In that specific context, I am inclined to agree with Kacelnik: animals, mosses, and even bacteria might easily enough be called 'rational' *if what we mean by that is B-rational*.

4. Conceptual Dynamite

However, I argue that *biological rationality* as presented by Kacelnik is flawed in the same way as Nietzsche's idea of the *Übermensch*: being too easy to misunderstand or misconstrue in a specific dangerous manner, while not *needing* to be in order to serve its philosophical purpose. Worse, if that flaw were to be exploited, then B-rationality could contribute to each kind of harm discussed above: it could demean certain identities, reinforce patriarchy in popular culture, and lead to oppressive state policies.

The problem I see is that the idea of rationality, in every context but Kacelnik's taxonomy, is heavily charged with normative value. People use it to evaluate each other's decision-making, and rationality is almost always valued over irrationality. And while an educated, good-faith reader of Kacelnik will not mistake him for saying that all persons ought to maximize inclusive fitness in their individual paths through life, it would be easy for a non-scholastic reader or, worse, a bad actor to derive the following distortion:

According to Kacelnik, it is ***biologically irrational*** *to act contrary to evolution.*

In the wrong hands, this distortion would act like dynamite: a tool created for beneficial use in one context (mining) but transposed to another (warfare) where it can cause enormous harm. Combining the normatively charged concept of *rationality* with the scientific, neutral-seeming prefix 'biological' is dangerous. It could be used to argue (mistakenly, but perhaps convincingly nonetheless) that

https://doi.org/10.1017/S001221732400026X Published online by Cambridge University Press

any course of action that does not align with natural selection (a concept Kacelnik says is closely related to inclusive fitness and central to B-rationality) (Kacelnik, 2006, pp. 93–94) is fundamentally irrational and thus not a proper part of the human condition. Humans are, after all, commonly supposed to be rational animals.

Bad enough already, of course, for a concept to be potentially confusing. But in this particular case, I argue that the error (or distortion) above leads almost immediately to a dangerous inference: we should prevent people from people acting in ways that seem incompatible with evolutionary fitness. Consider the following examples:

- 'Voluntary' (i.e., not strictly necessary as a matter of survival) hysterectomies, which have increased in the United States following the reversal of *Roe v. Wade* (Gomez, 2002), clearly result in the recipient being unable to pass on their genes;
- Gender-affirming care for transgender and non-binary patients (both surgical and pharmaceutical) can make reproduction difficult or impossible, regardless of the direction of transition; and
- Needed medications that can incidentally compromise reproductive function are often denied to people of child-bearing age. (Plank, 2022)

On a distorted reading of Kacelnik, the interventions discussed above might become the centre of a culture-war battle surrounding whether procedures that harm a person's reproductive capacity are fundamentally — and again, this is not a term Kacelnik himself uses — 'biologically irrational.' This would, I hasten to emphasize, be a distortion: Kacelnik in the original does not fixate on personal reproductive success, but rather on inclusive fitness. But that distinction is too fine to disarm the conceptual dynamite; it would be trivial to either miss, misunderstand, or maliciously omit.

Kacelnik might be expected to respond that his target audience would do none of those things — and I am confident that this would be true! No competent, good-faith scholar reading Kacelnik would leap to the conclusion that he is himself calling any of the medical interventions discussed above inappropriate on supposed B-rationality grounds.

However, my concern is different. I am worried about good-faith lay readers unfamiliar with the fact that scholars sometimes redefine terms like 'rationality' in narrow and counter-intuitive ways, and I am even more worried about bad-faith readers of any description looking for a Trojan horse in which to hide their agendas.

We already see extensive claims that interventions like birth control, abortion, hysterectomy, and gender transition are 'unnatural' or 'against biology.' B-rationality (and its implied opposite, B-irrationality) seems so similar to these ideas on a superficial level that even a well-meaning reader might conflate them; a hostile one could easily produce a misleading description of Kacelnik's concept and weaponize it in bad faith to justify state restriction of various human rights. In a society that already tends to excessively politicize scientific issues like climate change and vaccine efficacy (Talisse, 2019), and in which oppression along intersecting social lines remains rampant, these dangers are worsened.

https://doi.org/10.1017/S001221732400026X Published online by Cambridge University Press

5. Levels of Response

For the reasons above, I suggest that we consider three levels of response to this type of risk, both here and generally throughout the practice of philosophy. The greater the danger of a particular concept being co-opted and misused, the stronger a response is warranted.

1. **Focus on Precision:** Where possible, describe concepts as precisely as can be done without sacrificing needed generality. For example, in this case, we might re-describe B-rationality without the need to *maximize* inclusive fitness (which Kacelnik's deployment of the concept does not actually require on my reading).

   Thus, an organism would be B-rational *to the extent that* its actions are *consistent with sufficient* inclusive fitness, rather than binarily either B-rational or not B-rational. This narrower, more focused account would still (I argue) accomplish Kacelnik's goal of highlighting the ways in which non-human organisms or their behaviours might be rational while mitigating the risk of misuse.
2. **Avoid Politicized Terms:** Where possible, give concepts technical names that do not resonate with unrelated political concepts. In this case, we might recognize that combining 'biological' and 'rationality' mirrors the discourse around reproductive freedom, and consider a replacement name like 'ecological rationality' or, when *not* arguing about whether plants and animals are rational as such, something like 'allele-distributive efficiency.'

   Of course, an objector might argue that this would lead to a euphemism arms race. As time and society move forward, it may be necessary to once again neutralize our terminology to keep it from resembling unrelated political topics. Perhaps ecology will be politicized in future, or non-biologists will start to talk a lot more about alleles.

   But this is not a new or unique problem. For generations, we have had to neutralize and re-neutralize terminology as language and discourse evolves. We renamed seats and parking spots 'accessible' instead of 'handicapped,' we began using person-first language ('person of colour' instead of 'coloured person'), and we are pushing for the use of 'Settlers and First Nations' as a more ethical replacement for 'Canadians and Aboriginals.' And many concepts have had to be repeatedly reworked; consider how many terms for people with mental disabilities have become slurs over time and have had to be replaced with new language.

   Scholastic terms of art should not expect to be immune to these pressures, especially if harm might result from continuing to use them in their current forms. We do not actually live and work in hermetically sealed ivory towers; we can help and harm the world around us. We, in a word, matter.

   Our industry, like every other industry, owes a duty to society — especially marginalized members of society — to at least try to mitigate the harm we might inadvertently do in the course of our work.
3. **Emergency Pause:** In extreme cases, I think it can be best overall to simply stop discussing concepts that are too dangerous to continue using. As fascism continues to resurge throughout the world, it might be the wrong time to talk about benevolent dictators, Plato's philosopher-kings, or Nietzschean

https://doi.org/10.1017/S001221732400026X Published online by Cambridge University Press

Übermenschen. Any philosopher who has advocated for these concepts, even in hypothetical or counter-factual worlds, might turn on the radio one morning to find that their work has been co-opted and *they* were now the "insane philosopher" being misquoted by some bad actor.

Of course, this level of response is the biggest ask. I, myself, must now do some defensive engineering by specifying that I do not mean that we should stop talking about political concepts! Rather, I mean to say only that there can in rare situations arise a set of circumstances in which it would be reckless to publish certain ideas. For example, the marketplace of ideas certainly needs a discussion of whether conquest could ever be justified — but if it was the day after Russia commenced its attack on Ukraine, it might be prudent to pause that discussion on an emergency basis lest one come to be seen as sincerely arguing that Vladimir Putin is doing nothing wrong.

By the way, in an earlier version of this article, this section was titled 'Avoid Harmful Concepts.' In an attempt to practice what I preach, I have renamed it to emphasize that this level of response should be used only as a last resort and even then only temporarily. It would have been too easy, in my opinion, to distort the earlier heading and make me seem to argue for something I don't believe!

In most cases, the risk of politicized co-opting is low and we can rely on the lowest level of response. After all, precision is already a desideratum of philosophical writing. But, in more charged cases, the higher levels might be necessary too.

For example, if asked, I would suggest Levels 1 and 2 for B-rationality. I think a re-engineered description that allows B-rationality to be scalar rather than binary coupled with a less inflammatory title would suffice to allay my concerns. I see no need to pause discussion of the concept, especially since the case for animal rationality is also important and B-rationality is important to that cause.

6. Conclusion

For all of the above reasons, I believe that it is important that scholars — especially those working in or adjacent to normatively laden topics — be aware of the risks associated with the increasing politicization of society and discourse. We ought, therefore, to design our concepts such that they are resistant to misuse, whether accidental or deliberate, and thus less likely to become associated with harmful discourse or even oppressive policies. While I have focused this article on the example of Kacelnik's B-rationality, I believe the same problem applies to many other concepts currently being discussed in philosophy and other scholarship. I urge us all to consider whether our concepts are susceptible to misuse and, if so, what level of response is needed to sufficiently mitigate that risk.

Competing interests. The author declares none.

References

Chalmers, D. J. (2020). What is conceptual engineering and what should it be? *Inquiry*, 1–18. https://doi.org/10.1080/0020174x.2020.1817141

Golomb, J., & Wistrich, R. S. (2001). Nietzsche's politics, fascism and the Jews. *Nietzsche-Studien*, *30*(1), 305–321. https://doi.org/10.1515/9783110172409.305

Gomez, M. (2022, August 25). More women are seeking sterilization since Roe vs. Wade was overturned. *Los Angeles Times*. https://www.latimes.com/world-nation/story/2022-07-19/women-sterilization-roe-vs-wade

Kacelnik, A. (2006). Meanings of rationality. In S. Hurley & M. Nudds (Eds.), *Rational Animals?* (pp. 87–106). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198528272.003.0002

Plank, L. (2022, July 11). Abortion bans are stopping these women from getting medication for their chronic illness: Women of childbearing age are being told that having a uterus precludes them from having access to vital medication. *MSNBC.com*. https://www.msnbc.com/opinion/msnbc-opinion/post-roe-abortions-aren-t-only-healthcare-beingdenied-women-n1296928

Shields, G. L. (2017). "A place where every decent guy will find himself eventually": Delineating the friend zone as a site of sexual violence. Doctoral dissertation, University of Texas at Austin. https://doi.org/10.15781/t2rf5m073

Talisse, R. B. (2019). *Overdoing democracy: Why we must put politics in its place*. Oxford University Press. https://global.oup.com/academic/product/overdoing-democracy-9780190924195?cc=ca&lang=en&

Cite this article: Luan, B. (2024). Conceptual Dynamite: An Argument for Thoughtful Language in Conceptual Engineering. *Dialogue* 63(3), 467–474. https://doi.org/10.1017/S001221732400026X