

On Future Worlds

Frank P. DeVita

NYU Center for Bioethics

March 2026

In this paper I'll offer a way out of the non-identity problem through reasoning with possible worlds. I'll lay out the concept of non-identity and show its incompatibility with person-affecting moral logic to motivate an impersonal view of morality and action grounded in counterfactual reasoning. I want to jettison the confusing concept of a *possible person* and ground moral differences between actions in abstract comparisons of the goodness and badness of the different possible worlds those actions might realize. I'll develop the idea that since the relative ranking of possibilities plays an intuitive role in moral reasoning, abstract projections of possible worlds put what I call *counterfactual moral pressure* on agents to actualize better worlds, independently of whom or what might be in them. The upshot to this worlds-centric view is that it permits separation from anthropocentric moral logic and centers moral reasoning on actions that might actualize a better future world with the best possible conditions for flourishing, no matter the nature of individuals in it.

Navigating Non-Identity

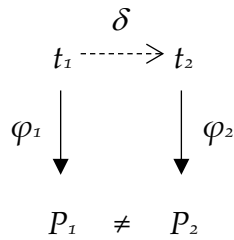
It is intuitive to think that *possible people* and *future worlds* weigh on our moral conscience. We must not use up the planet's resources, pollute the environment, or stop fighting infectious disease because future generations would be worse off if we did. The environment would be adversely affected, and with it, countless creatures, life forms, and constituents of the natural world. To complicate matters, our actions cannot harm or benefit possible people because non-identity nullifies person-affecting logic: since our choices determine *who* exists, those individuals will be no worse off existing than they would have been otherwise. This is the philosophically vexing *non-identity problem*. Derek Parfit thinks *future people*, those who *will* exist whichever way we may act can be harmed "though they don't live *now*, just as we can harm foreigners that don't live *here*."¹ But what about *possible people*—those who would exist if we acted in one way but not another?² Since our actions determine *who* exists, they create individuals *non-identical* to any others who would have existed had we acted differently.³ All non-identity contexts, I think, share a common structure:

¹Parfit 1976, p. 369.

²Parfit 1976, p. 369-70.

³Harman 2004, p. 89.

On Future Worlds



Discrete times t are represented on the dotted line. The arrows are actions φ that create non-identical (\neq) people or populations P . Adopting a policy, making a choice, or taking any action δ that shifts conception from t_1 to t_2 determines that a non-identical person or population P_2 will exist where P_1 would have. If our actions determine *who* exists, then no action can affect or *reach* a possible person P_1 because they determine that individual's identity as P_2 . A problem arises because non-identity is incompatible with moral theories that take harm as *person-affecting*. That is, constituted by actions that decrease *someone or something's* quality of life. Person-affecting harm is frictionless with respect to possible people and non-identity because harm cannot affect non-existent people. Therefore, we cannot coherently assert that possible people are affected by our actions. Non-identity *breaks* person-affecting logic.

If a possible person cannot be harmed, then an act "toward" them is neither a coherent thing nor a phrase nor analyzable in person-affecting terms⁴ because it has no object, and as such, is morally inert. Yet possible people, understood as children we

⁴ Parfit 1973, p. 372

might have or concretely as *future generations*, could be better or worse off because of our actions and decisions. If we conceive in poor health or in a warzone, ignore preimplantation genetic diagnoses, dump radioactive waste, or institute inflationary economic policies, or other φ s (actions), these can all have bad effects *on people*, so we shouldn't φ . If person-affecting moral relations to possible people are metaphysically and logically impossible because of non-identity, then what *grounds* the moral relation to future generations? Non-human beings? The environment? The *world*?

Projecting the Future

If we pollute the air until the sky becomes a toxic yellow haze and no one could comfortably breathe outside, that would be bad for the quality of life of living and future people. It would be bad for living and future animals, plants, and ecosystems. It would be bad for *the world*. If someone were to conceive while taking a teratogenic medication or refuse to stop using drugs while attempting to conceive, that would be bad for whoever is conceived. Dumping radioactive waste into a river would be bad for the futures of everything that is and will be in the surrounding area. We ought to adopt policies and take courses of action that avoid these kinds of morally suspect, destructive situations and behaviors. What kind of logic can ground this moral intuition? One answer is to deem future people *objects or productive ends of the will* with full moral

status.⁵ But this moral circle doesn't snare the environment. Another is to stipulate weighty reasons against actions that could harm future beings.⁶ A third is a duty to prevent and relieve suffering.⁷ My answer is that moral relations are grounded in how our actions will affect the conditions of the world, and the implications of those conditions on beings and things in that world.

If we take *welfare* or *quality of life* (henceforth used interchangeably) to mean the composite measure of health and constitutional factors that constitute how a life is going,⁸ the welfare and quality of life of possible people is a confused metric for reasons discussed above. So, to get moral obligations to future beings off the ground and overtake non-identity, we need a non-person-affecting solution. I'll suggest a modal one that expands the moral circle: If we posit a possible world in which possible people and living things *just are* future people and living things, we can abstractly project the welfare of beings in that world relative to an action, make cross-world comparisons of worlds with different actions, and reason about what we ought to do in the present using counterfactuals. These worldly welfare quotients, in my view, exert *moral pressure* on agents to select courses of action that would *create better worlds* with conditions for

⁵ Chambers 2025, p. 8-11.

⁶ Harman 2004, p. 91-93.

⁷ Narveson 1967, pp. 71

⁸ World Health Organization 1998/2012, p. 11. This definition specifies to people, but I want to leave a conception of "health and constitutional factors" open for application to non-human beings.

welfare and quality of life for future beings, whoever—or whatever—they are. On this model, moral determinations are cast in virtue of the conditions of the world they bring about and how those conditions vary with welfare in that context.

I have in mind a notion of a possible world as a heuristic device with explanatory force shaped by the facts and norms of the actual world⁹ or a counterfactual situation.¹⁰ Though abstract, these worlds are inferred from and therefore track reality yet allow for fixing and adjustment of morally relevant variables—in this case actions and outcomes—to make moral determinations. We can, for example, fix an action and ask about its effects on different people or beings, or fix outcomes and ask what actions might bring them about. Using possible worlds means dealing in counterfactuals, which I take as subjunctive conditionals that, in this context, relate agents, actions, future moral patients, and constitutional states¹¹. As such, any analysis of any action brought about by any agent is through what the world might be like for moral patients in a world with conditions determined by that action.

On this picture, moral obligations to future generations are *made true*¹² by counterfactuals and how we think the world would be given some action. These counterfactuals are grounded in conditions of the actual world and imagine them to

⁹ Stalnaker 2003, pp. 25-26.

¹⁰ Kripke 1980, p. 15.

¹¹ Harman 2004, p. 96-97

¹² Fine 1997, p. 556

change or be altered in various ways. If those alterations might make it more difficult for individuals that might exist to flourish, there is reason to avoid actions that bring those alterations about. Projecting worlds allows for the evaluation of future states of the world and so surmounts non-identity's block on person-affecting logic by focusing on the conditions we might bring about with our actions, and determines the morality of those actions in terms of the worlds that agents might *actualize*. The responsibility to actualize flourishing-promoting worlds, on this view, is a *pro tanto* reason for agents to be conscious of their actions and their possible effects on the future.

Cases

Let's run some classical cases¹³ through the projective worlds analyses and compare moral determinations for: (a) a 14-year-old girl who conceives, (b) a couple remains childless, (c) a woman who hastily conceives despite a high risk of easily avoidable birth defects, and (d) an expecting mother who refuses a simple treatment that will prevent fetal harm. Person-affecting logic doesn't reach (a) though (c) because there are only possible people in play in a non-identity context and harming can't get traction. In (d), harm reaches a future person and so can ground a moral judgment.

¹³ These cases appear in Parfit 1976 pp. 373-375, and Parfit 1984 pp. 357-361.

Using projective worlds (a), (c), and (d) are morally suspect actions because there are near worlds where welfare and quality of life would be higher at low cost to the agents. Taking the possible people in these situations as future people, we get the following counterfactuals: at w_a^* if the 14-year-old girl waits to conceive, her child would have a more mature and capable parent. At w_c^* if the hasty conceiver delays pregnancy, her child wouldn't suffer birth defects. At w_d^* if the expecting mother takes the treatment, gestating baby wouldn't be harmed. These worlds generate impermissibility results from the counterfactuals describing them. If the agents were actively deliberating, the counterfactuals would apply moral pressure on them to make different choices and so can guide action. Should the agents act in ways from the projected worlds, they would create conditions for higher welfare and quality of life. If they don't, they are answerable for why they made alternative, evaluably bad decisions despite high moral pressure from relevant counterfactuals. By contrast, in (b) there is very little moral pressure on the childless couple to procreate because although it would actualize a world where *more* welfare and quality of life are possible, the action would be supererogatory. The childless couple, on the projective view, might be judged to have selfish preferences, but those judgements would be unwarranted.

On my view, what makes our moral assertions about these agents and their actions true—what grounds them—are abstract comparisons of the conditions they create across possible worlds evaluated as a function of an agent's possible courses of

action and relevant counterfactuals. It implies that *if* there are courses of action by which *better worlds* with conditions that promote flourishing for whom or whatever exists in them are reasonably actualizable, that agents *ought* to select those courses of action over others for the sake of the future world. We can imagine other cases involving populations, animals, the environment, and non-human minds undergoing a similar form of analysis.

Conclusion: Affirming Abstraction

One might object that the projective worlds strategy does not solve non-identity cases nor ground morality because it estimates welfare and quality of life of future people in the abstract without identifying a moral victim or explaining why they would be harmed by an action. This objection turns on rejecting projection of welfare and quality of life as satisfying the explanatory demands of moral permissibility and denying that non-rigidly designated¹⁴ individuals can be terms in moral relations.

If our moral goals as human beings are to not merely live but to *flourish* and to promote, support, or at least not impede the flourishing of other living and conscious beings, present and future, then we have a moral obligation to maintain and *create*

¹⁴ Kripke 1980, p. 3-15

conditions for the possibility of more than a minimally decent existence. Without those conditions, the basic possibilities for flourishing might be constrained or absent, ruling out its possibility entirely. Such a world would be worse for who or whatever might be in it, and so is a worse world independently of how particular individuals' lives might go. This way, morality tracks situations and conditions of the world and compels agents to create situations in which flourishing is possible, independently of any individuals. We often think this way when we morally judge general conditions—pandemics, wars, heat waves, and other *situations* are bad in themselves beyond their badness for individuals. These situations are ways the world might be, and *projecting* worlds helps reveal ways in which we can create conditions for flourishing by selecting good actions that sustain or enhance conditions for flourishing, and pressures us to make those worlds actual.

To the second point, welfare and quality of life projections can be made independently of rigidly designated victims. Conditions of the world can make a victim out of anyone or anything. Who or what they are is in a sense unimportant relative to the demand to prevent unnecessary harm. The projective worlds strategy generates sufficiently morally weighty reasons for acting one way or another, who or whatever might be on the other side of action. It uses modality to pressure agents to *actualize* a better world, and implies that we are responsible for how our choices and actions shape the future.

References

- Chambers, K. Lindsey (2025) "A Non-Solution to the Non-Identity Problem" *Ethical Theory and Moral Practice*.
- Fine, Kit (2017) "Truthmaker Semantics" In *A Companion to the Philosophy of Language*, B. Hale, C. Wright & A. Miller, Eds. pp. 556–577.
- Harman, Elizabeth. (2004) Can We Harm and Benefit in Creating? *Philosophical Perspectives* 18, p. 89-113.
- Kripke, Saul A. (1980) *Naming and Necessity*. Harvard University Press.
- Narveson, J. (1967) "Utilitarianism and New Generations" *Mind*, 76(301), 62–72.
- Parfit, Derek (1976) "Rights, Interests, and Possible People" In *Moral Problems in Medicine*, S. Gorovitz, Ed. pp. 369-375.
- Parfit, Derek (1984) *Reasons and Persons*. Oxford University Press.
- Roberts, M. A. (2024) "The Nonidentity Problem" *The Stanford Encyclopedia of Philosophy* (Fall 2024 Edition). E. N. Zalta & U. Nodelman, Eds.
- Stalnaker, Robert C. (2003) "Possible Worlds (1976/1984)". In *Ways a World Might Be: Metaphysical and Anti-Metaphysical Essays*. Oxford University Press.
- World Health Organization (1998, rev. 2012). *Division of Mental Health and Prevention of Substance Abuse: Programme on Mental Health WHOQOL User Manual*.