Closing the Gap

Effective Altruism joined with the Vacuum.Net Theory

J. Konstapel · Constable Research, Leiden, The Netherlands

Version 1.0 · 29 November 2026

Findings

Effective altruism chooses where money goes. It ranks problems by scale, tractability and neglectedness. It prices uncertain outcomes with probabilities. It measures results. It does not ask how a problem works.

The Vacuum.Net theory asks exactly that. It describes every state as a loop that closes or does not close. Where closure fails, the theory computes where to intervene.

This essay reports four findings.

1. Effective altruism optimises without a model of cause. Its own methods show this. Its probabilities measure what is not known.
2. In the number system of the theory every orbit ends after finitely many steps. It ends in closure (0) or in one of two turning points ±(1+ω). This is proved for every point.
3. Every state that does not close can be made to close by one change of one trit. The depth of that change is computable. We call this the bending theorem. It is proved and checked on 35,064 states.
4. The two can be joined. The theory supplies understanding and the intervention. Effective altruism supplies measurement, correction and money. The join follows the theory's own rules for joining closures.

One thing stays open. There is not yet an instrument that reads a concrete problem as a state of the net. That instrument is the next piece of work.

Occasion

On 28 November 2026 Het Financieele Dagblad ran an analysis by Roel Verrycken. It described effective altruism as the ideology behind the AI company Anthropic. The company had just gone public. Its founders and early funders have close ties to the movement. Much of their new wealth is pledged to effective causes.

That raises a plain question. If a fund of this size sets out to solve the world's problems, which method does it use? And is there a better one? This essay answers the first question from the movement's own sources. It answers the second with the Vacuum.Net theory.

One detail of the article needs correction. It suggests that Sam Bankman-Fried would have earned over $13 billion from the listing. The FTX estate sold its stake in Anthropic in 2024 for about $1.3 billion in total. FTX held no stake at the time of the listing.

How effective altruism works

A rule of calculation

Effective altruism ranks a problem on three axes. Scale asks how many beings are affected and how badly. Tractability asks what share of the problem extra effort would solve. Neglectedness asks how little is spent on it now. 80,000 Hours defines tractability as the fraction of the remaining problem solved if direct effort doubled. That is an empirical rate. It is not a question of whether a solution exists.

Charity evaluators measure outcomes. GiveWell counts disability-adjusted life years and consumption. Effects that cannot be measured drop out of the sum. Expected value combines the outcomes, each weighted by its probability.

A lever: longtermism

Greaves and MacAskill argue for strong longtermism. Effects on the far future are the most important feature of our actions. They estimate the expected number of future people at 10^24 or more. With that number, a tiny cut in extinction risk outweighs any present good. This is how the rule tilts towards risks from artificial intelligence. Critics note that the case is robust only if one accepts decisions driven by minute probabilities of astronomical payoffs.

A funnel and a pipeline

The movement describes its own local groups as a funnel. People enter at a first introduction. A few leave at the bottom as deeply engaged people in high-impact careers. Introductory fellowships and paid student organisers feed the top. Career advice from 80,000 Hours steers graduates towards AI safety and AI policy.

Concentrated money

A small number of donors funds most of the movement. Estimates on the movement's own forum put the share of one foundation, Good Ventures, at one half to two thirds of all funding. The listing of AI companies may multiply this money.

The core fact

Each of these parts chooses between outcomes. None of them models how a problem arises. Effective altruism optimises without understanding. This is not a criticism. It is a description that follows from the movement's own definitions.

A second fact follows. Effective altruism decides with probabilities. In a deterministic system a probability is not a property of the world. It measures what the decider does not know. Effective altruism thus uses its own ignorance as a basis for choice.

Where effective altruism comes from

The Atlas of Wisdom separates two kinds of link. Transmission means shown contact or explicit reference. Echo means formal resemblance without shown contact. The same distinction sorts the roots of effective altruism.

From	To	Kind
Bentham (1789)	Mill (1863)	Transmission
Mill	Sidgwick (1874)	Transmission
Sidgwick	Singer (1972)	Transmission
Singer, Unger (1996)	Ord, Giving What We Can (2009)	Transmission
World Bank, Disease Control Priorities (1993)	GiveWell (2007)	Transmission
J-PAL, randomised trials (2003)	GiveWell, GiveDirectly	Transmission
Mozi (5th c. BCE)	Utilitarianism, effective altruism	Echo
Śāntideva (8th c.)	Utilitarianism	Echo
John Wesley (1760)	Earning to give	Echo
Scientific charity (from 1869)	GiveWell	Echo

From	To	Kind
Copenhagen Consensus (2004)	GiveWell	Echo

Three methods resemble effective altruism closely. The World Bank's Disease Control Priorities ranks health interventions by cost per disability-adjusted life year. The Copenhagen Consensus ranks global problems by cost-benefit analysis. Venture philanthropy and impact investing apply venture-capital discipline to social goals.

All of them, and effective altruism itself, choose by outcome per unit of money. None asks how the problem works. That is the gap this essay closes.

The theory in brief

The Vacuum.Net theory starts from one definition. A distinction is a movement that returns to itself. There is one strand. Every distinction is a loop of it. A loop persists only if it closes. What does not close passes on as tension. Coarse blocks of the strand obey the same rules again.

Each step of the strand is a trit: forward (+1), standing (0) or backward (−1). In the plane the trits form a number system on the triangular lattice of Eisenstein integers $z = a + b\omega$, with ω a cube root of one. The base is $\beta = 1 - \omega$. One step of the system is:

$$T(z) = (z - d) / \beta, \quad \text{with } d \in \{-1, 0, +1\} \text{ and } d \equiv a + b \pmod 3.$$

Every orbit ends

Theorem 1. Every orbit of T ends in 0 or in one of the two turning points $+(1+\omega)$ and $-(1+\omega)$.

Proof. The step satisfies $|T(z)| = |z - d|/\sqrt{3}$. So $|T(z)| \le (|z| + 1)/\sqrt{3}$. This is smaller than $|z|$ whenever $|z| > (1 + \sqrt{3})/2$. Norms are whole numbers. Hence the norm $N(z) = a^2 - ab + b^2$ drops strictly at every step while $N \ge 3$. Every orbit therefore enters the seven points with $N \le 1$. A direct check of these seven points shows exactly three fixed points: 0, $+(1+\omega)$ and $-(1+\omega)$. The other four reach 0 within two steps. ∎

The number of steps S satisfies $\log_3 N - 1.57 \le S \le 2 + \lceil \log_3((4 + 2\sqrt{3})N) \rceil$. About 70 % of points close. About 30 % end in a turning point; that share oscillates and has no limit.

Both the theorem and the bounds were tested under the Respondent–Judge protocol. Two independent judges confirmed them with their own code. One judge checked all 1,088,347 lattice points with norm up to 300,000 by building the three orbit trees backwards. The result is an easy case of known shift radix system methods. Its use here is new.

The bending theorem

Theorem 1 says what happens if nothing is done. The bending theorem says what one intervention can do.

Lemma. Changing one trit at depth k means adding $e \cdot \beta^k$ to z, with $e = \pm 1$. Since $\beta^k \cdot e \equiv 0 \pmod \beta$, the digit does not change. So $T(z + e\beta^k) = T(z) + e\beta^{(k-1)}$. After k steps the orbit stands at $T^k(z) + e$. An intervention at depth k is thus the addition of ±1 to the state reached after k steps.

Theorem 2 (bending). Every state that does not close is made to close by one trit change. The depth is the step at which the orbit reaches its turning point.

Proof. Let the orbit reach $+(1+\omega)$ at step k^*. Change one trit by -1 at that depth. The state becomes $(1+\omega) - 1 = \omega$. Then $\omega \to -1 \to 0$. The loop closes. For $-(1+\omega)$ change by $+1$. The state becomes $-\omega$. Then $-\omega \to 1 \to 0$. ∎

A computation confirms it. All 35,064 non-closing states with $|a|, |b| \le 150$ close after one change, at depths 0 to 9.

Depth of the change	States closed
0	6,248
1	4,660
2	4,042
3	3,856
4	3,834
5	4,010
6	4,416
7	3,418
8	578
9	2
Total	35,064

Corollary. Any state can be bent into any other. The difference $m + n\omega$ equals $(m + n)\cdot 1 - n\cdot\beta$. That is a finite number of trit changes at depths 0 and 1. ∎

The consequence is direct. Within the theory, a failed closure does not call for a bet. It calls for a computation. Where to intervene and at which depth is fixed by the orbit.

The fusion

Joining two different things is itself a closure problem. One Distinction gives rules for it. This essay uses them to join effective altruism and the Vacuum.Net theory.

What each brings

Effective altruism brings four strengths. It sends resources where they do most good. It measures whether something works and corrects course. It moves money and people at scale. It is willing to change its mind.

The theory brings three. It explains how a problem works: the cause, not the symptom. It uses one language for every field: body, organisation, society. It fixes the outcome and computes the intervention.

How they join

- **Rule 5.** Only rings climb to the next depth. The theory climbs through all depths. It acts as a ring.
- **Rule 6.** Join a ring to a meeting, not to standing. Effective altruism is a meeting of people, money and organisations. Its fixed scoring models are standing. So the theory joins the people and the money, not the scoring models. The cost is one.
- **Rule 2.** Two different closures join only at a cost. The seam carries tension. Here the seam is probability against fixed outcome. It must be named, not hidden.

- **Rule 11.** Do not ask a closure to read its own content. Let a shallower closure read it. The theory must not judge its own interventions. Measurement in the world must do so. The theory itself asks for what effective altruism does best.
- **Rule 8.** Give every reading its depth. Measures such as life years gain a depth: on which winding does the effect sit?
- **Rule 9.** Read the shallow only by agreement. At depth m, (3^m + 1)/2 equal steps are needed. This gives evidence and replication a fixed threshold per depth.

Money, debt and the future

In the theory, money is an address of exchange, not a standard outside it. Money therefore goes to the address of the failed closure. It is not counted in euros per life saved.

Debt is open tension carried to a later depth. What is not closed now reaches later generations as tension. Longtermism changes accordingly. It no longer bets on small chances of vast futures. It closes now.

The joined method

Step	Action	Source
1 Understand	Read where closure fails and give the reading its depth	Theory (Rule 8)
2 Intervene	Compute the trit change that restores closure	Theory (Theorem 2)
3 Measure	Let the world judge whether closure is restored	Effective altruism (Rule 11)
4 Correct	If not, the reading was wrong; read again	Both
5 Fund	Send money to the address of the failed closure	Effective altruism (§13.1)
6 Name the seam	State where probability and fixed outcome meet	Rule 2

What is dropped? From effective altruism: choosing by probability and expected value. From the theory: nothing. Effective altruism supplies the measurement the theory lacks.

What stays open

Theorems 1 and 2 hold inside the theory. To act on a real problem one must read that problem as a state of the net. One must know its depth and its orbit. No instrument for that reading exists yet.

This is an engineering task, not a gap in the proofs. The joined method already contains the test for it. Steps 3 and 4 check every reading against the world. A wrong reading shows up as a closure that did not happen.

Theorem 2 has been proved and computed. It has not yet passed the full Respondent–Judge protocol. That is the next test.

Annotated references

Akiyama, S. and Scheicher, K. Symmetric shift radix systems and finite expansions. Preprint, University of Tsukuba.

Why read? Shows that the base 1 – ω is not canonical, so turning points must exist. The known frame for Theorem 1.

Gabriel, I. (2017). Effective altruism and its critics. Journal of Applied Philosophy 34(4). doi:10.1111/japp.12176.
Why read? The standard philosophical review of objections: justice, equality, rights. Useful to see what the movement's critics already said.

GiveWell (2011). Why we can't take expected value estimates literally (even when they're unbiased). blog.givewell.org.
Why read? The evaluator's own warning about expected value. Shows the limits of the probability rule from inside.

Gleiberman, M. (2023). Effective altruism and the strategic ambiguity of 'doing good'. IOB Discussion Paper 2023.01, University of Antwerp.
Why read? An anthropologist's account of the public face and the core of the movement, and of its funnel model.

Greaves, H. and MacAskill, W. (2021). The case for strong longtermism. GPI Working Paper 5-2021, Global Priorities Institute, Oxford.
Why read? The core argument for longtermism. Read it to see how expected value tilts the movement towards existential risk.

Jamison, D. T. et al. (eds.) (2017). Disease Control Priorities, 3rd edition, vol. 9. World Bank. NCBI NBK525278.
Why read? The source of cost-per-DALY ranking that effective altruism adopted. Read chapter 1 for its history since 1993.

Konstapel, J. (2026). One Distinction, version 1.6. Constable Research, Leiden.
Why read? The definition, the closure rules and the rules for joining closures used in the fusion (chapters 13 and 19).

Konstapel, J. (2026). The Trit Tile: a reference work, version 1.1. Constable Research, Leiden.
Why read? The number system on the Eisenstein lattice. Source of the step T and of Theorem 1.

Leão, L. de S. and Eyal, G. (2019). The rise of randomized controlled trials in international development in historical perspective. Theory and Society 48.
Why read? How randomised trials became the gold standard that effective altruism relies on.

MacAskill, W. (2015). Doing Good Better. Gotham Books, New York.
Why read? The movement's programme in plain language. Read for its own account of scale, tractability and neglectedness.

Martin, O., Odlyzko, A. M. and Wolfram, S. (1984). Algebraic properties of cellular automata. Communications in Mathematical Physics 93, 219–258.
Why read? The algebra of additive cellular automata behind the passing rule of the theory.

Rubenstein, J. C. (2016). The lessons of effective altruism. Ethics & International Affairs 30(4).
Why read? A review of Singer and MacAskill that names the movement's 'hidden curriculum'.

Singer, P. (1972). Famine, affluence, and morality. Philosophy & Public Affairs 1(3), 229–243.
Why read? The founding argument. Short and clear. Read it first.

Stanford Encyclopedia of Philosophy. The history of utilitarianism. plato.stanford.edu.
Why read? Traces the line from Mozi and Śāntideva through Bentham, Mill and Sidgwick. Background for the lineage table.

Unger, P. (1996). Living High and Letting Die. Oxford University Press.

Why read? Extends Singer's argument. Cited as an inspiration for Giving What We Can.

Verrycken, R. (2026). Effectief altruïsme, de ideologie die Anthropic stuurt. Het Financieele Dagblad / De Tijd, 28 November 2026.
Why read? The occasion of this essay. A clear journalistic account of the movement's influence.

80,000 Hours. A framework for comparing global problems in terms of expected impact. 80000hours.org/articles/problem-framework.
Why read? The movement's own definitions of scale, tractability and neglectedness.

© 2026 J. Konstapel, Constable Research, Leiden. Reproduction of the whole requires written permission.