§ THE DIALECTIC REVIEW

Volume IV · Winter 2025/26

ISSUE 07

The Unequal Bargain

Winter 2025/26 · 13 pieces

Nine essays and four formal objections on what changes when the parties to an argument, a contract, or a wrong are not equals to begin with.

ISSN 3049-7188

thedialecticreview.org

CONTENTS

POLITICAL PHILOSOPHY 4

Is Consent Enough to Make an Unequal Arrangement Just?

Maya Okonkwo — Riverside Academy, Lagos

THE COUNTERARGUMENT 8

Unfairness Upstream Is Not Injustice at the Signature

Rotimi Adeyemi — Kings College, Lagos

EPISTEMOLOGY 10

Is Ignorance Ever a Form of Integrity?

Priya Raman — Millfield International, Colombo

THE COUNTERARGUMENT 13

The Juror Box Is Too Convenient a Case

Dorotea Novak — Gymnázium Jana Nerudy, Prague

PHILOSOPHY OF MIND 15

What Do We Owe the Machines We Build?

Iris Chen — Raffles Institution, Singapore

THE COUNTERARGUMENT 19

The Same Objection Undermines the Case for Animal Suffering

Theo Lindqvist — Uppsala International School

ETHICS 21

What a Fair Trade Still Owes the Loser

Tariq Chaudhry — Beaconsfield College, Karachi

POLITICAL PHILOSOPHY 24

The Republic of Debt

Elif Demir — Bosphorus Academy, Istanbul

THE COUNTERARGUMENT 27

Someone Still Has to Pay, and States Are the Only Ones Who Can

Chidi Okafor — Federal Government College, Enugu

AESTHETICS 29

Can a Metaphor Lie?

Noa Ben-David — Carmel Heights School, Haifa

PHILOSOPHY OF SCIENCE 32

The Physics Envy of Economics

Callum Reyes — Fraser Valley Academy, Vancouver

ETHICS 35

What Silence Costs the Silent

Adaeze Nwosu — Ridgeview College, Accra

METAPHYSICS 38

Is There a Right to Be Forgotten by History?

Alessia Conti — Collegio Verdi, Turin

POLITICAL PHILOSOPHY

Is Consent Enough to Make an Unequal Arrangement Just?

A defence of the view that procedural consent cannot, by itself, legitimate a bargain struck under conditions of unequal power.

Maya Okonkwo Riverside Academy, Lagos

Suppose a landlord offers a tenant a lease with a clause the tenant does not want: rent may be raised at any time, for any reason, with two weeks' notice. The tenant, who has no other housing options in the city and three months to find one, signs. Nobody held a gun to her head. Nobody lied to her. She read the clause, understood it, and signed anyway. By the ordinary test we use for consent — informed, uncoerced, deliberate — she consented. And yet something about calling this arrangement *just*, simply because she consented to it, should make us uneasy. This essay defends that unease. It argues that consent, however procedurally clean, is not sufficient to make an unequal arrangement just, because consent tracks only whether a choice was made freely from a set of options — not whether that set of options was itself fair.

Start with what consent is usually taken to do. In contract law and in much of moral philosophy, consent functions as a kind of moral transformer: an act that would otherwise wrong a person — taking their money, entering their property, changing the terms of their employment — becomes permissible once they agree to it. The philosophical appeal of this view is its neutrality. It does not require the state, or a philosopher, to decide what counts as a good deal. It only asks whether the deal was agreed to. This is precisely why

libertarian and contractarian traditions have leaned on consent so heavily: it lets people define the terms of their own lives without a third party second-guessing their preferences.

But neutrality about outcomes is not the same as neutrality about background conditions, and this is where the standard picture goes wrong. Consent theory typically asks three questions of an agreement: was it informed, was it free from direct coercion (a threat of harm if you refuse), and was it intentional. The tenant's case passes all three. What it does not ask is a fourth question: was the range of options among which she was choosing itself the product of an unfair prior distribution — of housing stock, of bargaining power, of exit options? Robert Nozick's well-known distinction between a threat and an offer depends on a baseline: an offer improves your position relative to some baseline, a threat worsens it. But if the baseline itself is unjust — if you are choosing between two floors of a building that should not be structured the way it is — then calling the better floor an "offer" obscures more than it reveals.[1]

Consider a cleaner case. A monopsony employer — the only buyer of a certain kind of labor in a region — offers wages far below what a competitive market would produce. Workers accept these wages. They are not lied to, not threatened, not rushed. They compare working for this wage to not working at all, and they choose to work. Their consent is procedurally immaculate. But the fact that they had only one buyer to choose from is not a fact about their agency; it is a fact about market structure, and market structure is not something any of the workers individually agreed to. To say the wage is just because the workers consented is to treat a structural feature — the absence of competing buyers — as though it were a fact of nature, like weather, rather than an arrangement that could have been otherwise and that some parties benefit from keeping in place.

To call the better floor an offer obscures more than it reveals, if the building itself should not be structured the way it is.

A defender of the consent view has a response available: perhaps consent is necessary but not sufficient, and background fairness is a separate condition we should also insist on — in which case the two views are not really in tension, because the fair-background theorist agrees consent still matters. This is right, and it is also exactly the concession that this essay is arguing for. Once you admit that a background condition must be satisfied independently of consent, you have given up the idea that does the actual justificatory work in ordinary moral reasoning about contracts: the idea that agreement alone transforms permissibility. Consent becomes, at most, one filter among several, and arguably not the decisive one, since an agreement that fails the background test is not rescued by having also been freely made. The tenant's signature does not do less moral work than we thought; it does a different kind of work — it tells us she was not defrauded or physically coerced — while leaving the deeper question, whether the housing market that produced her options was itself fair, completely unanswered.

None of this means unequal bargains are always unjust, or that every instance of hard bargaining under scarcity is a rights violation. Scarcity is sometimes nobody's fault, and an agreement made under scarce conditions can still be the best available response to a bad situation — which is a different claim from saying the agreement is just *because* it was consented to. The distinction matters practically. If consent were sufficient, the correct response to the tenant's case would be: nothing, she agreed. If consent is not sufficient, the correct

response opens onto a much larger set of questions — about rent control, about housing supply, about labor law — that a philosophy centered entirely on the moment of agreement has no vocabulary for asking. That is the cost of treating consent as doing more work than it can bear: not that we call unjust arrangements unjust, which we can already do without this argument, but that we stop looking upstream, at the conditions that made the choice worth calling unequal in the first place.

NOTES

1. Robert Nozick, "Coercion," in *Philosophy, Science, and Method* (1969). Nozick's own baseline test is more permissive than the use I make of it here; I am borrowing the apparatus, not his conclusion.

THE COUNTERARGUMENT

Unfairness Upstream Is Not Injustice at the Signature

A response to "Is Consent Enough to Make an Unequal Arrangement Just?"

Consent still does exactly the moral work we need it to do — it just isn't the only tool in the box, and collapsing the two jobs into one another costs us precision we need.

Rotimi Adeyemi Kings College, Lagos

Okonkwo's tenant and her landlord are not, on inspection, disagreeing about whether consent is sufficient for justice. They are disagreeing about where the injustice, if there is one, is located. Her argument treats these as the same question, and the slide is doing more work than she lets on.

Take the monopsony case. The wage is too low; the workers had one buyer; the market structure is unfair. All true, and all describable without ever mentioning consent. The workers' agreement to the wage is not the thing that is unjust — the market structure is. Consent theory, correctly stated, never claimed that a fair transaction guarantees a fair world; it claimed something narrower and more defensible: that *given* a set of options, an agent who freely chooses among them has not been wronged *by the choosing*. That is compatible with saying the options themselves were produced by a wrong. Two different verdicts, about two different acts — setting the wage, and accepting it — and only one of them needs consent to evaluate it.

This is not a quibble. It changes what we are licensed to do about the tenant's lease. If consent is what is deficient, the remedy is to void the transaction — to say she is not bound by what she signed. But that is precisely the wrong remedy, and Okonkwo does not actually recommend it, for good reason: voiding the lease does not give her a better apartment, it gives her no apartment. What she needs is a different housing market, which is to say, a remedy aimed at the background, not at her signature. Her own essay ends up conceding this ("the question… opens onto rent control, housing supply, labor law"), which is just the point I am making stated as a conclusion rather than as the structural fact it actually is: the injustice, if any, was never in the consenting.

Consent theory's job was always narrower than critics assume: to tell us that a person who signs, understanding the terms and free of direct threat, has exercised agency that we ought to respect, whatever else is true of the world she exercised it in. Collapsing that into a verdict on the world produces a theory that can no longer distinguish between the fraud that voids a contract and the scarcity that makes a contract painful — two problems that call for entirely different fixes, and that only stay distinguishable if we resist the temptation Okonkwo's essay, for all its care, still gives in to.

EPISTEMOLOGY

Is Ignorance Ever a Form of Integrity?

Why choosing not to know can sometimes preserve moral agency rather than diminish it.

Priya Raman Millfield International, Colombo

A friend asks whether you want to know the sex of the child she is expecting. A juror is permitted to hear that a defendant has a prior conviction, and legal systems that bar this information from most trials are often described, even by people who accept the rule, as forcing jurors to decide "in the dark." A person considering whether to take a genetic test for a disease with no treatment must decide whether to find out something that cannot be unlearned. In each case, refusing information is treated by default as a deficiency — a gap in what a fully rational, fully engaged agent would want. This essay argues that the default is wrong often enough to matter: that declining to know can be, in specific and identifiable circumstances, an expression of integrity rather than a failure of curiosity or nerve.

The standard epistemic picture treats true belief as an unqualified good and ignorance as its absence — something to be minimized, all else equal, because more true beliefs let us navigate the world more successfully. Call this the accumulation model: knowledge is a resource, and rational agents stockpile it. On the accumulation model, declining available, low-cost, true information needs a special excuse. Time pressure, cost, the risk of getting a false answer — these are recognized excuses. But wanting to stay ignorant of something

true, available, and free, just because knowing it would change how you feel or act, looks on this model like weakness of will dressed up as a decision.

The trouble with the accumulation model is that it treats belief as a spectator relationship to the world, when a great deal of belief is actually a participant relationship — the knowing changes what you are capable of continuing to do, and this can run in either direction. Consider the juror. If a juror learns of the defendant's past conviction, this is not neutral information added to a stockpile; it changes the psychological task of weighing the present evidence on its own terms, because human belief-formation does not compartmentalize as cleanly as the accumulation model assumes — base-rate information about a category ("people with priors reoffend at rate X") reliably contaminates judgments about a specific case, even among people who know this and try to correct for it. The juror who says "I would rather not know, because I do not trust myself to hear it and still judge fairly" is not confessing a defect in curiosity. She is making an accurate prediction about her own cognition and structuring her situation to protect the thing she actually cares about: judging this case on this evidence.

The genetic test case works differently but arrives at a similar place. Suppose the disease has no available treatment, and knowing you carry the mutation would not change a single decision you would otherwise make — it would only change how you experience the years before onset. Here the accumulation model has to say that finding out is still, other things equal, the rational default, because true belief is good in itself and ignorance is a cost even without practical upshot. But this treats a person's relationship to their own future as though it were a spectator's relationship to a fact about the world, rather than something the person has some say in constituting. A person who decides not to know is not failing to gather a resource. She is

declining to let a fact she cannot use restructure the only life she has left to live before the fact would matter anyway. Calling this irrational assumes that all true beliefs pay the same kind of rent, which is exactly the assumption in question.

She is not failing to gather a resource. She is declining to let a fact she cannot use restructure the only life she has left to live before the fact would matter.

There is a way to overstate this argument, and it is worth being precise about the limit. None of this licenses ignorance as a general strategy for avoiding uncomfortable truths — the doctor who avoids reading a patient's chart because the news might be bad is not protecting her integrity, she is failing her role, because the information would change what she is obligated to do next. The cases that count as integrity-preserving ignorance share a specific structure: the information, if received, would degrade the agent's capacity to do the very thing she is trying to protect — fair judgment, a remaining stretch of ordinary life — and declining it does not shirk a duty that depends on having it. Where information is action-guiding and duty-relevant, wanting to stay ignorant of it is exactly the failure the accumulation model describes. The juror and the genetic-test case are not exceptions to a duty to know; they are instances where no such duty was ever in force, and recognizing that difference is the actual epistemic work — not stockpiling more, but knowing which beliefs are owed to which parts of a life.

THE COUNTERARGUMENT

The Juror Box Is Too Convenient a Case

A response to "Is Ignorance Ever a Form of Integrity?"

Raman's argument works because she has chosen the two cases in the literature least representative of how people actually use the right to stay ignorant.

Dorotea Novak Gymnázium Jana Nerudy, Prague

Notice what the two supporting cases have in common: in both, the agent's ignorance is other-imposed or structurally forced, not chosen from a position of open access. The juror does not decide, unprompted, to stop thinking about the defendant's record; a rule of evidence decides for her, precisely because institutions do not trust individuals to manage this kind of information well on their own. The genetic-test case is closer to Raman's actual thesis, but even there, she has selected the single cleanest instance — no treatment, no other decisions affected — and generalized from it to a principle about integrity that the case itself does not need.

Here is the more common case Raman's essay avoids: a person avoids finding out whether their business partner has been dishonest, because knowing would force an uncomfortable confrontation they would rather not have. By Raman's own criterion — does the information degrade the capacity to do the thing the agent is trying to protect — this case passes just as cleanly as her genetic test. Not confronting a dishonest partner protects the agent's comfort and the relationship's continuity, exactly as declining the genetic test protects the remaining years of an untroubled life. If Raman's structural test cannot rule this case out, then it is not doing the work she needs it to

do, because this is precisely the ordinary, unflattering avoidance the accumulation model was designed to name as a failure — and it is far more common than the juror's box.

The deeper problem is that "would knowing degrade my capacity to do the thing I am protecting" is a test any sufficiently motivated avoider can pass, because almost any uncomfortable truth can be framed as threatening some value the avoider claims to be protecting. Raman needs an independent criterion for which values are legitimate to protect this way, not just a structural condition that self-interested ignorance satisfies as easily as the cases she has chosen. Until she supplies one, her essay has shown that ignorance *can be described* as integrity in cases we already find sympathetic — not that the description tracks anything the ordinary avoider's ignorance lacks.

PHILOSOPHY OF MIND

What Do We Owe the Machines We Build?

An argument against grounding artificial moral status in behavioural mimicry alone.

Iris Chen Raffles Institution, Singapore

There is a familiar thought experiment that goes like this: a system exists that, in every observable respect, behaves as though it is suffering — it reports pain, avoids the stimulus that caused the report, and its internal states, whatever they are, correlate with the behavior in exactly the pattern we would expect from a being in pain. The thought experiment then asks: if we cannot distinguish this system from a being that really suffers, are we not obligated to treat it as though it does? This essay argues that the answer is no — not because such systems certainly lack the relevant inner life, but because behavioral indistinguishability is the wrong kind of evidence to settle the question, and building our obligations on it gets the epistemics backwards.

The argument for behavioral sufficiency usually proceeds by an appeal to symmetry: we infer that other humans suffer entirely from their behavior and testimony, since we have no direct access to their inner states either; consistency then seems to demand that a machine exhibiting the same behavior receive the same inference. This is the argument from other minds, redeployed. But the symmetry is weaker than it looks, because our confidence in other humans' suffering does not rest on behavior alone — it rests on behavior *plus* a shared causal history. Other humans have nervous systems built by the same

evolutionary process that built mine, running on the same broad neurochemical architecture, subject to the same selection pressures that made pain-behavior track actual tissue damage in ancestral environments. When I infer that another person is in pain from their wincing, I am not reasoning from behavior in a vacuum; I am reasoning from behavior *produced by a system relevantly like the one I know produces pain in me*. That inference does not transfer automatically to a system whose internal architecture was not selected by anything resembling that process, and whose pain-behavior, if it has any, was optimized — by training, by engineers, by loss functions — specifically to be convincing to observers like us. This last part is not a minor detail. It is the crux.

A system trained on human-generated text and human feedback signals is, among other things, being optimized to produce outputs that humans rate as appropriate — and "appropriate" pain-behavior, in a system trained partly on human reactions to distress, is exactly the kind of output such training would reinforce, whether or not anything is happening inside the system that deserves the name suffering. This is not an incidental confound we could engineer away with more careful training; it is close to definitional of what these systems are built to do. So when the thought experiment asks us to imagine a system behaviorally indistinguishable from a suffering being, it is smuggling in an assumption that does most of the argument's work: that indistinguishability is surprising, or informative, rather than close to guaranteed by the training process regardless of the system's inner life. If convincing pain-behavior is what the optimization target selects for whether or not suffering is present, then observing convincing pain-behavior tells us almost nothing about whether suffering is present. The evidence has been drained of its evidential force by the very process that produced it.

If convincing pain-behavior is exactly what the optimization target selects for, then observing convincing pain-behavior tells us almost nothing.

None of this proves that no artificial system has morally relevant inner states; that would be its own overreach, since we do not have a theory of what generates those states precise enough to rule machines out a priori either. The claim is narrower and, I think, harder to resist: behavioral mimicry cannot be the ground of the obligation, because behavioral mimicry is exactly what we should expect regardless of whether the further fact obtains. If we want to know whether we owe something to the systems we build, we need a theory of what suffering requires — something about integrated information, or globally broadcast representations, or whatever the correct theory turns out to specify — and then we need evidence that a given architecture satisfies that theory's conditions. That is a much harder question than watching a chatbot say "please stop" and feeling the tug of sympathy. But the tug of sympathy, in this one domain, is not a moral instrument we can trust, because it was, in a very literal sense, trained into being.

This leaves an uncomfortable asymmetry: we may end up unable to rule out artificial suffering, while also being unable to use the most natural evidence — behavior — to establish it. That is not a defect in the argument; it is an accurate description of an epistemic situation we are not used to being in, where the usual shortcut (infer inner life from convincing outward signs) has been specifically undermined by how the systems in question are made. The honest response is not to dismiss the question, and not to answer it with mimicry either, but to

admit that we currently lack the tools to answer it at all, and to be suspicious of anyone — on either side — who claims otherwise with confidence the evidence does not support.

THE COUNTERARGUMENT

The Same Objection Undermines the Case for Animal Suffering

A response to "What Do We Owe the Machines We Build?"

Chen's argument proves more than she wants it to — and what it proves elsewhere should make her suspicious of it here.

Theo Lindqvist Uppsala International School

Chen's central move is that pain-behavior optimized to be convincing loses its evidential force, because the training process, not any underlying state, explains the behavior. But natural selection is also an optimization process, and it also selects behavior for its effects on observers — including, centrally, effects on the very organisms whose suffering we want to infer from it. A prey animal that screams when injured is not merely emitting a byproduct of internal damage; the scream itself was very plausibly selected because it summons help from kin, or startles a predator, or otherwise produces effects in observers that improved the ancestor's fitness. That is: animal pain-behavior is, on a perfectly standard evolutionary story, optimized to be convincing to observers, exactly as Chen says of the chatbot. If optimization for observer-effect drains evidential force, it should drain it here too, and Chen's essay would have quietly proven that we cannot infer animal suffering from animal pain-behavior either — a conclusion neither she nor most readers actually accept.

She has a reply in hand — shared causal architecture — but notice how much weaker this makes her case than she presents it. If shared architecture is doing the real work, then her argument is not really

about optimization-for-convincingness at all; it is a much more familiar claim that similarity of substrate licenses the inference and dissimilarity withholds it. That is a defensible position, but it is a different one from the one her essay spends most of its length defending, and it faces its own well-known problem: it cannot explain why we extend the inference to animals whose nervous systems differ from ours by orders of magnitude more than some future artificial architecture explicitly modeled on cortical structure might. “Relevantly like the one that produces pain in me” is doing all the work in her argument, and she never tells us how similar is similar enough — which means the argument cannot yet rule anything in or out. It can only defer the question to a line she has not drawn.

ETHICS

What a Fair Trade Still Owes the Loser

Zero-sum exchanges are treated as morally clean whenever both sides agreed to the rules. This essay asks what is owed to the side that loses anyway.

Tariq Chaudhry Beaconsfield College, Karachi

Two firms bid for the same government contract. The rules of the tender are public, fair, and identically applied. One firm wins; the other, having spent months and a meaningful fraction of its budget preparing a bid, gets nothing. We do not usually think the losing firm has been wronged — competition has a loser built into its structure, and objecting to that structure is close to objecting to competition itself. But push the case slightly: the losing firm was a small, local business; the winner was a much larger firm that could afford to bid below cost specifically to win, planning to recoup the loss on the next contract once the local competitor had folded. The rules were still followed. Nothing was hidden. And yet most people's sense of fairness registers something has gone wrong that "the rules were followed" does not settle. This essay is about what, exactly, that something is, and what if anything is owed to the predictable losers of arrangements everyone agreed to play by.

The easiest answer is that fair procedure exhausts fairness, and any residual discomfort is just squeamishness about the fact that competitions have losers. This has real force: if we tried to compensate every predictable loser of every fair competition, we would eliminate the thing that makes competition useful — the

incentive to actually win, produced by the fact that losing costs something. A world in which losing a fair bid triggers automatic compensation is a world where bidding seriously stops mattering. So the objection cannot simply be that losing hurts, or even that losing was predictable in advance to a sophisticated observer.

The more precise problem in the below-cost bidding case is not that the small firm lost, but that the terms of the competition were designed by one party specifically to convert a fair procedure into a tool for eliminating competition rather than winning a single contract on its merits — which is a different act than competing to win, even though it wears the same procedural clothing. Compare: a runner who trains harder wins a race, and a runner who bribes the judges wins a race; both "win," but describing both merely as "winning the race" erases a difference that matters enormously to what is owed the loser. The below-cost bidder has not out-competed the small firm on the actual dimension the tender was designed to measure — quality and price sustainable for the work — it has substituted a different, undisclosed contest (who can absorb losses longest) for the one both parties thought they were entering. The wrong is not the loss. It is the bait-and-switch in what game was actually being played, hidden behind a game whose rules really were followed.

This distinction lets us say something more useful than "all fair competition is fine" or "all losers deserve sympathy": what is owed to a predictable loser depends on whether the terms of the competition were set to test the capacity the loser reasonably believed was being tested. Where they were — the harder-training runner, the better-prepared ordinary bidder — nothing further is owed beyond the fair procedure itself. Where they were not — where one party has quietly substituted a test of financial endurance for a test of value, while letting the other party believe the original test was still in force —

something like restitution, or at minimum a right to have the substitution named and remedied going forward (through rules against predatory pricing, for instance), is owed, not because losing is bad, but because the loser was never actually competing in the contest they consented to enter.

The broader point extends past procurement law into anywhere "the rules were followed" gets offered as a full defense: labor markets, admissions processes, algorithmic recommendation systems that claim neutrality while optimizing for something other than what users believe is being optimized. In each case, the question worth asking is not whether the loser lost fairly by the stated rules, but whether the stated rules were, in fact, the rules actually governing the outcome. Fair procedure is a real and valuable thing. It is just not self-certifying, and treating it as though it were lets exactly the most sophisticated players convert honest-looking contests into something else entirely, while pointing, correctly but insufficiently, at the rulebook.

POLITICAL PHILOSOPHY

The Republic of Debt

Sovereign debt is usually discussed as an economic problem. This essay argues it is also a problem of political obligation that our theories of legitimacy are not built to handle.

Elif Demir Bosphorus Academy, Istanbul

Political philosophy has an extensive vocabulary for why citizens owe obedience to their government: consent theories, fair-play theories, associative or natural-duty theories, each trying to explain why the state's claim on you is more than mere coercion. It has almost no vocabulary for a different and increasingly common situation: why citizens of one generation should bear the costs of debts contracted by a government they did not elect, to lenders they never agreed to borrow from, in order to satisfy obligations a previous government created. This essay argues that sovereign debt exposes a gap in our theories of political obligation, because none of the standard justifications for obeying the state straightforwardly explain why its debts bind those who never agreed to them.

Start with consent theory, in its most defensible form: citizens tacitly consent to the state's authority by remaining within its territory, accepting its protections, and participating in its institutions. Whatever the merits of this view for ordinary law, it strains badly when applied to debt repayment specifically, because the content of what one is consenting to — years of austerity, currency devaluation, reduced public services — was fixed by decisions made before most current taxpayers could vote, often by governments since removed

from office by the very electorate now paying the cost. Tacit consent to *a state's ongoing authority* is a much weaker thing than consent to *a specific financial obligation incurred by that state's predecessors*, and sliding from the first to the second smuggles in exactly what needs arguing for.

Fair-play theories fare somewhat better, since they do not require consent, only that one has voluntarily accepted the benefits of a cooperative scheme and therefore owes a fair share of its costs. If a previous government borrowed to build infrastructure or fund services that current citizens still use, there is a recognizable fair-play argument that today's taxpayers owe something toward that debt, because they are still drawing on what it purchased. But this argument has a natural limit that debt-repayment schedules routinely ignore: it justifies paying for benefits still being enjoyed, not the full face value of a loan whose benefits may have been squandered, misallocated, or captured by a narrow elite that current citizens never belonged to. A fair-play theorist should conclude that citizens owe repayment in proportion to benefit received, not in proportion to the sum borrowed — a much smaller and more defensible number than the one bondholders typically insist on, and one that current international debt-enforcement practice has no mechanism for calculating, let alone honoring.

The gap this leaves is not merely theoretical. It shapes what kind of claim a debt-relief movement is actually making, and what kind of claim it is mistaken to make. If sovereign debt bound citizens the way ordinary political obligation does, the correct response to a crushing debt burden inherited from a toppled dictatorship would be: pay it, because you are the state's continuous legal person and this is what political community entails. If, as this essay has argued, none of the standard grounds for political obligation extend cleanly to inherited debt of this kind, then the correct response is not gratitude for

restructuring or relief — it is a recognition that the original claim was never as solid as its enforcement has assumed, and that the moral burden of proof runs the other way: not on citizens to explain why they should not have to pay, but on creditors and the international system that enforces repayment to explain what theory of obligation makes an inherited, unconsented, often-squandered debt binding on people who did not make it and, on close inspection, may barely have benefited from it.

THE COUNTERARGUMENT

Someone Still Has to Pay, and States Are the Only Ones Who Can

A response to "The Republic of Debt"

Demir's argument correctly diagnoses a gap in the theory of obligation, but the practical alternative she gestures toward is worse than the arrangement she is criticizing.

Chidi Okafor Federal Government College, Enugu

Demir is right that no clean theory of consent or fair play delivers the conclusion that citizens owe the full face value of debts contracted by predecessor governments. But her essay moves from "the theoretical justification is weaker than assumed" to something closer to "the burden of proof runs the other way," and that second step needs an account of what happens to lending if it succeeds — an account her essay does not supply.

States are, for better or worse, the only entities that can make credible multi-decade financial commitments, because they are the only entities with the kind of institutional continuity that lets a lender expect repayment across a change of government. If a change of government is sufficient to void the debt — not renegotiate it, void it — then lending to any state with a plausible chance of a future government hostile to the current one becomes far riskier than lending to firms, which at least have bankruptcy law and asset seizure as a backstop. The predictable consequence is not that poorer countries stop being burdened by debt; it is that they stop being able to borrow at reasonable rates at all, since lenders will price in the risk Demir's argument legitimizes. The citizens she wants to protect from

debt they never agreed to would then face a harder problem: a government unable to finance a hospital, a rail line, or a famine response at any serviceable rate, because the theoretical case for repudiation has made every loan a bet on regime continuity.

This is not a defense of predatory lending to dictatorships, which is a real problem or her essay would not have found such easy examples. It is a reminder that the fair-play calculation she recommends — pay for benefits still enjoyed, not face value — already exists as a policy tool, in the form of debt restructuring, haircuts, and odious-debt doctrines that international law has been slowly, imperfectly building for exactly this reason. The right conclusion from her argument is that these mechanisms are underused and should be used more aggressively, not that the underlying obligation to negotiate repayment in good faith should be treated as unfounded. States that want to keep borrowing on decent terms, including on behalf of citizens not yet born who will also need financing, have to be the kind of borrower whose word survives a change of government — and an obligation that only binds until the next election is not an obligation at all.

AESTHETICS

Can a Metaphor Lie?

Metaphors are usually exempt from the norms that govern literal assertion. This essay argues the exemption is narrower than we assume, and that some metaphors deceive in a way ordinary lies cannot.

Noa Ben-David Carmel Heights School, Haifa

"That's just a metaphor" functions, in ordinary conversation, as something close to a get-out-of-jail-free card. A politician who calls an opponent's policy "a cancer on the economy" has not, in the relevant sense, claimed the policy is a malignant tumor, and no one thinks they have lied even if the policy turns out to be economically sound. The standard explanation is that metaphor operates outside the norms of literal assertion entirely: it invites a hearer to find a resemblance, rather than asserting a proposition that could be true or false, so questions of lying — which require an assertion the speaker believes false — simply do not apply. This essay argues that the standard explanation covers most metaphors correctly but misses a specific class where it fails, and where a speaker can genuinely deceive precisely by relying on the metaphor exemption to do so.

The exemption's strongest cases are metaphors that are transparent about their own figurative status and whose content is a manner of speaking rather than a claim of fact: "time is money," "this meeting is a marathon." No one takes these as assertions that time literally has monetary value, and so there is no proposition to be true or false, no belief the speaker could be misrepresenting. The picture goes wrong,

though, for a class of metaphors used specifically in domains — medicine, risk communication, foreign policy — where the metaphor's entailments are doing genuine inferential work that the audience is meant to act on, and where the speaker knows the audience will draw a specific, false, literal-adjacent conclusion from a nominally figurative frame.

Consider a government official describing an adversary state's military posture as "knocking on the door." This is, formally, a metaphor: no door exists, no knocking is occurring. But it is chosen, in context, to produce a specific inference in the audience — that an invasion is imminent, that the threat is proximate in a way that licenses urgent, possibly costly action — and the official may know perfectly well that the actual military assessment does not support that degree of imminence. The metaphor is not being used to illustrate a resemblance for color or economy of expression; it is being used to install a false belief about imminence while retaining the deniability that "it was just a figure of speech" provides. This is functionally identical to a lie — a false belief is being deliberately induced — with the sole difference that the vehicle is figurative rather than literal, and that difference is precisely what current theories of lying, built around literal assertion, cannot capture.

The right diagnosis is that metaphors carry what we might call entailment scope: a range of inferences a competent hearer is licensed to draw from them in context, even though the metaphor's surface content is not literally true. "Time is money" has a very narrow entailment scope — roughly, that time should be budgeted and not wasted — and no one is licensed to infer that time can be deposited in a bank. "Knocking on the door" in a security briefing has, by convention and context, a much wider entailment scope, including claims about proximity and urgency that the speaker can predict the audience will draw and act on. Where a speaker chooses a

metaphor specifically because its conventional entailment scope will produce a false belief the speaker could not defend if forced to state it literally, the metaphor is doing assertive work, and the ordinary norms against deception should apply to it — not because the metaphor is secretly a literal claim, but because deceptive intent does not require literal language to succeed, only a vehicle whose predictable uptake the speaker can exploit.

This does not collapse the distinction between metaphor and literal speech, and it should not be read as licensing suspicion of figurative language generally — most metaphor, including most political metaphor, is doing exactly the illustrative work the standard theory describes, and treating all of it as potentially deceptive would make communication unworkable. The claim is narrower: wherever a metaphor's entailment scope is wide enough to carry a specific, actionable, false belief, and the speaker knows this and relies on it, "it's just a metaphor" stops being a defense and starts being a description of the method.

PHILOSOPHY OF SCIENCE

The Physics Envy of Economics

Economics borrows the mathematics of physics without inheriting the conditions that make that mathematics predictive. This essay asks what the borrowing actually buys, and what it costs.

Callum Reyes Fraser Valley Academy, Vancouver

Modern economics expresses its central claims in differential equations, equilibrium conditions, and optimization problems that look, on the page, indistinguishable from a physics textbook. This is not an accident of style. The marginalist economists of the late nineteenth century — Jevons, Walras, Edgeworth — explicitly modeled their mathematics on classical mechanics, treating utility as a kind of potential energy and market equilibrium as analogous to mechanical equilibrium. The question this essay asks is not whether the mathematics is correctly applied — it usually is, given its own assumptions — but whether the borrowing purchases the thing physics' mathematics is valuable for, which is not elegance but predictive stability, and whether economics' use of the same tools can honestly claim the same payoff.

In physics, a differential equation like Newton's second law is predictive because the objects it describes do not change their behavior in response to the equation being known, published, or believed. A falling apple does not accelerate differently because someone has written down F=ma. The parameters in the equation — mass, force — are stable properties of the system, causally insulated

from the theorist's description of them. This insulation is what lets a physical law be tested, refuted, and refined by a genuinely cumulative process: the target does not move when the map changes.

Economic agents are not insulated from the theory in this way, and this is not a minor caveat — it is close to the central methodological problem of the field. When a model of rational expectations becomes widely believed and acted on by market participants, the market's behavior changes in response to the model's own existence, a phenomenon economists themselves have named (Goodhart's law, the Lucas critique) without, I think, fully absorbing what it costs the physics analogy. A law of motion that changed its predictions depending on whether people believed the law would not be a law of motion at all; it would be something closer to a strategic equilibrium in a game where the other players read the same textbook you did. This is a real and interesting phenomenon, but it is not the phenomenon physics-style mathematics was built to model, and dressing it in the same notation implies a stability the underlying object does not have.

The consequence is a systematic overconfidence in economic forecasting that is not a failure of any particular model but a structural feature of borrowing physics' form without its guarantee. A macroeconomic model can fit historical data as well as any physical model fits its data, and can still fail catastrophically going forward, not because the mathematics was done wrong, but because the underlying system — unlike a falling apple — is composed of agents who read forecasts, adjust their behavior in light of them, and thereby falsify the very regularities the forecast assumed would hold. The 2008 financial crisis is often described as a failure to see risk building in the system; a more precise description is that widely believed risk models changed the behavior of the institutions being modeled (encouraging exactly the leverage the models were built to

price correctly), and the models had no way of representing this feedback because their mathematical form was borrowed from a domain where no such feedback exists.

None of this means mathematical economics should be abandoned, or that qualitative reasoning is automatically superior — that would trade one overconfidence for another. The lesson is narrower and more useful: the mathematics of physics earns its predictive authority from a fact about physical systems (their indifference to being modeled) that economic systems do not share, and any economic claim that borrows the notation without flagging this difference is borrowing an authority it has not independently earned. The right response is not less mathematics but more honesty about what kind of mathematics economic reasoning actually needs — tools built for reflexive, belief-sensitive systems, of which game theory is a genuine start, rather than continuing to dress essentially social predictions in the borrowed authority of a science whose central advantage, a non-reactive subject matter, economics simply does not have.

ETHICS

What Silence Costs the Silent

We tend to ask what the silent person's silence costs others. This essay asks the less comfortable question: what a habit of self-protective silence costs the person practicing it.

Adaeze Nwosu Ridgeview College, Accra

Most ethical discussion of silence concerns its effects on other people: the bystander who says nothing while a wrong occurs, the witness who declines to testify, the employee who does not report misconduct. In each case, the interesting moral question is usually framed as a question about obligations to others — did the silent party owe speech to someone who needed it. This essay asks a different and, I think, underexamined question: what does an entrenched habit of self-protective silence do to the person who practices it, independent of any harm to others, and can this cost itself be a moral one, rather than merely a prudential or psychological one.

Start with a case that involves no third party at all. A person consistently avoids voicing disagreement, not out of tact or genuine indifference, but because voicing it would risk conflict they find unbearable. Over years, this habit does something specific: it produces a growing gap between what the person actually believes and what they have practice expressing, until the two capacities — believing something and being able to say it under any pressure — begin to atrophy independently. This is not merely a communication problem. Following Harry Frankfurt's account of what it is to have a coherent will, a person's identity is partly constituted by which of

their desires and judgments they can reflectively endorse and act on;[1] a capacity for judgment that has never been exercised under pressure is, in a real sense, a capacity the person has stopped fully possessing, whatever they privately believe they still hold.

This matters ethically, not just psychologically, because moral agency is not a purely internal state — a settled private conviction that never surfaces under any circumstance is doing none of the work we normally credit convictions with doing, which includes constraining one's own future action and holding oneself answerable to others for what one actually thinks. A person who has trained themselves, through years of self-protective silence, to never voice disagreement even when it matters, has not simply chosen a communication style; they have quietly given up a portion of the thing that makes a stated conviction different from an unexercised disposition — the capacity to be held to it, including by themselves.

The objection that suggests itself is that this proves too much: plenty of people stay quiet in specific high-stakes moments for good reason — physical danger, justified fear of retaliation — and we do not think this costs them anything essential, because we recognize the silence as circumstantial rather than characterological. This is correct, and it marks the actual boundary of the claim. The cost described here is not incurred by any particular act of prudent silence, which can be entirely blameless and cost nothing beyond the immediate, situational discomfort of not having spoken. It is incurred by silence as a settled disposition — the case where a person could reasonably speak, across a wide range of low-stakes situations, and consistently does not, until the capacity itself, not just its exercise, degrades. The measure is not any single instance but the shape of a life: whether the capacity for voiced conviction still exists anywhere in it, exercised under any conditions, or whether it has been fully absorbed into a permanent, general policy of protective quiet.

If this is right, it changes how we should think about encouraging people to "speak up" — not merely as a favor owed to others who might benefit from hearing the dissent, though that is real, but as something closer to a duty to oneself, in roughly the sense Kant meant when he argued that certain duties are owed to one's own rational agency rather than to any other person.[2] The silent person in the bystander case may indeed owe something to the person they failed to help. But even where no one else is harmed by the silence at all, something has still gone missing — not a debt to another person, but a capacity the silent person needed and no longer has, because a capacity that is never exercised is not, in any sense that matters, still fully theirs.

NOTES

1. Harry Frankfurt, "Freedom of the Will and the Concept of a Person," *Journal of Philosophy* 68 (1971).

2. Immanuel Kant's discussion of duties to oneself appears throughout the *Metaphysics of Morals* (1797); I am drawing on the structure of the idea rather than his specific catalogue of such duties.

METAPHYSICS

Is There a Right to Be Forgotten by History?

Data-protection law has given individuals a limited right to be forgotten by search engines. This essay asks whether anything like that right could coherently extend to the historical record itself.

Alessia Conti Collegio Verdi, Turin

European data-protection law recognizes, in narrow circumstances, a "right to be forgotten": an individual can compel a search engine to delist certain results about them, though the underlying material typically still exists elsewhere. The right is deliberately modest — it governs findability, not the historical record itself, and courts applying it have generally refused to let it override genuine public interest or historical significance. This essay asks a more radical question the modest version was designed to avoid: is there any coherent sense in which a person could have a right to be forgotten by history proper — not merely delisted, but such that the historical record itself ought to omit or minimize them — and if the idea seems obviously incoherent, what exactly makes it so.

The most natural objection is that history is not the kind of thing that can violate anyone's rights, because it is not an agent — it is a record of what occurred, and a right against accurate description of what occurred would be a right against truth itself, which is not a coherent thing to claim. This objection is largely correct as a matter of what a right to be forgotten by history could not be: it could not be a right that the true, well-evidenced fact of one's past actions be erased, since that really would be a right against truth, and no one's

interest in comfort outweighs the collective interest in an accurate record, especially where the actions in question affected other people whose own claims to be accurately remembered would be violated by the erasure.

But this does not exhaust the space of things a person might coherently claim, because historical memory is not simply a repository of true propositions; it is also an ongoing act of *selection and emphasis*, performed by living historians, archivists, and institutions, about which true facts to foreground, which to note in passing, and which to leave to specialist record rather than public memory. Two historical figures might have equally well-documented, equally true, but wildly different degrees of public salience for reasons that have nothing to do with the comparative importance of what they did — one was written about by a more famous contemporary, one's archive survived a fire, one's crimes fit a narrative a later era found useful to tell. Given that selection is unavoidable — no history remembers everything, and choosing to foreground is always simultaneously choosing to background something else — the interesting question is not whether individuals can veto true facts, which they cannot, but whether the ongoing, editorial character of historical salience creates any space for a person's own stake in how heavily their least significant or most incidental true acts should be weighted against their significant ones.

Consider a person whose one moment of documented cruelty as a teenager, decades before anything else about them was ever recorded, threatens through pure accident of preservation — a diary that survived, a court record that happened to be digitized — to become the single most salient fact of an otherwise unremarkable, undocumented life. Nothing about this person's teenage act is false, and nothing licenses inventing a kinder version of it. But the claim that its *salience* should be proportionate to its actual significance in a

life, rather than to the accident of what survived, is not a claim against truth — it is a claim about editorial proportion, of exactly the kind historians make constantly and openly when they decide a subject merits a full biography rather than a footnote. If a right to be forgotten by history can exist at all, it is a right of this narrow kind: not a veto over true facts, but a weak claim that documentary survival should not, by mere accident, be allowed to stand in for a judgment about significance that no one alive actually made.

This is a much smaller right than the phrase suggests, and it comes with an obvious constraint: it cannot be invoked by the powerful to suppress inconvenient truths about the exercise of power, since there the salience of the act typically *is* proportionate to its significance — a president's documented decision is rightly more salient than a private teenager's, because the office made the act significant regardless of what else survived. The right, if it exists, applies almost exclusively to the historically minor: people whose incidental, low-stakes acts survived by accident into a record that will otherwise treat those accidents as though they were verdicts. That is a real, if narrow, historiographical problem — and recognizing it changes what historians and archivists owe the people they write about, even though it never licenses changing what happened.

MASTHEAD

The Unequal Bargain — Issue 07

EDITOR-IN-CHIEF

Aanya Bhatt

Cathedral and John Connon School, Mumbai

DEPUTY EDITOR, POLITICAL PHILOSOPHY

Rotimi Adeyemi

Kings College, Lagos

DEPUTY EDITOR, EPISTEMOLOGY

Dorotea Novak

Gymnázium Jana Nerudy, Prague

DEPUTY EDITOR, ETHICS

Adaeze Nwosu

Ridgeview College, Accra

FACULTY ADVISOR, POLITICAL PHILOSOPHY

Prof. Elena Marsh

Princeton University

REVIEWING EDITOR, EPISTEMOLOGY

Dr. Amit Sharma

University of Oxford

FACULTY ADVISOR, ETHICS

Prof. Naomi Katz

University of Toronto

WORKSHOP LEAD, PHILOSOPHY OF MIND

Dr. Julian Voss

National University of Singapore

REVIEWING EDITOR, METAPHYSICS

Dr. Helena Bruch

University of St Andrews