§ THE DIALECTIC REVIEW

Volume II · Autumn 2024

ISSUE 03

The Ethics of Knowing

Autumn 2024 · 6 pieces

Six essays on belief, doubt, and the moral obligations that attach to how — and how carefully — we come to think what we think.

ISSN 3049-7139

thedialecticreview.org

CONTENTS

EPISTEMOLOGY 3

What You're Not Allowed to Believe on Purpose

Nadia Petrov — Novi Sad Gymnasium, Novi Sad

ETHICS 6

The Ethics of Changing Your Mind in Public

Julius Achebe — Brookstone International School, Lagos

POLITICAL PHILOSOPHY 9

Ignorance as a Social Achievement

Amara Diallo — École Internationale de Dakar

EPISTEMOLOGY 12

Can You Know Something You Can't Explain?

Felix Brandt — Gymnasium am Rhein, Cologne

ETHICS 15

The Duty to Believe Well

Saoirse Byrne — Ashgrove Academy, Cork

PHILOSOPHY OF SCIENCE 18

What Maps Leave Out on Purpose

Ren Takahashi — Seiwa Academy, Kyoto

EPISTEMOLOGY

What You're Not Allowed to Believe on Purpose

We cannot simply decide to believe something the way we decide to raise an arm. This essay asks what, exactly, is impossible about direct doxastic control, and why it matters morally.

Nadia Petrov Novi Sad Gymnasium, Novi Sad

Try, right now, to believe — not merely imagine, not merely say aloud, but actually come to believe — that the chair nearest you is made of solid gold. Most people find they cannot do this, however strong the incentive offered for succeeding. This is the phenomenon philosophers call the impossibility of direct doxastic voluntarism: belief does not respond to the will the way ordinary bodily action does, and this essay asks why, and what follows morally from the answer.

The standard explanation appeals to belief's function: a belief is supposed to track truth, to represent how the world actually is, and this function would be undermined if beliefs could be adopted simply because we wanted to hold them, independent of evidence. On this view, the impossibility of believing at will is not an unfortunate limitation on our freedom but a constitutive feature of what belief is for — a state that responds to evidence rather than desire is functioning correctly as a belief; a state that responded to desire instead would be functioning as something else, perhaps closer to a wish or a resolution, wearing belief's grammar without doing belief's job.

This explanation is broadly right, I think, but it leaves an important asymmetry unexplained: while we cannot believe at will directly, we clearly can influence what we come to believe indirectly, over time, through choices fully within our control — which sources we expose ourselves to, which arguments we take the trouble to seriously engage with rather than dismiss, which communities we spend our time in, all of which shape, gradually and without any single moment of "deciding to believe," what ends up seeming true to us. Blaise Pascal's famous advice to the person who wants to believe in God but cannot simply will it — go through the motions, attend, participate, and belief will follow — is exactly an appeal to this indirect route, and its psychological plausibility (habituation genuinely does shape credence over time) is part of what makes it interesting rather than merely cynical.

This indirect control is where the ethics of belief actually lives, and where most everyday moral evaluation of believers is implicitly located, once you look closely. We do not blame people for failing to instantaneously believe true things or disbelieve false ones on command, since this is not a capacity anyone has. We do blame people for the far more common and far more culpable pattern of curating, over years, an information environment that reliably produces convenient beliefs — reading only sources that confirm what they already want to think, treating disconfirming evidence as an occasion to search harder for a dismissive counter-argument rather than to actually weigh it, surrounding themselves exclusively with people who already agree. None of this involves believing at will in the impossible, direct sense. All of it involves a long sequence of small, perfectly voluntary choices whose foreseeable cumulative effect is a belief system insulated from correction — and this is something a person can be responsible for, in exactly the way they cannot be responsible for failing to believe the chair is gold on command.

This reframing matters because it locates epistemic responsibility somewhere real and actionable, rather than somewhere impossible. Telling someone "you should believe differently, right now, by choosing to" asks for something no one can do and is, strictly, a kind of category error. Telling someone "the way you have spent the last several years curating what you read and who you talk to has made it predictable that you would end up believing this, and you had ample opportunity to curate differently" asks for something real, is often true, and is where genuine epistemic blame, when it is deserved, should actually be aimed — not at the belief itself, which was never under direct voluntary control, but at the long chain of freely chosen exposures that made holding it, eventually, almost unavoidable.

ETHICS

The Ethics of Changing Your Mind in Public

Public reversal is usually treated as either admirable intellectual honesty or embarrassing inconsistency. This essay argues both readings miss what actually makes a public reversal creditable.

Julius Achebe Brookstone International School, Lagos

A public figure who reverses a previously stated position faces a peculiar and asymmetric reception: if the reversal moves toward a view their existing critics already held, it is celebrated as intellectual honesty and growth; if it moves away from a view associated with a specific political or social identity, it is frequently punished as betrayal or evidence the original position was never sincerely held. This essay asks what would make a public reversal genuinely creditable, independent of which direction it happens to move in, and argues the answer has almost nothing to do with the content of the new position and almost everything to do with the visible structure of the reasoning that produced the change.

The temptation, from a purely epistemic standpoint, is to say that any reversal driven by genuinely new evidence or a genuinely better argument is creditable, full stop, regardless of its political valence, since responsiveness to evidence is exactly what a good believer should exhibit. This is true as a description of ideal epistemic virtue, but it does not by itself explain why observers are right to be more suspicious of some public reversals than others, because the observer, unlike the person reversing, typically cannot directly verify whether the reversal was actually driven by evidence or was instead driven by

a change in the reverser's incentives — a shift in audience, funding, political alliance — dressed up, after the fact, in the language of following the evidence.

This is where the visible structure of the reasoning becomes the thing that actually matters, because it is often the only thing an outside observer has access to. A reversal accompanied by a specific, checkable account of what new evidence or argument produced it, an acknowledgment of what the person previously got wrong and why their earlier reasoning failed on its own terms rather than merely asserting a new conclusion, and ideally a track record of similar honest reversals in directions that did not happen to align with any change in the person's incentives, gives observers real, checkable grounds for treating the reversal as genuine intellectual movement rather than incentive-tracking with an epistemic gloss. A reversal that arrives without this structure — a bare assertion of the new view, timed suspiciously close to a change in the person's audience or funding, with no specific account of what evidence did the work — is reasonably treated with more suspicion, not because reversal itself is suspect, but because the observer has been given nothing to distinguish it from the incentive-tracking alternative.

This explains the asymmetric reception pattern without needing to appeal to tribal loyalty as the whole explanation, though tribal loyalty surely plays some role too. A reversal that moves toward a person's existing audience's preferences is, all else equal, more likely to actually be incentive-tracking, precisely because the incentive to make that particular move was present and available; a reversal that moves away from a person's existing audience's preferences is less likely to be incentive-tracking, because it works against the obvious incentive, which is part of why it reads, correctly on average, as more credible evidence of genuine reasoning. This is not proof in any individual case — people do sometimes have principled reasons to

reverse toward their own side's preferences, and people do sometimes cynically reverse away from their side for other, non-epistemic advantages — but it explains why the asymmetric prior is not simply irrational tribalism; it is a reasonable, if fallible, heuristic about which direction of reversal is more likely to have overcome, rather than been produced by, incentive.

The practical upshot for anyone changing their mind publicly and wanting to be believed is straightforward, even if it is rarely followed: show the specific reasoning, name what you got wrong the first time and why, and do this whether or not the new position happens to be convenient. The practical upshot for observers is to resist both automatic postures — neither crediting every reversal as growth nor dismissing every reversal as betrayal — and instead ask the only question that is actually answerable from the outside: has this person shown their work, in a way that would let me tell the difference between following an argument and following an audience?

POLITICAL PHILOSOPHY

Ignorance as a Social Achievement

Some ignorance is passive, a mere absence of information. This essay is about the other kind: ignorance that had to be actively built, maintained, and defended against correction.

Amara Diallo École Internationale de Dakar

Most everyday ignorance is unremarkable: there are simply more facts in the world than any person has time to learn, and not knowing most of them requires no explanation beyond finite attention. This essay is about a different and more troubling category, one the historian of science Robert Proctor named agnotology: ignorance that is not a passive gap but an actively constructed and maintained social product, requiring ongoing effort to produce and sustain, in the way a hole in the ground requires ongoing digging rather than simply being left when nobody bothers to fill it in.[1]

The clearest documented cases come from industries whose products caused harm that scientific evidence would otherwise have made widely known well before it did: tobacco companies' internal documents, later exposed through litigation, show a deliberate strategy not of denying the evidence of harm outright — an approach vulnerable to simple refutation — but of manufacturing the appearance of ongoing scientific controversy where, internally, none seriously existed, funding a steady supply of alternative studies, professional skeptics, and media appearances specifically designed to make the actual, settled state of the evidence look, to a public that

could not directly evaluate the underlying science, like an unresolved debate among experts. This is ignorance as an active production, requiring sustained investment, coordination, and effort — considerably more effort, in fact, than simply letting the public learn the truth would have required.

What makes this a distinctively political phenomenon, rather than merely an unusually elaborate form of lying, is that its target is not primarily belief in a specific false proposition but confidence in the process by which true beliefs are normally arrived at. A person told a specific lie can, in principle, be corrected by a specific piece of counter-evidence. A public made to believe that a topic is subject to genuine, ongoing scientific controversy — when among people with direct access to the evidence there is none — has been given something harder to correct, because any single piece of counter-evidence can be absorbed into the manufactured controversy as "one side's view," rather than registering as the decisive correction it would be if the public's picture of the state of expert agreement were accurate. This is why agnotology campaigns specifically target trust in expert consensus itself, rather than simply denying particular facts: undermining confidence in the process is more durable, and more resistant to any single correction, than lying about any single conclusion the process has reached.

This has a specific and uncomfortable implication for how citizens ought to allocate epistemic trust in domains known to have been targets of this kind of campaign. The ordinary epistemic advice — "consider both sides, weigh the arguments yourself" — is exactly the vulnerability agnotology exploits, because it treats the appearance of two sides as itself evidence that two genuinely comparable epistemic positions exist, when the entire point of a manufactured-controversy campaign is to produce that appearance without the underlying substance. A more resistant heuristic, harder to manipulate though

not immune to it, is to ask not "what are the arguments on each side" but "who has direct, verifiable financial or institutional interest in a specific answer to this question, and does the pattern of disagreement track expertise or track interest" — a question that would have flagged the tobacco campaigns' manufactured experts far more reliably than evaluating their arguments' surface plausibility ever could, since the arguments were specifically designed, by people with genuine scientific training, to be surface-plausible to a lay audience.

None of this licenses reflexive distrust of every apparent scientific disagreement, which would trade one failure mode for another — genuine scientific controversy is common and healthy, and dismissing it as manufactured whenever it is inconvenient is itself a form of the same manipulation in reverse. The point is narrower: ignorance is sometimes a fact about what a population has not yet learned, and sometimes a fact about what has been actively, expensively, and skillfully built to prevent them from learning, and telling these apart requires looking past the arguments themselves, which manufactured controversy is specifically designed to make look legitimate, toward the much harder to fake question of who benefits from the argument continuing.

NOTES

1. Robert N. Proctor and Londa Schiebinger, eds., *Agnotology: The Making and Unmaking of Ignorance* (2008).

EPISTEMOLOGY

Can You Know Something You Can’t Explain?

Expert intuition often outperforms explicit reasoning, yet cannot be articulated by the expert who has it. This essay asks whether this counts as knowledge, and what follows if it does.

Felix Brandt Gymnasium am Rhein, Cologne

Experienced clinicians frequently report a sense that something is wrong with a patient before any test confirms it, and this sense is often borne out; experienced chess players recognize a strong or weak position within seconds, well before they could articulate the specific calculation justifying the assessment; veteran firefighters, in cases documented by researcher Gary Klein, have reported ordering evacuations moments before a structural collapse, on the basis of a felt wrongness they could not, in the moment, put into words, but which tracked subtle cues — an unusual quietness, a strange heat pattern — their trained perception had registered without their explicit awareness catching up.[1] This essay asks whether these cases count as knowledge, given that the standard analysis of knowledge as justified true belief seems to require an articulable justification the expert, by hypothesis, cannot provide.

The traditional worry is that without an articulable reason, we cannot distinguish a genuine, reliable intuition from a lucky guess that happened to be right, and justification is supposed to do exactly this distinguishing work — separating knowledge from mere true belief arrived at by chance. If the clinician cannot say why she suspects something is wrong, on what basis do we credit her suspicion as

knowledge rather than a coincidentally correct hunch, indistinguishable, from the inside, from all her other hunches that turned out wrong?

The reliabilist tradition in epistemology, associated with Alvin Goldman, offers a way through this that does not require the knower to be able to articulate their justification at all: what matters for knowledge, on this view, is not that the believer can state the reasons for their belief, but that the belief was produced by a process that is, as a matter of fact, reliably truth-conducive — a process that tends to produce true beliefs across the relevant range of cases, whether or not the believer has conscious access to how the process works.[2] On this view, the clinician's trained perceptual intuition, if it has in fact been calibrated by years of exposure to cases where the same subtle cues correlated reliably with the same underlying conditions, can constitute knowledge in exactly the sense a more explicit, articulable inference would, because reliability, not articulability, is what makes a true belief count as known rather than merely guessed.

This does not mean all confident, inarticulate intuitions should be trusted equally, and the reliabilist framework actually gives us a precise way to say which should and should not be: the question is not whether the intuition feels confident — confidence is notoriously uncorrelated with actual reliability in many domains — but whether the specific domain and the specific person's training history are of the kind known to produce reliable calibration. Klein's own research on firefighters and chess masters identifies the conditions under which expert intuition reliably outperforms explicit reasoning: a domain with consistent, learnable statistical regularities, and extensive practice with fast, clear feedback on whether each judgment was right or wrong, which is exactly the training regime that calibrates perceptual pattern recognition to track real regularities rather than spurious ones. Domains lacking this — highly

unpredictable systems, or professional contexts where feedback on judgment quality is slow, ambiguous, or absent — do not produce this kind of reliable expert intuition no matter how confident the practitioner feels, which is exactly why we should trust an experienced firefighter's inarticulate sense that a building is about to collapse far more than an inarticulate hunch from a stock picker in a notoriously feedback-poor, low-predictability domain, even though both hunches present, from the inside, with the identical phenomenology of confident, unexplainable knowing.

This gives experts and non-experts alike a genuinely useful, checkable question to ask before trusting an inarticulate intuition, in place of the unhelpful "trust your gut" that gets applied indiscriminately: does this domain have the structure — real regularities, fast and clear feedback, extensive deliberate practice — known to produce reliable calibration, or does it merely have the surface trappings of expertise (years of experience, professional confidence) without the specific training conditions that make intuition track truth rather than merely feel like it does. The inarticulate feeling of knowing is the same in both cases. Whether it constitutes actual knowledge is not a fact about the feeling at all, but a fact about the process, largely invisible to the person having the feeling, that produced it.

NOTES

1. Gary Klein, *Sources of Power: How People Make Decisions* (1998).
2. Alvin Goldman, "What Is Justified Belief?" in *Justification and Knowledge*, ed. George Pappas (1979).

ETHICS

The Duty to Believe Well

If belief is not directly voluntary, can we still be under a genuine moral duty to believe responsibly? This essay defends a qualified yes.

Saoirse Byrne Ashgrove Academy, Cork

William Kingdon Clifford's famous claim that "it is wrong always, everywhere, and for anyone, to believe anything upon insufficient evidence" is usually introduced in epistemology courses only to be immediately challenged: if belief is not under direct voluntary control, as the impossibility of doxastic voluntarism suggests, how can there be a moral duty governing it at all, given that "ought implies can" — the widely accepted principle that a genuine obligation must be something the obligated party is actually capable of fulfilling or violating through their own agency?[1] This essay defends a qualified version of Clifford's claim that survives this objection, by locating the duty where genuine voluntary control actually exists.

The objection is forceful against a naive reading of Clifford: if I cannot simply choose to believe or disbelieve a given proposition at will, then a duty phrased as "do not believe X" looks incoherent, addressing a directive at a capacity I do not have, in the same way a duty to "digest your food more slowly through an act of will" would be incoherent, addressing a directive at a physiological process outside voluntary control. If this is what Clifford's principle requires, the objection succeeds, and the principle should be abandoned or radically reinterpreted.

But recall the earlier point about indirect doxastic control: while I cannot believe at will directly, I have extensive, entirely voluntary control over the inputs that shape what I come to believe — which evidence I actively seek out and which I avoid, how much effort I put into checking a convenient claim against the same scrutiny I would apply to an inconvenient one, whether I actually engage with the strongest available version of an opposing argument or content myself with its weakest, most easily dismissed version. Clifford's duty, properly reconstructed, is not a duty over the belief itself but over this entire upstream process of evidence-gathering and evidence-evaluation, all of which is genuinely, ordinarily voluntary in exactly the way the belief's final output is not. "Ought implies can" is satisfied, because the relevant "can" attaches to the input-gathering process the duty is actually about, not to the belief-formation process the objection mistakenly assumes it targets.

This reconstruction changes what culpable believing actually looks like, and importantly, makes it something we can identify without needing direct access to anyone's inner cognitive process. A person who ends up with a false, harmful belief is not thereby blameworthy — the belief itself was never under their direct control, and people of perfectly good epistemic conduct sometimes still end up wrong, since even careful evidence-gathering does not guarantee a true conclusion. What can make a person blameworthy is a demonstrable pattern in the upstream process: did they seek out relevant, available evidence before forming a confident view, or form the view first and treat evidence-seeking as an afterthought, if it happened at all; did they subject convenient and inconvenient claims to comparable scrutiny, or apply a double standard that let convenient claims pass with a glance while inconvenient ones faced an unreasonable burden of proof; did they engage the strongest form of opposing arguments, or content themselves with a weak version convenient to dismiss. These are all observable, or at least reconstructable, facts about

conduct, not facts about an inner state no one, including the believer, has reliable introspective access to — which makes this version of Clifford's duty something we can actually apply, to others and to ourselves, without pretending we can inspect the inaccessible moment of belief-formation itself.

This leaves genuine, non-culpable disagreement fully intact as a real and common phenomenon: two people can each conduct the upstream process responsibly — seeking relevant evidence, applying even-handed scrutiny, engaging serious opposing arguments — and still arrive at different conclusions, because the evidence available to each was different, or because reasonable interpretation of shared evidence genuinely diverges. Clifford's duty, correctly reconstructed, was never meant to guarantee convergence on truth; it was meant to locate moral responsibility precisely at the boundary of what an epistemic agent can actually control, which turns out not to be their beliefs, but the far more mundane, far more voluntary, and far more neglected question of how they went about forming them.

NOTES

1. William Kingdon Clifford, "The Ethics of Belief" (1877).

PHILOSOPHY OF SCIENCE

What Maps Leave Out on Purpose

A map that included everything would be useless. This essay argues scientific models are maps in exactly this sense, and asks what follows for how we should judge a model's honesty.

Ren Takahashi Seiwa Academy, Kyoto

A map that attempted to represent every detail of the terrain it depicts — every rock, every blade of grass, every momentary puddle — would not be a more accurate or more useful map; taken to its logical endpoint, as Jorge Luis Borges's famous one-paragraph fable of the empire whose cartographers made a map the exact size of the empire itself illustrates, it would be a useless, unfoldable duplicate of the territory, containing all the same information as the terrain and therefore none of the compression that makes a map worth consulting in the first place.[1] This essay argues that scientific models are maps in exactly this sense — useful only because of what they deliberately leave out — and that this reframes what it means to criticize a model as unrealistic.

A common complaint against simplified scientific models — economic models that assume perfectly rational agents, physical models that assume frictionless planes, epidemiological models that assume homogeneous mixing within a population — is that these assumptions are false, and a model built on false assumptions cannot be trusted. This complaint, taken generally, proves too much, because every useful model in every field makes some simplifying assumption that is, strictly, false of the actual world; the question was never

whether a model's simplifications are true (they never are, by design) but whether the specific simplifications made are the right ones for the specific purpose the model is being used for.

This is the same design question a cartographer faces, and thinking it through in cartographic terms clarifies what "the right simplification" actually means. A subway map that represents stations as evenly spaced points along straight or gently curved lines, discarding true geographic distance and direction entirely, is a spectacularly unrealistic map by the standard of geographic accuracy — and is exactly the right map for a passenger who needs to know which line to transfer to, since accurate topology (what connects to what) is what that purpose requires, and accurate geography is not merely unnecessary for that purpose but would actively clutter the representation with irrelevant precision. The same map would be a disastrously bad map for someone trying to walk between two stations by the most direct surface route, because for that purpose, exactly the geographic information the subway map discarded is exactly what is needed.

Applied to scientific models, this reframes the right question to ask about a simplifying assumption from "is this assumption true" — which it almost never is, and asking this question invites either naive rejection of every useful model or a shrugging "all models are wrong" that abandons the possibility of discrimination entirely — to "does this specific simplification preserve what this specific use of the model actually needs to track, and does it discard only what that use does not need." A frictionless-plane model in introductory mechanics is not lying about the world; it is doing exactly what the subway map does, discarding a real feature (friction) that is irrelevant to the specific conceptual point being illustrated (the relationship between force, mass, and acceleration in the idealized case), while remaining

honest, to a reader who understands the genre, about the fact that this feature has been discarded rather than smuggling it in as though it had been accounted for.

This gives a sharper standard for judging a model's honesty than either "is it realistic" or "all models are simplifications so none can be judged": a model is used dishonestly not when it simplifies — all useful models must — but when it is deployed for a purpose that actually depends on precisely the feature the model discarded, without acknowledging the mismatch. An economic model that assumes perfectly rational agents is a perfectly honest and useful tool for illustrating certain baseline dynamics of exchange, exactly as a subway map is honest about topology; the same model becomes dishonest the moment it is deployed to make confident predictions about a domain — a financial crisis driven substantially by panic, herd behavior, and other systematically irrational dynamics — where the very thing it assumed away is the thing doing most of the causal work. The map was never wrong to leave out the terrain's every contour. It becomes a lie only when someone tries to use it to navigate exactly the contour it was built, quite reasonably, to ignore.

NOTES

1. Jorge Luis Borges, "On Exactitude in Science" (1946), a one-paragraph fable often cited in philosophy of science discussions of modeling and idealization.

MASTHEAD

The Ethics of Knowing — Issue 03

EDITOR-IN-CHIEF

Dominic Alvares

St. Xavier's Collegiate School, Kolkata

DEPUTY EDITOR, EPISTEMOLOGY

Felix Brandt

Gymnasium am Rhein, Cologne

FACULTY ADVISOR, ETHICS

Prof. Naomi Katz

University of Toronto

REVIEWING EDITOR, EPISTEMOLOGY

Dr. Amit Sharma

University of Oxford