Original study • DOI: 10.2478/biocosmos-2025-0001 • BIOCOSMOS • 1 • 2025 • 12-17

BioCosmos

WHAT IS LIFE? AN OPERATIONAL DEFINITION

David S. Oderberg*

Department of Philosophy, University of Reading, Reading RG6 6EL, UK

Abstract

One way of defining life is via a real definition, which gives the essence of life. Another approach is an operational definition, which shows how living things can be tested or measured in a way that is distinctive of the biological. Although I give a real definition elsewhere, in this paper I provide an operational definition, echoing Canguilhem's dictum that life is what is capable of making mistakes. Biological mistakes are central to the behaviour of organisms, their parts and sub-systems, and the collections to which they belong. I provide an informal definition of a biological mistake. I contrast mistakes with mere failures and malfunctions. Although closely related phenomena, each is distinct. After giving some brief examples of mistake-making and how it can be tested, I reply to some objections to the very idea of a biological mistake.

Keywords

life • mistakes • normativity • reductionism

There are two ways to approach that most difficult and fundamental of questions: what is life? One is to give a *real* definition – a statement of the *essence* of life (assuming it has one). A real definition tells you what an *object* is. It doesn't tell you primarily (or even at all!) about the meanings of words, facts about grammar, or information about what is in someone's head. The real definition of *gold*, for instance, is that it is a metal whose atomic number is 79. The real definition of a *fish* is that it is a water-dwelling vertebrate with gills at maturity. These definitions give you the very *essence* of the thing being defined – what it is that separates it from every other kind of thing in the universe.

The other way of approaching the question 'what is life?' is to give an *operational* definition: one that, speaking loosely, makes life *measurable* and *testable* – that characterises it in such a way as to give the experimental biologist something work on. These are not mutually exclusive approaches, of course, and they also interpenetrate at the more abstract level, but they are methodologically quite different.

In previous work I have articulated what I contend to be the real definition of life: a living being is that which has the power of *immanent causation* – causation that originates with an agent and terminates in that agent for the sake of its self-perfection (1, 2). It is not merely teleology that defines life, but the kind of purposive behaviour directed at the well-being of the organism – whether or not the behaviour be characterised by self-consciousness, language, free will, responsibility, knowledge of purpose, among other properties of particular kinds of living thing. Immanent causation is to be distinguished from ordinary, common-or-garden *efficient* causation, which is, to co-opt the old saying, just 'one damn thing after another'. With immanent causation, a living being acts not just *on* itself and does not merely do things *to* itself, as when a volcano erupts or a wave crashes on the shore. A living being acts on itself, and does things to itself, *for* itself. To act *for* itself means that it acts to keep itself in a good state, functioning well, alive, healthy, integrated, and so on. No inorganic being, no matter how large or complex, does this. To the objection, 'What about computers that maintain their operating systems in good health?' and so on, the reply is that these are *artefacts* – man-made objects with no intrinsic teleology of their own. Their ends and objectives are programmed *into* them, so what they appear to do for *themselves* is really done *for us*. If there were no 'us', there would be no self-maintenance by a computer or any artefact, because there would not even *be* artefacts – only inorganic things with no intrinsic purposive behaviour of their own. Their derivative 'life' – what they do based on what we *direct* them to do – would vanish without their users and creators.

There is much that can and has been said about such a definition, but here I focus on the fact that it is not immediately

---

Corresponding author e-mail: d.s.oderberg@reading.ac.uk obvious how to ‘cash out’ such a definition in experimental terms. What kinds of novel, testable hypotheses might be generated by such a high-level metaphysical definition? We can, however, *operationalise* this way of understanding life by developing an insightful remark by the French philosopher and physician Georges Canguilhem (3): ‘life is what is capable of error’. Living things *make mistakes*; they get things wrong. A dog may forget where he buried his bone. A frog may be a millisecond too late darting its tongue at an insect. A fish takes the bait. Broody domestic hens are notorious for trying to hatch golf balls and other vaguely egg-resembling objects. Animals are fooled by camouflage, traps, prey that hide. The living world is full of tricks – deception, subterfuge, misjudgment, mistiming, miscommunication. It is astonishing how much an organism – from the very large to the microscopic – has to *get right*, how many mistakes it must *avoid* – just to approach the end of its natural lifespan and propagate its own kind. That the environment is inherently dangerous and often unstable – certainly quite unpredictable – renders mistake-making a certainty and its avoidance a necessity.

Not only whole organisms, but also collectives of organisms and even *parts* of organisms make mistakes. Flocks of birds fly into skyscrapers. Antibodies are fooled by pathogens – the classic example is antigenic mimicry, which is thought to be at the root of many autoimmune diseases; it is not only the self antigens that fool the antibodies in many cases but the non-self antigens – such as viruses – themselves (4, 5). Blood platelets – essential for the formation of clots – are activated by exposure to collagen released by endothelial injury. However, in an individual with a certain genetic abnormality, they will not respond correctly – as they are supposed to – and the individual will suffer potentially fatal blood loss (6).

It is arguably the case that if not actual mistake-making then at least the *potential* for making mistakes is universal across biological systems. The connection with the real definition of life offered above is clear enough: if it is definitive of organisms (by which I mean, for convenience, whole organisms, pluralities, and also parts and sub-systems of organisms) that they act for their well-being, then we should expect them not always to get things right. Given the already-noted instability of nature – both external to the organism and in its own internal constitution – we should expect mistakes to be found everywhere and at all times. Perhaps we can at least conceive of an organism in a hyper-stable environment, with a simple life cycle, constitutively as simple as biology allows, and thereby incapable of making a mistake – at least in its actual world. Such an organism is likely not a denizen of our world, and if it were it would not take much imagination to posit a close possible world in which it could and did make mistakes.

Mistake-making is then, as it were, a concrete entry on the debit side of life’s accounting. It comes with the territory if something is alive. It indicates the fundamental *normativity* of life. Perhaps we could even give a definition of life halfway between the highly abstract and the highly concrete: life is that which is subject to *norms*. Note that we are, all the while, talking about *material* life – the life of bodily substances, whether or not – as in the human case, I believe – there is an immaterial element as well. On an extended view, we could even say that purely spiritual beings – angels, for example – are subject to norms of behaviour and are potential mistake-makers. God, by contrast, although purely spiritual is subject to no norms, does not act to perfect Himself since already perfect, and cannot err.

Normativity has very little place in current philosophy of biology. It is often equated, wrongly, with ‘values’, ‘oughts’, ‘shoulds’, ‘prescriptions’, ‘commands’ (7). It is thought to involve the ‘spooky metaphysics’ of teleology (8). It is considered to be a matter of perspective, or context, or theoretical interest (9). It is hard not to detect a certain cant (certainly not Kant!) in such pronouncements, as though it were philosophically unproblematic literally to collapse the distinction between health and disease, the normal and the pathological, welfare and what might be called ‘ilfare’. If anything in the natural world has a claim to intuitive obviousness, it is that there are ways in which things go *well* or *badly* for an organism. We know it with certainty in our own case, yet why should our case be special?

The point can be brought out vividly by the simple observation that electrons do *not* make mistakes, neither do protons, or carbon atoms, or lumps of gold or grains of sand, or continents or rivers. The only mistakes in physics are those made *by* physicists, found in textbooks or research articles. There are mistaken *theories* or *hypotheses*. Instruments may give erroneous readings. But the proper *objects* of physics do not makes mistakes. No proton ever went the *wrong* way in a particle accelerator, except relative to some theoretical perspective or expectation. By clear contrast, however, mice do make mistakes – it’s the second mouse that gets the cheese, so the proverb goes – and so do ants, grasshoppers, antelope, and we humans.

How should we define biological mistakes? Leaving aside technical details (10), a biological mistake is a kind of behaviour that departs from a standard of correctness for the organism in its environment. The way it so departs is by threatening the organism’s ability to ‘get on’, or ‘act effectively’, in its environment. There is no mathematical precision when evaluating what counts as a threat. We know that if a deer gets caught in the headlights, then if the developing causal connection between the vehicle and the deer is not broken

– say, by the deer's coming to its senses and bounding over the fence, or by the car's braking hard – then the deer will end up injured or dead. We could say the same if it were fifty metres from the vehicle yet heading inexorably in that direction. There is no way to set a cut-off point beyond which a connection is too remote to count, at least narrowly and relatively insignificantly, as a mistake. More work should be done in this area, but for now I will say only that what this teaches us is that mistakes shade into general *risk*. Life is full of risk: you risk your life, to some degree, whenever you get out of bed in the morning. But it is not a mistake simply to live. It is, however, a mistake to play on your mobile phone while crossing a busy road.

Mistakes are not the only marker of biological normativity. There are also malfunctions and what I call 'mere failures'. Again, the 'functions debate' in philosophy of biology is a minefield of its own (11). For present purposes, a malfunction should be thought of as any systemic breakdown in an organism that threatens welfare – sickness being the most obvious. Sickness is a departure from correctness – being healthy. But it need not involve a mistake: being invaded by a parasite is not of itself a mistake, but drinking parasite-infested water is. Again, a mere failure also departs from correctness: it is what simply *happens* to an organism, not something the organism *does* to threaten its well-being. Being hit by lightning, or by a car, or rained on by radioactive particles, are mere failures that need not involve doing anything wrong. They *might* – say if you went outside during a thunderstorm against all advice – but mistake-making is not entailed either by malfunctions or mere failures.

Interestingly, the converse is also true: mistake-making does not entail either malfunction or mere failure. The latter should be evident, but even the former is true. Our initial intuition might be that making a mistake must involve a malfunction. But consider that a hen trying to hatch a golf ball, or a fish taking bait, are not malfunctioning (in the usual run of things): they can be in the best shape possible for their kind, and yet still get things wrong due to their intrinsic limitations. A hen does not have the discriminatory capacity to distinguish an egg from something looking quite a bit like one, nor a fish the ability to distinguish bait from real food. That's just how it is with these kinds of organism. It would be a waste of a poultry farmer's time to try to *teach* hens not to sit on golf balls; in fact this mistake is *exploited* by poultry farmers to manage broody hens. Most fish cannot be trained to avoid bait either, much to the relief of fishermen the world over. These latter kinds of mistake I term *unavoidable*. They are an inevitable effect of the natural constitution of certain kinds of organism. It is likely, *a priori*, that *every* species is liable to unavoidable mistakes, given their finiteness; much of the research into heuristics and biases is premised on such a fact about us humans (12). Experimentally speaking, it would be a fascinating task to map the limits of unavoidability in a given species – just one aspect of treating mistake theory as a way of operationalising organic teleology.

The theory of biological mistakes is not supposed to supplant or compete with any other productive biological frameworks we already have. Rather, it is a different way of looking at life with a view to developing novel and testable hypotheses. Biologists investigate mistakes all the time: indeed the language of biological mistakes is far less problematic for biologists, in my experience, than it is for philosophers who are used to thinking of mistakes from the inside, as a human-centred phenomenon. What mistake theory does is seek to organise the various concepts and phenomena involved in thinking about and observing mistakes, with a view to more systematic investigation.

Along with testing for unavoidable versus avoidable mistakes, another avenue for research is testing for mistake prevention, minimisation, and correction mechanisms. For a given kind *K*, what capacities do *K*s have for keeping themselves out of the way of certain kinds of mistake-making, or for keeping mistakes to a minimum, or correcting mistakes that have been made? Correction could involve getting back onto the right path there and then, or learning from the mistake in such a way as to avoid the same one in future. It is highly likely that if *K*s are mistake-prone but unable to prevent, avoid, correct or minimise those mistakes (at least some combination thereof), they are already extinct or on the way to becoming so. Again, one should expect *a priori* that no such kind does exist or ever has existed – otherwise how could it even have managed more than a few seconds of existence in the first place?

As I have emphasised, mistake theory should enable the generation of novel, testable hypotheses. For example, dopamine neurons have been proposed to perform an *evaluative* function with respect to birdsong – in particular, that of zebra finches (13, 14). Dopamine spikes correlate significantly with fluctuations in a combination of measures such as song pitch, frequency, and Wiener entropy (a measure of where the sound stands between a pure tone and white noise), such that higher spikes are associated with closeness of the song to the one it learned from its father and lower spikes with greater fluctuations away from that template. Dopamine seems to act like a mistake detection and correction system, maintaining the song's relative fidelity to what the finch learned. From this we can hypothesise that not only is the bird itself capable of mistake-making, but the dopamine neurons *themselves* might be capable of performing their evaluative function mistakenly. This latter possibility will need further investigation: does the dompaninergic system monitor or regulate the song evaluation function of dopamine neurons, and if so how? Further, how is the standard of correctness for song production represented by the zebra finch, especially since each one has a unique correct song that it is taught? Given that birdsong is a form of communication, mistake theory focuses on whether and how a bird may *get it wrong* in, for instance, attracting a mate through song. Interestingly, the dopaminergic error signalling that operates when the bird practises alone is turned off in the presence of a potential mate and retuned to feedback from the other bird (15). It would be informative to investigate whether this retuning was itself mistake-prone.

Since relative fidelity is at stake, not mathematically exact correspondence, we might hypothesise that the bird, via its dopaminergic sub-system, evaluates the correctness of the song *as correct* rather than as containing a specific number of pitches in a certain order, for example. But, one might object, how could a bird evaluate something as correct without the concept of correctness? In reply, we need to know what it means for an organism to possess a concept: perhaps there are kinds of practical mastery, for example, which constitute a kind of concept possession but without the abilities associated with reason, language, and self-awareness. If the bird is able to keep its song faithful to the original, maybe that is enough for us to assign to it a kind of mastery of the concept of a *correct birdsong*, for limited biological purposes. Still, we do not have to go even that far. It is doubtful a bird has the concept of *food* any more than that it has the concept of *correctness.* And yet birds are good at finding food and avoiding eating non-food; similarly, perhaps, for song: the zebra finch can learn its song, can sing, and can keep its song correct – for the most part. That requires an ability to *assess* the song as right or wrong, albeit instinctively. Perhaps the bird has some kind of primitive *aesthetic* standard; in other words, the ability to tell whether the song *sounds pleasant* (for mate-finding) or *good enough* (for an alarm call), and so on. One could test such a hypothesis, crudely speaking, by investigating whether the bird is counting or measuring units of song as opposed to evaluating the song as a whole pattern. (Needless to say, such testing is extremely difficult to do, but progress is being made all the time). If it can be established that it evaluates a song for rightness or wrongness, then it is likely capable of making a bad judgment, e.g. evaluating a song as pleasant enough to attract a mate when in fact it departs too far from the template to facilitate reproductive success. Evidence for a mistaken judgment would be found in whether the song did successfully attract a mate; if not, we could investigate whether and how quickly – perhaps too quickly for any calculations to be made – the mistake-prone bird was able to improve its song the next time around.

There are of course various objections that can be levelled at the very idea of a biological mistake (10). I want to focus on two here. One is that since mistakes presume the reality of biological normativity, and biological normativity is reducible to physics and chemistry, there are no real biological mistakes. I have already pointed out that denying the reality of normativity (I will omit ‘biological’ from now on for convenience) means denying evident distinctions between health and disease, welfare and ilfare, and related phenomena. We can, however, say more (16). Consider a Laplacean all-knowing demon who comprehended all the physical and chemical facts in a given situation where one might postulate a biological mistake. Suppose there to be three physico-chemical pathways, as we might call them, traced by an organism – itself (ex hypothesi) nothing but a skinful of physics and chemistry, to put it metaphorically. The organism follows path A, which the mistake theorist points out, in agreement with simple observation, is a mistake – say, mislocating a prey. It then backtracks and follows path B, which the mistake theorist, again conforming to observation, identifies as wandering around doing nothing in particular – neither mistaken nor correct relative to the objective of *finding prey*. It then backtracks and follows path C, which is identified as finding prey – and catching and eating it. From the normative perspective, these pathways are quite different in their meaning – not just for us as observers, but for the organism itself. One is right, one is wrong, one is neutral. One promotes survival, one threatens it (a small amount, by delaying the hungry organism’s getting its lunch), and one is ‘neither here nor there’ – normatively *indifferent*, as we might say.

The problem for the Laplacean demon – who is also a reductionist about normativity – is how *he* can make this distinction. He has all the physico-chemical facts at his disposal, yet physico-chemically speaking pathways A, B, and C are all *on a par*. None of them have any properties that mark them out as mistaken, correct, or indifferent. They are, as we can call them, *mere physico-chemical variations*. It is, in principle, impossible for the demon to read off the normative properties from the physico-chemical ones. In which case he is faced with a choice. He can deny normativity altogether, in eliminativist fashion: there really *is* no distinction between sickness and health, doing well in life and doing badly, flourishing and suffering, and so on. This, as suggested earlier, seems wholly implausible on its face. Or else, the demon must concede that since normativity cannot be read off the physics and chemistry on their own, it must be an *addition of being* (to use a term from David Armstrong) – even if the normativity *supervenes* on the physical and the chemical. (Or maybe it is emergent in some sense we cannot explore here.) This is a familiar non-reductionist line of argument but no less potent for that.

A second objection concerns the relation of mistakes to the environment. On mistake theory there is no such thing as a pure or absolute mistake. The job of every organism is to get on well in its environment – to 'make a living', as Mark Okrent (17) puts it. This applies to humans as well as to other organisms. To drink boiling water is a mistake – even if, suppose fancifully, your stomach miraculously shrugs off the assault and you are hailed by the crowds as a superman. We need to qualify, however. Assuming you had no idea this would happen on drinking the boiling water, what you did was a mistake relative to the environment you were in – your ordinary daily surroundings, let us suppose. You got off lucky, to be sure, but you still acted against the norms of health and survival. But if you *knew* you had this miraculously strong stomach, then if, suppose further, you were demonstrating your prowess to an audience at a talent show, you would *not* have made a mistake in that environment. Note that we are not considering *morality* or *responsibility*, merely what does or does not serve the person's welfare. We could substitute a bird and a mating display to make the same point. A bird that shows off colourful plumage – with no mate in sight but a predator standing by and alerted to the bird's presence – has made a mistake, whether it knows there is a predator and no mate or not. Suppose – again somewhat fancifully – that the bird is attacked mid-display by the predator but fights it off, emerging battered but in generally good shape. A potential mate then comes along, picks up cues as to the first bird's health and fitness, and happy reproductive success ensues. Again, we should say that the bird made a mistake relative to the environment in which a predator was standing by, but not relative to the predator-free environment in which a mate was standing by. These are just facts about what the bird has done – normatively laden facts about behaviour that does or does not threaten effective action in a given environment.

We can, of course, add further facts into the mix in order to make subtler evaluations; this is just what mistake theory is designed to encourage. If the bird *knew* it could fight off the nearby predator, exploiting its success in order to attract a mate, then we could plausibly assess the situation as one in which no mistake was made. We should say that there was a single environment, containing both predator and potential mate, in which the bird implemented a strategy for mate attraction, with ensuing success. In other words, what counts as an environment for mistake-theoretic purposes can involve what an organism *knows*, what it *plans*, and what its capacities are. An environment is just a spatio-temporal region in which an organism carries on its life cycle. The region is defined at least partly by the biologically relevant causal relations obtaining, where the organism is directly or indirectly affected by other things and/or capable of directly or indirectly affecting other things. A sceptic might say: 'All organisms carry on their life cycles in the universe as a whole, so this is the only real environment in which mistakes can be made, especially given the "butterfly effect" and the causal connectedness of all things.' It is hard to know what to say that would convince the sceptic otherwise, nor is it clear that it is worth saying anything to such a sceptic. Given that mistake theory is about operationalising more abstract ideas about purpose and normativity, it is good enough to defer to biologists: they do not care about the boundaries of the universe when assessing whether an organism is doing well or not. Nor do they consider the earth as a whole, or even the vast bulk of it. They are interested only in whether an organism – and its conspecifics, within a broader evolutionary context – is adapted to its niche, which is the specialised, relatively narrow spatio-temporal region in which its lives, maintains itself, and reproduces its kind. If that is good enough for biologists, it is good enough for mistake theorists.

The more stable a region is, the less mistake-making is possible. With change comes greater potential to get things wrong. That is why organisms generally prefer stability and predictability, at least when it comes to staying alive and healthy. The mistake theorist, from the experimental perspective, can choose how broad a region to consider when evaluating behaviour for actual or potential mistake-making. A mistake in one environment may lead to success in a larger environment – thus making it interesting to observe an organism's strategies over time for coping with different challenges coming from novel causal interactions. There is a difference between, say, a fish that accidentally swims upstream due to a sudden change in the current, thereby exhausting itself and dying of fatigue, and a salmon swimming upstream so as to reproduce. The first makes a mistake, the second doesn't, and we make the distinction in terms of the different environments. They are two radically different, albeit superficially similar, kinds of behaviour. But what if the salmon dies of exhaustion itself before managing to reproduce? The evaluation will depend on the detail. For example, if it was overwhelmed by a current stronger than it was capable of overcoming, it would have been subject to a mere failure (see above). If it was diseased and never capable of achieving its objective, it would have been subject to a malfunction (again see above).

I submit that although there are genuinely important questions about the relation of action to the environment, and hence about the individuation of mistakes, we should put aside more factitious concerns about whether mistakes are possible at all given the infinite divisibility of spatio-temporal regions, or the causal connectedness of the entire universe, or the potentially huge variability in mistake-making from one period in an organism's life cycle to the next. What mistake theory enables us to do is to organise many disparate ideas about action, success and failure, welfare, avoidance, prevention, learning, among others, into a general framework for interrogating living systems. In an age in which Big Data can overwhelm high-level theory, the investigation of mistakes enables biologists to stay close to the data while still entertaining the Big Ideas of teleology and normativity.

References

1. Oderberg DS. Synthetic life and the bruteness of immanent causation. In: Feser E. (ed.) *Aristotle on method and metaphysics*. Basingstoke: Palgrave Macmillan; 2013. p.206–235.
2. Oderberg DS. *Real essentialism*. London: Routledge; 2007.
3. Canguilhem G. *The normal and the pathological*. New York: Zone Books; 1991. p.22.
4. Wildner G. Antigenic mimicry – the key to autoimmunity in immune privileged organs. *Journal of Autoimmunity*. 2023;137: 102942. doi: 10.1016/j.jaut.2022.102942
5. Oldstone MBA. Molecular mimicry and immune-mediated diseases. *FASEB Journal*. 1998;12: 1255–1265. doi: 10.1096/fasebj.12.13.1255
6. Arthur JF, Dunkley S, Andrews RK. Platelet glycoprotein VI-related clinical defects. *British Journal of Haematology*. 2007;139: 363–372. doi: 10.1111/j.1365-2141.2007.06799.x
7. Garson J. *What biological functions are and why they matter*. Cambridge: Cambridge University Press; 2019. p.15.
8. Cummins R, Roth M. Traits have not evolved to function the way they do because of a past advantage. In: Ayala FJ, Arp R. (eds.) *Contemporary debates in philosophy of biology*. Oxford: Wiley-Blackwell; 2010. p.72–85, at 83. (The authors do not use the word 'spooky').
9. Searle JR. *The rediscovery of the mind*. Cambridge, MA: MIT Press; 2002. p.238–239.
10. Oderberg DS, Hill J, Bojak I, Gibbins JM, Austin C, Cinotti FM. Biological mistakes: what they are and what they mean for the experimental biologist. *British Journal for the Philosophy of Science*. 2023. doi: 10.1086/724444
11. Garson JA. *Critical overview of biological functions*. Dordrecht: Springer Cham; 2016.
12. Kahneman D, Slovic P, Tversky A. (eds.) *Judgment under uncertainty: heuristics and biases*. Cambridge: Cambridge University Press; 1982.
13. Duffy A, Latimer KW, Goldberg JH, Fairhall AL, Gadagkar V. Dopamine neurons evaluate natural fluctuations in performance quality. *Cell Reports*. 2022;38: 110574. doi: 10.1016/j.celrep.2022.110574
14. Gadagkar V, Puzerey PA, Chen R, Baird-Daniel E, Farhang AR, Goldberg JH. Dopamine neurons encode performance error in singing birds. *Science (New York, NY)*. 2016;354: 1278–1282. doi: 10.1126/science.aah6837
15. Roeser A, Gadagkar V, Das A, Puzerey PA, Kardon B, Goldberg JH. Dopaminergic error signals retune to social feedback during courtship. *Nature*. 2023;623: 375–380. doi: 10.1038/s41586-023-06580-w
16. Hill J, Oderberg DS, Bojak I, Gibbins JM, Austin C, Cinotti FM. *Mistakes and correctness: a new look at the reductionism debate in biology*. (forthcoming).
17. Okrent M. Nature and normativity: biology, teleology, and meaning. London: Routledge; 2018.