Teorema, 2025, 44(1), pp. 1-7
https://doi.org/10.30827/trif.34326

The place of choice and norms in a world of inevitable events

Jesús Zamora Bonilla

Universidad Nacional de Educación a Distancia

Abstract: Some aspects of Robert Sapolsky's arguments against free will in his book Determined are clarified. In the first place, it is argued that our choices are determined not in the sense that physical laws are always deterministic, but in the sense that, even if random events may happen, our choices cannot be but the ones they have been, given everything that did not depend upon us (including indeterministic events). In the second place, the determination of our choices is not equivalent to their not being conscious, deliberate choices, for the difference between these and other kinds of psychological or physiological events does not consists in an imaginary ontological openness, but just in following different (and more complex) neurological routes. Lastly, it is argued that the inevitability of our choices only precludes moral assessment according to some limited moral theories.

Keywords: *free will*, *determinism*, *inevitability*, *fatalism*, *moral relativism*.

Resumen: En este artículo se aclaran algunos aspectos de los argumentos de Robert Sapolsky contra el libre albedrío en su libro Determined. En primer lugar, se sostiene que nuestras elecciones están determinadas no en el sentido de que las leyes físicas sean siempre deterministas, sino en el sentido de que, incluso si ocurren eventos aleatorios, nuestras elecciones no pueden ser otras que las que han sido, dado todo aquello que no dependía de nosotros (incluidos los eventos indeterministas). En segundo lugar, que nuestras elecciones estén determinadas no equivale a que no sean elecciones conscientes y deliberadas, ya que la diferencia entre estas y otros tipos de eventos psicológicos o fisiológicos no consiste en una supuesta apertura ontológica imaginaria, sino simplemente en que siguen rutas neurológicas distintas (y más complejas). Por último, se argumenta que la inevitabilidad de nuestras elecciones solo impide la evaluación moral según algunas teorías morales limitadas.

Palabras clave: *libre albedrío*, *determinismo*, *inevitabilidad*, *fatalismo*, *relativismo moral*.

Received: 30 June 2025 **Accepted**: 4 July 2025 **Published**: 30 July 2025

Teorema. Revista Internacional de Filosofía
ISSN/ISSN-e: 1888-1254

* This work has been supported by Research Project PID2022-142120NB-100, financed by the Ministry of Economy and Competitiveness, Spanish Government.

The most frustrating thing of philosophical debates is that most often they seem incapable of making any of their participants change his or her mind on the debated point, and perhaps no topic is more representative of this predicament than that of free will. Analytic philosophy has historically tried to face this problem by insisting in the clarification of concepts as a fundamental step in the debate, but often this has only led to an almost malign proliferation of different possible senses for each term, creating a semantic forest that mostly helps to discourage future participants to make the effort of studying and learning all those Byzantine definitions, and inviting them, as a result, to enter de debate as if most of that work of conceptual gardening had never taken place, trusting again, instead, on their own prejudices. This sad introit has probably no other role than expressing my own pessimism on the possible efficacy of my subsequent arguments: though I will try (or will seem to be trying) to offer some clarifications on three points on which Sapolsky's rejection of free will has been (in my opinion) badly misunderstood, I have basically no faith on the chances these arguments may have of persuading anybody who happens to have contrary views. But, well, that's what one is paid for [*tongue in cheek*]. The three points I am referring are the following ones: in the first place, a confusion between 'determinism' (as a general thesis about nature) and the way in which our behaviour is 'determined'; second, a confusion between our behaviour being determined and our supposed incapacity of making choices; and third, a about what logically follow about the morality of our actions if we accept that our choices are determined. I tend to align with Robert Sapolsky on most of what he says about these topics, and so I think that it is mostly his critics who are subject to the three s, though probably he has not done enough to avoid them (and, in the case of the third one, it is possible that he is still also a victim of it to some degree… though perhaps this is just a misunderstanding on my part). Hence, I would like my paper be read as an effort of defending Sapolsky's claims from other people's misunderstandings, in spite of the fact that at some point I might sound as trying to 'correct' him, but only in the sense of "well, perhaps if the argument were expressed in this way, it would be clearer that Sapolsky is right".

1. The Inevitable

Regarding the first point, I confess I love the title of Sapolsky's book: *Determined*. The problem is that the claim that our choices and our behaviour is totally determined by factors we don't control can easily lead to assume that the author is committed to a couple of claims I don't think he accepts (that the universe is governed by strictly deterministic laws, and that we lack the capacity of making conscious, voluntary and deliberate choices -i.e., that we lack 'agency'), and can make other authors think that by refuting those claims they are refuting Saposlky's arguments. I will speak about agency in the next section, and concentrate now on the first point, that about determinism. Whether the 'ultimate' physical laws (if there is something like that) are deterministic or indeterministic is a question that I think Sapolsky's book leaves clearly open, as testified by the fact that he devotes significant efforts to show how free will is incompatible with *indeterministic* physical laws. The important point, instead, is whether the combination of deterministic and indeterministic physical events leave any place for our choices being really free (i.e., '*undetermined*'). The 'determinism' implied in the book's title is not that of physical deterministic laws, but better the 'fatalism' (if I may use this word, and please, erase from it all supernatural connotations) involved

in the following assertion:

If everything on which you had no control had been exactly the same, then your choice would have been exactly the same.

This claim is compatible both with determinism and with indeterminism because amongst the things on which you have had no control, there can be things that happen in an indeterministic way (for example, a synapsis in my brain firing -instead of not firing- due to a quantum jump, when both firing and not firing had been possible -but each with a fixed probability- from the wave function describing the system the synapsis belongs to). I doubt that these types of indeterminate 'quantum jumps' regularly occur within our warm, wet and messy bodies, which are very inauspicious for the quantum superpositions required for 'jumps' to take place. But I admit this is a scientifically unsettled question, and I prefer to leave it open. The point, hence, is that the state of the world immediately previous to our choice, together with the possible quantum events that happen in that instant, completely determined our decision. A world in which *everything* had been exactly equal (including the possible quantum random events) until the moment we make a choice, would have inevitably lead to our making exactly the *same* choice. All our decisions are, hence, *inevitable*. It is not that the history of the universe could not have been different 'in principle' (if some *other* things had been different); it is that, given the history of the universe till the moment we have to choose, and given every chancy event that may happen in that moment and that is not under our control, our choice could not have been a different one.

Some authors had pretended to escape this fatalist conclusion by taking refuge in the distinction between 'lower' and 'higher' ontological levels, and in the idea that there is a line of causation from the high to the low levels, or what they call *top-down causation*. But causation is neither 'downward' nor 'upward': causation is merely *forward*, from the past state of the system to the future state of the system. After all, the state of the system at the higher levels *supervenes* on its state at lower levels (i.e., if all the universe happened to be the same at the level of subatomic particles, your rhododendron would have flowered exactly the same day as it did, no matter how pompously you pronounce that the flower exists at a higher ontological level than its molecules), and this entails that the 'efficient causation' manifested by 'higher levels' is just *borrowed* from that of the most basic levels: the only thing capable of making a particle *move* from one place to another is the aggregate physical forces or the surrounding physical fields, and this is as true for a cosmic ray traversing the intergalactic space, as for any of the atoms in your rhododendron's blossoms. Hence, assuming that from our brains emerge some kind of 'higher ontological level' (call it 'agency', 'self', 'will', or whatever) that has the capacity of making the whole system behave in a certain way *instead* of another, implicitly assumes that this ontological level is an uncaused cause (or, less metaphysically: interpreting it as something that is a cause of some effects, but is not the effect of other, *previous* causes, which obviously it cannot be the case: *everything is caused* -beside the quantum jumps). Furthermore, there is a simple argument showing that this 'uncaused cause' is inconsistent with physics: for, if it were real, that would mean that the 'self' (or whatever) would have the capacity of making that an indeterminate physical system (which, according the quantum laws, has *a certain probability* (less than one) of 'collapsing' on some states, and other probabilities of doing it on other possible states) collapsed *always* in the same state (just if the self happens to freely want it), violating in this way the probabilities that the quantum laws determine. (I don't resist the temptation of pointing to the fact

that this arguments suggest a way of testing the physical assumptions that underly the idea of free will: if free will were real, it would be possible to construct a physical system that quantum laws leave indeterminate -with a determinate probability of behaving in a way or another-, but that always behaves in the same way -because it freely chooses so-. Show that this is physically possible, and you would have proved that free will is physically possible).

In conclusions, our choices and actions can be *determined* even if *determinism* is not the right view on physical laws.

2. Condemned to choose

Jean Paul Sartre famously asserted that 'we are condemned to be free', a claim that deniers of free will obviously refuse. But critics of free-will-deniers, and particularly critics of Sapolsky, very often confuse the thesis that free will does not exist (i.e., that our choices are not free), with the (wrong) claim that *we don't make* (*'real'*) *choices*, or that we don't have control over our behaviour. Of course, Sapolsky never claims that we don't make (voluntary, deliberate, conscious) choices, for a big part of his books is precisely devoted to describing how we *do* it (and what is possible and not possible for us *in* doing it). What Sapolsky's and the other free-will-deniers say is that our choices are not free, not that they are not choices. In this sense, free-will-deniers could rephrase Sartre's dictum as 'we are condemned to choose (and are condemned to choose in the only way it was possible for us to choose given everything on which we had no control)'. Unfortunately, Sapolsky's choice (☺) of words is perhaps a little bit infelicitous regarding this point, for he often writes things like 'We are no captains of our ships; our ships never had captains' (*Determined*, p. 386). Of course this must be taken as just a metaphor, whose literal content is rather the very sensible idea that there is no centralised, uncaused nucleus of agency from which our will autonomously emerges, but that our nervous system is a much messier thing, better depicted as the combination of a myriad of processes, many of them unconscious, and none of them with the magical capacity of influencing our body without being determined by (the sum of everything) that happens in it. But many readers are reasonably tempted to ignore the subtle difference between the (wrong) idea that we have no (conscious, deliberate) control on our decisions, and the (right) idea that, in each particular instance, we can only have exerted that control in the way we did it (given everything on which we didn't have control, as we saw in the past section). After all, conscious, deliberate, voluntary action is a very different *type of activity* from the neurological and physiological point of view than other types of behaviours on which we certainly don't have voluntary control; there are even muscles that we cannot move voluntarily, and muscles that (in normal circumstances) can only be moved deliberately. I don't think Sapolsky would deny that humans have certain cognitive capacities, absent to a higher or lower degree in depending which other animal species, and totally inexistent in entities like trees or volcanos (and surely in existing computers), that we can call 'agency', or even 'conscious agency', though he rightly argues that our sense of agency can be subject to numerous and frequent cognitive illusions. He won't also deny (I guess) that our brains' creating conscious experiences has something important to do with the way in which our behaviour is much more complex and flexible than that of other creatures. The only relevant point leading to the conclusion that *free will does not exist* is that *conscious will is not free*, not that *there is no conscious will*. Repeating the

argument of the first section, we can say now that of course we often take conscious, deliberate, voluntary decisions!; the only think we deny is that we might have taken a different conscious, deliberate, and voluntary decision if everything that didn't depend on us would have been exactly the same. I.e., our cognitive process of deliberation, with all the elements it may contain (some conscious, some unconscious), would have been exactly the same if the world repeated itself in everything in which we had no choice.

This reflection suggests that perhaps one important thing missing in Sapolsky's book is a bigger attention to the role of *consciousness* in our behaviour. If what we do depends so strongly on things we do not only not control, but of which we are totally unconscious, then what is consciousness *for*? If so many critics of Sapolsky had thought his vision of man is that of a puppet, a mere automaton, then why not insisting in the way consciousness helps to modulate our acts, and why not trying to illuminate the way in which that role of consciousness doesn't undermine the fact that conscious cognitive process are as determined as unconscious ones? One important hint in this respect is the fact that conscious perception of anything seems to happen always *after* the occurrence of what is perceived, and probably our voluntary decisions are no exception to this: as Libet showed, we notice that we have taken a decision a little later than 'taking' it. People usually interpret this as a simple (or simplistic) argument that shows that conscious decision is 'illusory', but from the fact that we really perceive (say) a movement of our foot a little later than when it really happens, it clearly doesn't follow that the foot's movement is 'illusory'; it only means that the *perceptual experience of* the foot's movement is not the same as the foot's movement *itself*. Similarly, in the case of Libetian experiments the fact that we *notice* that we have taken a voluntary decision a few milliseconds after something related to this has happened in our brains does not entail that those brain processes cannot be the ones in which 'taking a *voluntary* decision' *consist*. The '*sense* of voluntariness', or 'of agency', can certainly be a *perceptual construction* created to represent something that happens in our brain, which can be identified with 'voluntariness' or 'agency', and which is *neurologically different* from the conscious perception *of it*. That the two things ('agency', and 'the *sense* of agency') are different is on the basis of the experiments showing that the latter can *in some cases* be defective and lead to illusions, but it is reasonable to assume that in normal circumstances we *perceive* our agency because *it makes biological sense* that we do. What the distinction between agency, on the one hand, and, on the other hand, the sense of agency allows to do to us (as philosophers) is suspending the assumption that our conscious perception of our voluntary agency provides us with a transparent, immediate, and infallible description of how our voluntary agency physically works at the neurological level. In fact, the dismal truth is that the *phenomenology* of voluntary choice might be, in principle, no more illuminating of the *physical nature* of voluntary choice than what the phenomenology of warmth and cold illuminates the real physical nature of temperature and heat.

By the way, my most preferred guess about the biological role of the *consciousness of agency* is that its main function has probably to do with facilitating learning. Animals can learn some things unconsciously (in the sense that they don't realize that they are learning), and even they can learn to associate some stimuli they don't happen to perceive consciously to begin with. But in cases of more complex and flexible behaviours, it seems we learn much more efficiently by *paying attention* to what we do, to the circumstances in which we do it, and to its consequences. The feeling of

agency (which includes the *feeling* that we might have acted otherwise, and the feeling of *regret* for not having chosen the best option) would be what allows us to consciously monitor our own cognitive process of choosing, and hence, to learn efficiently what is what we should do when in the future we happen to face a similar circumstance. All of which, I insist, may perfectly occur in a physical world in which only one line of events can physically happen (given all chancy events we didn't have the power to make them happen in a way or another).

3. Everything permitted?

Sapolsky wisely focuses during a big part of the book in the problem which makes the rejection of free will more unpalatable to most people, philosophers and non-philosophers alike: if we are not free (if it is inevitable that we decide what we decide -given the state of the world at the moment of our deciding-), then what is the point of distinguishing good and evil, of establishing moral norms, of blaming and praising? If there's no free will, isn't then everything 'permitted', in some sense? I confess I feel a little bit unmoved by many of the arguments in this part of the book (though agreeing with most of them), because they are so clearly addressed to a moral culture (that of the USA) very different from the society, comparatively less obsessed with 'individual responsibility', I am familiar with (that of South-Western Europe). Actually, even taking into account the deeply Catholic roots of my own Mediterranean culture, it's funny in this context to notice how Spanish scholastic theologians of the early Modern Age managed to make coherent the existence of free will (and hence, of individual guilt, merit, and responsibility) with the omniscience of God: according to the philosopher Luis de Molina (16^{th} century), this apparent problem is easily solved by declaring that *God knows in advance what you will freely choose*... in a way not too different from how your own mother also knows it. In this sense, for many of us (in our Mediterranean culture) has never been a deep metaphysical mystery the idea that *you can be responsible for things that were inevitable for you to do.* A much more ancient seed of this idea is clearly shown in nothing else than the first known European literary works: the heroes of the *Iliad* and the *Odyssey* do nothing but following (as pupets, one might say) the plans of the gods; it is the gods who 'inspire' the wills and decisions in the hearts of the heroes... but this does not preclude at all that their companions or their enemies (or, for that matter, we readers) praise them or blame them, or see their destinies as something they *deserved*.

Sapolsky's main goal in this part of the book is trying to persuade his fellow citizens that it is good (?) to depart from a vision of crime and punishment based on individual moral responsibility, and replace it with a more 'humanitarian' one, which probably would even be more efficient in reducing crime. Though I am totally sympathetic with that goal, I also think that his argument is based on a flawed implicit assumption: that criminals 'do not really deserve' a harsh punishment, because after all they only did what they were inevitably to do (given the circumstances), which for many people would be equivalent to the claim that 'criminals *deserve* a better treatment'. I see this as a mistake because, in the first place, if we accept that moral desert is a fiction, or at least that it does not depend on a metaphysical capacity of having acted otherwise (as I explained in the previous paragraph), then there is *no logical contradiction* in wanting a harsh punishment system nevertheless. Homeric heroes used to be cruel, after all.

In the second place, the person to which the argument is addressed has no more free will than the criminal we are talking about with her, and hence she has no individual responsibility in having the kind of moral judgements she has. If Sapolsky plays with her the game of arguing what (moral or factual) conclusions she 'must' endorse, then because of the same logic we might 'play' that game *with the criminals*: trying to persuade them that there are some things they 'must' not do. And, well, it will be up to 'us' (the whole society, in a democratic system -if something like this exists-) establishing what will happen when someone does what he or she must not.

Lastly, in the third place, being myself a nihilistic relativist about 'moral facts', I can't help but find a little bit funny many people's appalment with the idea that 'moral norms would collapse if we stopped believing in free will'. Because actually most people wouldn't see the slightest problem in the case of *other types of social norms* if free will did not exist. Let's consider, for example, the rules of *sports*: we wouldn't desire to change them, much less eliminate them, just by learning that the players of the rival team are not humans, but programmed humanoid robots (and that are programmed in such a way that they can behave occasionally in a way that violates the rules). The same happens when we are playing a simulated sport against a computer, which obviously has no free will, but which can obviously commit faults (by the way, this is not the case in a chess computer game, an interesting comparison to discuss). Hence, if ceasing to believe in free will doesn't make us want to abandon the rules of sports, then it shouldn't make us want to abandon the rules of the game of morality either.

Acknowledgements

Research for this paper has benefited from the Ministerio de Ciencia e Innovación (España) research projects PID2021-123938NB-I00, PID2021-125936NB-I00 and PRX22/00154.