Review

More Determined Than Scientific? Unpacking the Fallacies in Sapolsky's Case Against Free Will

Poria Azadi *

Posted Date: 8 May 2025

doi: 10.20944/preprints202505.0388.v1

Keywords: free will; determinism; Robert Sapolsky; reductionism; emergence; constraint causation; Strawman Fallacy; performative contradiction

 doi:10.20944/preprints202505.0388.v1

Review

More Determined than Scientific? Unpacking the Fallacies in Sapolsky's Case Against Free Will

Poria Azadi

Independent Researcher, Iran; poria.azadi96@gmail.com

Abstract: Robert Sapolsky's *Determined: A Science of Life Without Free Will* synthesizes findings across disciplines to argue for hard determinism, concluding free will is an illusion incompatible with science. This critique contends the work, despite its breadth, suffers fundamental conceptual and methodological flaws. Primarily, it attacks a strawman definition of free will—equating it to an uncaused, acontextual neural event, an *a priori* empty set (FWS = ∅)—thereby sidestepping meaningful engagement with sophisticated accounts. Secondly, it exhibits logical incoherence, failing to define determinism clearly and oscillating between incompatible deterministic frameworks (hard/Laplacian vs. soft/contextual) without acknowledgment, leading to performative contradictions where the author implicitly exempts himself while advocating normative changes. Thirdly, it misappropriates concepts from complexity, chaos, and emergence, interpreting them reductionistically via a simplistic view of causality (treating it as a linear chain between noumena) and neglecting the causal efficacy of organization via constraint causation. Fourthly, it conflates biological historicity with deterministic necessity, ignoring how accumulated structure enables agency. Fifthly, it implicitly adopts an outdated behaviorist stimulus-response model. Finally, it recapitulates established arguments without offering the novel biological paradigm its subtitle claims. Sapolsky's determinism appears less a scientific conclusion and more an artifact of outdated metaphysics and insufficient engagement with theoretical biology and the philosophy of complex systems.

Keywords: free will; determinism; Robert Sapolsky; reductionism; emergence; constraint causation; Strawman Fallacy; performative contradiction

1. Introduction: Reproducing an Intellectual Impasse

Robert M. Sapolsky's *Determined: A Science of Life Without Free Will* seeks to provide a decisive scientific case against the existence of free will. Drawing upon extensive evidence from neuroscience, genetics, chaos theory, complexity, and other fields, Sapolsky endeavors to demonstrate that every human action is the necessary outcome of antecedent causes beyond individual control, leaving no room for genuine agency. He states his central thesis unequivocally: "we are nothing more or less than the cumulative biological and environmental luck, over which we had no control, that has brought us to any moment" (Sapolsky, 2023, p. 5). The author's engaging prose and skill in rendering complex scientific concepts accessible to a general audience are undeniable strengths of the work.

However, notwithstanding these merits, the analysis presented in *Determined* suffers from foundational conceptual and methodological flaws that undermine the validity of its central conclusion. Rather than offering a novel or challenging perspective, the book largely remains trapped within the classic intellectual impasses of the free will debate. This critique argues that Sapolsky commits several critical errors in reaching his deterministic conclusion: (1) the deployment of a caricatured definition of free will that *a priori* constitutes an empty set (Strawman Fallacy); (2) a lack of logical coherence arising from the failure to define determinism clearly and the oscillation between incompatible deterministic frameworks, leading to performative contradictions; (3) the misappropriation and reductionist interpretation of scientific findings from complexity, chaos, and emergence, neglecting organizational constraints and constraint causation via a simplistic view of

 doi:10.20944/preprints202505.0388.v1

causality; (4) a fundamental misconstrual of biological historicity as purely deterministic constraint; (5) an implicit adherence to an outdated behaviorist framework; and (6) the reiteration of long-standing hard incompatibilist positions without significant novel analysis, failing to substantiate the book's ambitious subtitle and potentially misinterpreting historical views on responsibility.

This analysis will meticulously examine these deficiencies, demonstrating how Sapolsky's absolute determinism results less from the compelling force of scientific progress and more from inadequate metaphysical presuppositions—particularly a simplistic view of causality as a linear chain acting upon noumena—and a failure to grapple deeply with the complexities of biological organization and agency as understood in contemporary theoretical biology.

2. The Strawman Fallacy: Defining an Empty Set (∅)

The primary and perhaps most fundamental flaw pervading *Determined* is its reliance on a definition of free will that is not merely weak or simplistic, but constitutes, *a priori*, an empty set based on the very scientific principles the book invokes. Sapolsky frames his challenge to proponents of free will ("free willers") thus:

> This is my challenge to the free willers: Find me the neuron that started this process in this man's brain, the neuron that had an action potential for no reason, where no neuron spoke to it just before. Then show me that this neuron's actions were not influenced by whether the man was tired, hungry, stressed, or in pain at the time... Show me a neuron (or brain) whose generation of a behavior is independent of the sum of its biological past, and for the purposes of this book, you've demonstrated free will. (Sapolsky, 2023, p. 14)

Let FWS be the set of phenomena satisfying Sapolsky's conditions for free will. An element x belongs to FWS if and only if x is a neural event (e.g., an action potential) such that:

1. x has no antecedent neural cause ($\neg\exists y$: Cause(y, x), where y is a prior neural event).
2. x is uninfluenced by the organism's physiological state Sphysio (e.g., fatigue, hunger, stress).
3. x is uninfluenced by the organism's sensory inputs Ssensory in the preceding moments.
4. x is uninfluenced by the organism's entire biological history Hbio (including genetics, development, past experiences, epigenetic modifications, etc.).

Condition (1) demands an uncaused neurological event, violating basic principles of neurophysiology where action potentials result from synaptic inputs or intrinsic membrane dynamics influenced by prior states. Conditions (2), (3), and (4) demand complete independence from the biological context and history that fundamentally constitute any living system. Basic neuroscience and biology confirm that neural activity is intrinsically embedded within, and responsive to, physiological states, environmental inputs, and accumulated history. Therefore, based on established scientific understanding (much of which Sapolsky himself details elsewhere in the book), no such neural event x satisfying all conditions {1, 2, 3, 4} can possibly exist within a biological organism.

Formally, the set defined by Sapolsky's criteria is necessarily empty: $\mathbf{FWS = \emptyset}$.

This definition is thus not merely a caricature; it is definitionally impossible within the naturalistic framework Sapolsky claims to uphold. It is self-contradictory to seek a biological phenomenon (a neuron firing) that is simultaneously independent of all biology.

The absurdity lies in the subsequent structure of the book. Having defined free will as an *a priori* impossibility (FWS = ∅), Sapolsky dedicates hundreds of pages to demonstrating, through an exhaustive review of causal factors from neurobiology to evolution, that no instance of FWS can be found. This extensive argumentation aims to prove that $\mathbf{\forall x, x \notin \emptyset}$ – a statement that is trivially true by definition of the empty set. The book expends immense scientific effort to demonstrate that a definitionally impossible phenomenon does not occur. This is akin to writing a treatise proving that square circles do not exist by meticulously measuring every circle and square.

This constitutes a profound category error and a misdirection of the entire debate. Instead of engaging with meaningful philosophical positions on free will—compatibilist accounts of freedom as acting from one's character or reasons (even if determined), or sophisticated libertarian accounts

 doi:10.20944/preprints202505.0388.v1

involving agent-causation or structured indeterminacy (Kane, 2002; O'Connor, 2000)—Sapolsky refutes a self-defined impossibility. As Aikin & Casey (2011) argue, such straw man fallacies fundamentally obstruct productive discourse by attacking misrepresented positions. The central argument of *Determined* rests not on a scientific refutation of plausible free will concepts, but on the elaborate demolition of a conceptual chimera defined *a priori* as non-existent within the natural world.

3. Logical Incoherence: Undefined Determinism, Conflated Frameworks, and Performative Contradiction

A second cluster of critical weaknesses lies in the book's ambiguous and inconsistent use of "determinism," its conflation of different causal frameworks, and the resulting performative contradictions inherent in its normative conclusions.

3.1. Undefined Determinism: Oscillating Between Hard and Soft Views

Crucially, Sapolsky never provides a precise, consistent definition of the "determinism" he champions. Instead, he oscillates between two fundamentally different philosophical conceptions, often blurring the lines between them:

- **Hard Determinism (Laplacian):** This view posits absolute necessity. Given the state of the universe at one time and the laws of nature, only one future is possible. Every event, no matter how small, is rigidly fixed. Sapolsky invokes this strong sense frequently, especially when emphasizing the "seamless stream" of causes stretching back indefinitely: "when you behave in a particular way... it is because of the determinism that came just before, which was caused by the determinism just before that, and before that, all the way down" (Sapolsky, 2023, p. 3). This implies a block universe where the future is as fixed as the past, leaving no room for objective chance or alternative possibilities.
- **Soft Determinism (Conceptual):** This perspective, while still deterministic, differs metaphysically. It might allow for objective chance at micro-levels (e.g., quantum mechanics) but argues that these micro-indeterminacies are irrelevant to macro-level behavior, which remains effectively determined by large-scale factors or constraints. Alternatively, it might focus less on absolute necessitation and more on the overwhelming *influence* of prior factors, arguing that even if tiny variations were possible, they wouldn't ground meaningful free will or responsibility because the *character* and *reasons* for action are shaped by factors beyond our control. The emphasis shifts from the impossibility of alternatives to the lack of control over the *sources* of our actions.

Sapolsky seems unaware of, or at least fails to articulate, the crucial distinction between these two positions. He uses the language of hard, Laplacian determinism ("no cracks," "turtles all the way down") to establish the *impossibility* of free will, but often shifts towards arguments more characteristic of soft determinism when discussing responsibility, focusing on the lack of control over genetics, upbringing, etc. (e.g., Chapter 9, Chapter 11). He doesn't clarify whether he believes *literally no event* could have been different (hard determinism) or whether macro-level outcomes are effectively determined despite potential micro-indeterminacy or simply overwhelmingly influenced (closer to soft determinism). This failure to define his core concept leaves the foundation of his argument unstable and ambiguous. Is he defending a specific metaphysical thesis about necessity, or a broader claim about overwhelming influence? The book never clarifies.

3.2. Conflating Physical and Sociobiological Factors

Compounding the ambiguity, Sapolsky often conflates arguments based on physical determinism (neural processes, physics) with arguments based on sociobiological or contextual determinism (genes, hormones, environment, culture). While both contribute to his overall deterministic picture, they operate at different levels of explanation and involve different kinds of

 doi:10.20944/preprints202505.0388.v1

causal relations. He seamlessly moves from discussing neurotransmitters (p. 45) to childhood trauma (p. 66) to cultural norms (p. 74) as if they are all simply links in the same linear chain of efficient causes, without adequately addressing the complexities of inter-level causation or the distinct nature of biological, psychological, and social explanations. This further obscures precisely what kind of "determinism" is being asserted.

3.3. The Performative Contradiction and Authorial Self-Exemption

This lack of clarity exacerbates the performative contradiction inherent in the book's normative aims. Sapolsky argues forcefully against blame and retribution:

> ...there can be no such thing as blame, and that punishment as retribution is indefensible... (Sapolsky, 2023, p. 9)
>
> ...we need to accept the absurdity of hating any person for anything they've done... (Sapolsky, 2023, p. 403)

He advocates for societal change based on this understanding:

> The goal of this and the next chapter is to explore the theme of the second half of this book, namely that regardless of it seeming unimaginable, we can change in these realms. (Sapolsky, 2023, p. 302)
>
> This is where the science has taken us. And we need to accept the absurdity... (Sapolsky, 2023, p. 403)

If hard (Laplacian) determinism is true, advocating for *us* to "accept," "change," or "stop blaming" is logically incoherent, as these actions themselves would be predetermined. Such advocacy only makes sense if Sapolsky implicitly adopts a weaker form of determinism or exempts himself and his readers from the absolute necessity he otherwise proclaims (Sobel, 2023). He acts as a rational agent attempting to persuade other rational agents to *choose* a different perspective, thereby presupposing the very capacities his strongest arguments deny.

Furthermore, this contradiction is amplified by Sapolsky's implicit self-exemption from the deterministic framework he imposes on others. While arguing that no one *chooses* their beliefs or actions, he presents his own conclusions as the result of reasoned scientific analysis and offers his book as a means to achieve a better, more rational, and more humane society by *changing* our beliefs and practices regarding free will ("This is the goal... to change fundamental aspects of how we behave," p. 244). He writes as if he, the author, has somehow transcended the deterministic chains he describes, enabling him to offer insights and solutions from an objective standpoint, seemingly unaware of the profound contradiction. This stance mirrors the paradox faced by any proponent of strong determinism who simultaneously advocates for specific societal changes or ethical viewpoints. It requires assuming a privileged position outside the deterministic system one is describing, a position that the theory itself denies is possible. His normative claims about punishment also risk oversimplifying complex historical and philosophical contexts where concepts of responsibility might not have directly mapped onto modern notions of free will, as seen in his discussion of historical punishments (Chapter 14) which he links to free will beliefs potentially anachronistically, ignoring periods where legal/philosophical responsibility was conceived differently.

This unresolved tension between the book's deterministic claims and its normative, change-oriented stance, combined with the lack of a precise definition of determinism, constitutes a significant logical flaw.

4. Simplistic Causality and Misunderstanding Emergence

One of the most striking and problematic aspects of *Determined* is Sapolsky's attempt to enlist concepts from complexity theory, chaos theory, and emergence as evidence *for* his deterministic stance (Chapters 5-8). This approach involves a significant misinterpretation of causality and

 doi:10.20944/preprints202505.0388.v1

emergence, driven by an underlying reductionist framework that is inconsistent with contemporary understandings of complex systems.

4.1. Causality as a Linear Chain: The "Turtles All the Way Down" Fallacy

Sapolsky's core argument rests on tracing causal influences backward in time, from the immediate neural precursors of an action (milliseconds before) to hormones (minutes to days), neuroplasticity (weeks to years), development (childhood, adolescence, fetal life), genetics, culture, and finally, evolution (Chapters 2, 3, 4, 9, 10). He repeatedly emphasizes the "seamless stream of influences" (p. 81), concluding that because every step is influenced by the previous one, and we didn't control the ultimate origins, there is no room for free will. This invokes the image of "turtles all the way down" (p. 2, 82), suggesting an infinite regress of prior efficient causes determining the present state.

> Why did that behavior occur? Because of biological and environmental interactions, all the way down. (Sapolsky, 2023, p. 4)
>
> Why did that moment just occur? "Because of what came before it." Then why did that moment just occur? "Because of what came before that," forever... (Sapolsky, 2023, p. 82)

While the tracing of influences is scientifically informative, the *causal interpretation* is simplistic and aligns with a classical, linear view of efficient causation. It implicitly assumes:

1. **Bottom-up Exclusivity:** Causation flows strictly from lower levels (genes, neurons) or earlier times to higher levels (behavior) or later times.
2. **Linearity & Sufficiency:** Effects are proportional to causes, and the chain is unbroken and singular, with each prior state being sufficient for the next.
3. **Causality between Noumena:** The argument implicitly treats this causal chain as operating on an underlying, fixed reality (noumena), rather than on observable scientific entities and their relationships within specific descriptive frameworks.

This view fails to account for the complexities of causation in biological systems, particularly the role of organization and constraints. Scientific causality concerns relations between scientific entities within descriptive languages (L1, L2, etc.). Ontological irreducibility arises when phenomena described in L2 cannot be fully explained in L1 without loss of understanding, due to informational and semiotic differences ("Irreducibility in Unfolding Futures"). Sapolsky's linear chain model, by focusing solely on antecedent efficient causes acting on a presumed fixed underlying reality (noumena), implicitly attempts reduction at the noumenal level, a philosophically problematic stance that overlooks how different levels of description capture distinct, irreducible causal relationships and how organization itself exerts causal influence via constraints.

4.2. Reductionist Treatment of Emergence and Downward Causation

Sapolsky dedicates chapters to emergent complexity (Chapters 7-8), discussing phenomena like ant colonies, slime molds, and neural self-organization. He acknowledges their complexity and unpredictability but interprets them through a strictly reductionist and deterministic lens.

> Put enough of the same simple elements together, and they spontaneously self-assemble into something flabbergastingly complex... With enough quantity, extraordinary quality just... emerges, often even unpredictably. (Sapolsky, 2023, p. 155)
>
> [Emergence explains] how purely deterministic constituent parts obeying simple rules could produce collective phenomena whose behavior is typically unpredictable. (Sapolsky, 2023, p. 190)

He explicitly rejects strong emergence and the idea that emergent states can alter the rules governing their constituents:

> Even if consciousness qualifies as the closest thing to true strong emergence, it's sure not a case of it. (Sapolsky, 2023, p. 197)

 doi:10.20944/preprints202505.0388.v1

> Downward causation doesn't cause individual building blocks to acquire complicated skills; instead, it determines the contexts in which the blocks are doing their idiotically simple things... Emergent systems can't make the bricks that built them stop being brick-ish. (Sapolsky, 2023, p. 201)

This reveals a fundamental misunderstanding of contemporary views on emergence and downward causation, particularly the concept of **constraint causation**. Emergence is not about components magically changing their fundamental properties or violating physical laws. It is about how the *organization* of components at a higher level creates *constraints* that channel the behavior of those components in ways not possible in isolation (Ellis, 2016; Juarrero, 1999).

- **Constraint Causation:** Higher-level structures (e.g., cell membranes, neural network architectures, metabolic cycles, social norms) act as constraints. They reduce the degrees of freedom available to lower-level components, selecting specific pathways and stabilizing functional patterns. This selection and channeling *is* a form of downward causal influence. It's not a new force, but the causal efficacy of organization itself. The "bricks" remain "brick-ish" in their fundamental properties, but the emergent "arch" constrains where and how those bricks can be placed and how forces are distributed among them.
- **Ontological Irreducibility:** This organizational level, defined by constraints, often requires its own descriptive language (L2) that is not fully reducible to the language of the components (L1) without loss of explanatory power regarding the system's function and behavior ("Irreducibility in Unfolding Futures," Section VI). The causal powers associated with the organization (e.g., the self-maintenance of an operational closure) are properties of the whole, irreducible to the parts.

Sapolsky's dismissal stems from viewing causality solely as forceful efficient causation between parts. He lacks the conceptual framework to understand how constraints, arising from organization, can exert genuine, irreducible causal influence. His view remains tethered to a simplistic, bottom-up mechanism that cannot accommodate the causal power of form and organization central to understanding life and agency (Moreno & Mossio, 2015; Noble, 2012).

4.3. Misappropriation of Chaos and Complexity

As discussed previously (Section 4.1), Sapolsky treats the unpredictability of chaotic systems as purely epistemic, ignoring the ontological challenge it poses to Laplacian determinism when combined with physical limits on precision (Gisin, 2020). He uses complexity and self-organization merely as examples of intricate determinism, failing to see them as potential sources of emergent autonomy grounded in organizational closure and constraint causation (Letelier et al., 2011; Rosen, 2012).

In essence, Sapolsky interprets all evidence through the lens of his presupposed linear, bottom-up, efficient-cause determinism acting on a fixed reality. He fails to engage with alternative causal frameworks (like constraint causation) or ontological perspectives (like process philosophy or non-reductive physicalism based on organizational emergence) that are arguably better suited to the phenomena he discusses. His treatment of emergence and causality remains simplistic and fails to capture the insights offered by contemporary systems biology and philosophy of complex systems.

5. Conflating Historicity with Deterministic Necessity

A related conceptual error is Sapolsky's tendency to equate the *historicity* of biological systems with deterministic necessity that precludes freedom. Throughout the book, he meticulously traces the influences of evolutionary history, cultural history, developmental history (fetal life, childhood, adolescence), and learning on present behavior. He argues that because these historical factors shape who we are and are outside our control, our current actions are determined and not free.

> ...all we are is the history of our biology, over which we had no control, and its interactions with environments, over which we also had no control, creating who we are in the moment. (Sapolsky, 2023, p. 85)

 doi:10.20944/preprints202505.0388.v1

This argument conflates *influence* with *necessitation* and fundamentally misunderstands the role of historicity in complex adaptive systems. While history undoubtedly shapes and constrains present possibilities, it does not necessarily eliminate them or render the present state rigidly determined in a Laplacian sense.

- **History Enables Agency:** Far from merely constraining, the accumulated history of an organism—its memories, learned skills, developed structures, epigenetic modifications—constitutes the very architecture that *enables* sophisticated agency. Memory allows for learning from past mistakes; developmental processes build the complex machinery required for perception and action; cultural history provides norms and tools. Historicity, embodied in the agent's organizational structure (its constraints), expands the space of possible actions and allows for more nuanced, context-appropriate responses compared to a system without such history. A system with memory and learned constraints has *more* degrees of freedom in its meaningful interactions with the world, not fewer, compared to a simple reactive system.
- **Historicity vs. Determinism:** That an agent's current state is a product of its history does not entail that its *next* state is fully determined by its current state plus external inputs. The historical constitution shapes the *possibility space* and the *probabilities* of transitions within that space, but within frameworks incorporating structured indeterminacy or operational closure, the specific outcome may still remain open until actualized. Sapolsky uses history as evidence for determinism, whereas a systems perspective sees history as the foundation upon which adaptive, potentially non-deterministic, agency is built.

By treating history solely as a source of deterministic constraints, Sapolsky overlooks its crucial role in creating the complex, adaptive organization that makes sophisticated agency possible in the first place. He mistakes the necessary conditions for agency (a history that builds structure) for conditions that preclude agency.

6. Neglect of Theoretical Biology: A Behaviorist Echo

A further significant conceptual deficit in *Determined* is its apparent neglect of major developments in theoretical biology, particularly those concerning self-organization, autonomy, and the active role of the organism, which have unfolded since the decline of classical behaviorism in the mid-20th century. Sapolsky's framework often implicitly reverts to a stimulus-response model, even when discussing internal biological factors, thereby failing to capture the essence of biological agency as understood in contemporary systems biology and related fields.

6.1. Internal States as External Stimuli?

Sapolsky meticulously details how various factors—from sensory inputs milliseconds before an action to hormone levels hours before, genetic predispositions, and cultural norms absorbed over decades—influence behavior. For example, he discusses how sensory cues like disgusting smells or attractive faces influence moral judgments (pp. 47-50), how hormone levels (testosterone, oxytocin, glucocorticoids) modulate aggression, trust, or impulsivity (pp. 52-57), and how stress or hunger affects decision-making (pp. 105-109).

While acknowledging these influences is crucial, Sapolsky's explanatory framework often treats these internal biological states (hormone levels, blood glucose, stress responses) as if they were *external stimuli* impinging upon a passive organism or brain, determining its response. He describes hormones "altering behavior" (p. 52) or "marinating your fetal brain" (p. 68), and stress causing the PFC to function poorly (p. 105). The underlying model resembles S → R (Stimulus → Response), where the "stimulus" can be an internal physiological state, but the organism itself remains largely a reactive entity whose behavior is *determined by* these factors, rather than an active agent *regulating* these factors as part of its self-maintenance.

6.2. Ignoring Self-Regulation and Operational Closure

 doi:10.20944/preprints202505.0388.v1

This framing overlooks decades of theoretical biology emphasizing the organism as an *autonomous, self-regulating system*. Concepts like autopoiesis (Maturana & Varela, 1980), organizational closure (Moreno & Mossio, 2015; Letelier et al., 2011; Rosen, 2012), and related ideas from systems biology (Noble, 2012) highlight that living systems actively maintain their internal milieu and organizational integrity through complex feedback loops and regulatory networks. Hormone levels, blood glucose, or stress responses are not merely external inputs determining behavior; they are dynamically regulated variables that are part of the organism's own self-maintaining organization. The organism actively works to keep these within viable ranges (homeostasis/allostasis) according to its intrinsic norms.

Sapolsky's linear causal chain (past events → current internal state → behavior) fails to capture this circular causality inherent in self-regulation. By treating internal states primarily as determinants of behavior rather than as components of a self-regulating system that *also* determines those states, he implicitly adopts a behaviorist stance that ignores the organism's active role in constituting and maintaining its own internal conditions. This prevents him from seeing agency as potentially arising from this very capacity for self-regulation and self-maintenance within the bounds of operational closure. The organism is reduced to a complex reactor, rather than an autonomous agent actively managing its internal and external interactions to ensure its viability. This perspective is surprisingly disconnected from the core tenets of theoretical biology developed since the 1950s, which emphasize the active, self-constructing nature of life.

7. Lack of Novelty and Unsubstantiated Ambition

Despite its extensive scope and engagement with modern scientific findings, a significant criticism of *Determined* is its ultimate lack of conceptual novelty and its failure to live up to the paradigm-shifting claim implied by its subtitle, "A Science of Life without Free Will."

7.1. Recapitulating Established Positions

The book's central thesis—hard determinism is true and incompatible with free will—represents a long-standing position in the history of philosophy. Sapolsky primarily reiterates classical incompatibilist arguments, marshalling contemporary scientific evidence in their support. While the synthesis is valuable, the underlying philosophical argument offers little beyond established hard incompatibilist or hard determinist positions, failing to engage deeply with sophisticated compatibilist or alternative libertarian views (Kane, 2002; O'Connor, 2000; Doyle, 2011).

7.2. The Unjustified Claim of a New Paradigm

The subtitle boldly suggests the book introduces a new scientific paradigm for biology, one fundamentally devoid of free will. This implicitly places the work alongside genuinely paradigm-shifting contributions in theoretical biology, such as Maturana and Varela's autopoiesis, Rosen's relational biology ((M,R) systems), or Kauffman's work on self-organization and the adjacent possible. These works introduced fundamentally new conceptual frameworks and formalisms for understanding life's organization.

Determined, however, offers no comparable theoretical innovation. It does not propose a new model of biological organization, causality, or information processing. Instead, it interprets existing scientific findings through the lens of classical determinism and reductionism. The "science of life without free will" presented is not a new biological paradigm but rather the application of an old philosophical stance (hard determinism) to contemporary biology, often by misinterpreting or neglecting concepts like emergence and organizational closure that are central to modern theoretical biology (Letelier et al., 2011; Moreno & Mossio, 2015). The book's ambition, as stated in the subtitle, remains largely unsubstantiated by its actual conceptual contribution.

8. Conclusion: Determinism as Presupposition, Not Scientific Outcome

 doi:10.20944/preprints202505.0388.v1

Robert Sapolsky's *Determined*, despite its impressive synthesis of scientific findings related to human behavior, ultimately fails to provide a convincing argument for the complete absence of free will. The absolute determinism and denial of meaningful agency that form the book's core message appear less as inescapable conclusions drawn from scientific evidence and more as products of conceptual flaws, logical inconsistencies, and reductionist interpretations.

Sapolsky initiates his argument by attacking a caricatured, *a priori* empty definition of free will (Strawman Fallacy). He undermines logical coherence by equivocating between incompatible and poorly defined deterministic frameworks and falling into performative contradictions. He misappropriates concepts from complexity, chaos, and emergence, interpreting them through a reductionist lens that fails to recognize constraint causation and ontological irreducibility. He conflates historicity with deterministic necessity and implicitly adopts an outdated behaviorist framework, neglecting core concepts from theoretical biology like operational closure. Finally, the work largely recapitulates established arguments, failing to deliver the promised new biological paradigm.

The determinism presented in *Determined* seems less compelled by the scientific data itself and more reflective of the author's underlying metaphysical assumptions—a commitment to a classical, mechanistic worldview operating on noumena (Smolin, 2014)—and a reluctance to embrace richer conceptual frameworks offered by complexity science, systems biology, process philosophy, and related fields. These alternative frameworks, focusing on concepts like organizational closure (Letelier et al., 2011; Rosen, 2012), emergence grounded in constraint causation (Ellis, 2020; Noble, 2012; Mascolo & Kallio, 2019), embodied interaction (Gallagher, 2020), and dynamic time (Bouton, 2021; Gisin, 2020), offer pathways to understand agency and freedom not as mysteries or illusions, but as emergent, natural properties of complex biological systems (Costa et al., 2022).

Ultimately, *Determined*, despite its narrative power and scientific scope, serves as a prominent example of the limitations of reductionist thinking and the risks of neglecting organizational and conceptual complexity when addressing profound philosophical and existential questions. Rather than providing a definitive resolution to the problem of free will, the book primarily illustrates the constraints of the paradigm within which it operates, presenting determinism more as a presupposition guiding the interpretation of science than an outcome rigorously derived from it.

References

Aikin, S. F., & Casey, J. (2011). Straw Men, Iron Men, and Argumentative Virtue. *Topoi, 30*(1), 45-55. https://doi.org/10.1007/s11245-010-9095-y

Bedau, M. A. (2019). The nature of life. In S. Schneider & M. Velmans (Eds.), *The Blackwell Companion to Consciousness* (2nd ed., pp. 754–770). Wiley Blackwell. https://doi.org/10.1002/9781119132363.ch54

Bishop, R. C. (2019). Towards a Middle Way: Non-Reductive Physicalism and Contextual Emergence. *Journal of Consciousness Studies, 26*(1-2), 148-172.

Bouton, M. E. (2021). Time: A Dynamic Non-Reductive Approach to Temporal Passage. *The British Journal for the Philosophy of Science, 72*(3), 757-784. https://doi.org/10.1093/bjps/axz059

Costa, F., Duarte, C., & Santos, M. (2022). Agency as informed autonomy: A naturalized approach to free will. *Biological Theory, 17*(2), 101-118. https://doi.org/10.1007/s13752-021-00387-6

Doyle, R. (2011). *Free Will: The Scandal in Philosophy*. I-Phi Press.

Dubois, D. (2020). Causal determinism in physics and beyond: False dichotomies and missed opportunities. *Entropy, 22*(5), 515. https://doi.org/10.3390/e22050515

Ellis, G. F. R. (2016). *How can physics underlie the mind? Top-down causation in the human context*. Springer.

Ellis, G. F. R. (2020). The Causal Closure of Physics in Real World Contexts. *Foundations of Physics, 50*(10), 1057-1097. https://doi.org/10.1007/s10701-020-00366-0

Gallagher, S. (2020). *Action and Interaction*. Oxford University Press. ISBN: 978-0198794430

Gisin, N. (2020). Mathematical languages shape our understanding of time in physics. *Nature Physics, 16*(2), 114-119. https://doi.org/10.1038/s41567-019-0748-5

Juarrero, A. (1999). *Dynamics in action: Intentional behavior as a complex system*. MIT Press.

 doi:10.20944/preprints202505.0388.v1

Kane, R. (Ed.). (2002). *The Oxford Handbook of Free Will*. Oxford University Press.

Letelier, J. C., Cárdenas, M. L., & Cornish-Bowden, A. (2011). From L'Homme Machine to metabolic closure: Steps towards understanding life. *Journal of Theoretical Biology, 286*(1), 100-113. https://doi.org/10.1016/j.jtbi.2011.06.033

Mascolo, M. F., & Kallio, E. (2019). Beyond free will: The embodied emergence of conscious agency. *Philosophical Psychology, 32*(4), 547-575. https://doi.org/10.1080/09515089.2019.1587910

Maturana, H. R., & Varela, F. J. (1980). *Autopoiesis and Cognition: The Realization of the Living*. D. Reidel Publishing Company.

Moreno, A., & Mossio, M. (2015). *Biological autonomy: A philosophical and theoretical enquiry*. Springer. https://doi.org/10.1007/978-94-017-9837-2

Noble, D. (2012). A theory of biological relativity: no privileged level of causation. *Interface Focus, 2*(1), 55-64. https://doi.org/10.1098/rsfs.2011.0067

O'Connor, T. (2000). *Persons and Causes: The Metaphysics of Free Will*. Oxford University Press.

Rosen, R. (2012). *Life Itself: A Comprehensive Inquiry into the Nature, Origin, and Fabrication of Life*. Columbia University Press. ISBN: 978-0231075657

Sapolsky, R. M. (2023). *Determined: A Science of Life Without Free Will*. Penguin Press.

Smolin, L. (2014). *Time Reborn: From the Crisis in Physics to the Future of the Universe*. Mariner Books. ISBN: 978-0544245594

Sobel, J. H. (2023). The Performative Contradiction in Denying Agency. *Philosophical Studies, 180*(3), 1123-1144. https://doi.org/10.1007/s11098-022-01871-8