IMPERIAL COLLEGE LONDON – PHILOSOPHY OF MIND

Do We Have Free Will?: An Inquiry into Digital Compulsion and the Autonomy of the User

Word Count: 1997
Chosen Essay Topic: 1

Joshua Ho
Module Leader: Mr. Greg Artus

A report submitted in partial fulfilment of the I-Explore module requirements

January 2026

1 Essay

1.1 Background and motivation

I entered this course motivated by a personal struggle with phone addiction, which felt like a direct challenge to my sense of agency. Coming from a STEM-focused background, I initially adopted a hardline physicalism, viewing addiction as a purely neurological malfunction measurable through brain activity.

At the London Neurotech Hackathon last year, my team used an EEG headband to model distraction as a data point, assuming that sufficiently detailed neural measurements could essentially “solve” the mind.

Figure 1: *Me at the 2025 London Neurotech Hackathon, attempting to find the ’mind’ in a stream of EEG data.*

However, this course has challenged the "brute facts" I once relied upon. If I am merely a machine responding to stimuli, who is the "I" that feels "crippled" by the addiction? By exploring the mind-body problem, I realized I was living in a state of philosophical contradiction. On one hand, I leaned on a reductionist view that mental states are just neural states; on the other, I appealed a "Cartesian" hope that my "self" could somehow exist independently of these biological urges to exert "free will".

1.2 Accounts of the mind

To resolve this contradiction, I must re-evaluate the account of the mind that best aligns with a eudaimonic way of living. Most contemporary neuroscience implicitly accepts Materialism, which is the view that the brain causes all aspects of the mind. This stands in stark contrast to Substance Dualism, which (following Leibniz's Law) argues that since the mind and body have different properties, they must be distinct substances.

While Dualism is often considered an impediment to neuroscience, it is supported by the work of Wilder [1], a materialist-turned dualist. Penfield observed that while he could stimulate a patient's brain to move their arm, the patient could clearly distinguish this "passive" movement from one they "willed" themselves. This suggests a "self" that monitors the machine from the outside. However, I find this unconvincing. just because its not been sampled doesn't mean black swans cannot exist? Besides, just because its not measured does not reject the possibility that these abstract concepts can only be triggered by a specific combination of simulations.

Seeking a more rigorous physicalism, I turn to Eliminative Materialism, which quite radically claims our common-sense understanding of the mind is wrong. In this view, "digital detoxing" is not a mental struggle of the will, but a mechanical recalibration of a chemical imbalance. Yet, this feels insufficient. If we eliminate "desire" and "willpower" as mere "folk psychology," we lose the very vocabulary needed to describe the struggle of addiction. Furthermore, Behaviorism fails to bridge this gap; as Hilary Putnam's "Super-Spartans" [2] thought experiment suggests, mental states cannot be reduced to outward behavior alone, as we can suppress behavior (like the urge to check a phone) while the internal desire remains.

These failures suggest that looking for the mind as a "thing" inside the brain, as I did at the Hackathon, is a Category Mistake. The mind is not a part of the machine, but the way the machine functions. This realization shifts the inquiry from the substance of the mind to the autonomy of the user: Do I have free will?

1.3 Do I have free will?

While people generally believe humans exhibit freedom to some extent, the "free will" problem remains a subject of debate, and is center to giving us a better account of phone addiction. Free will has traditionally been conceived of as a kind of power to control one's choices and actions, yet philosphical discussion has led to various strands of thought that should be disentangled.

Determinism posits that all events are determined by causes. I decide to commute to school by foot, which is why I do; I want to check my phone messages, so I do. We constantly assume this truth in our daily dealings; for instance, a tutor corrects a pupil's work with the expectation that this external intervention will cause a change in the pupil's future behavior.

However, Determinism must be distinguished from Fatalism, which suggests that "ends" are fixed regardless of the "means." I argue that Fatalism is philosophically bankrupt due to its lack of falsifiability. Consider a student with no engineering background who takes an engineering module; if they "doomscroll" under the assumption they are "fated to pass," they will almost certainly fail, as complex course content does not materialize "out of thin air." And yet, the fatalist would then make a retroactive claim that the student was simply "fated to fail", and that his original belief that he's "fated to pass" is deemed false. Such a stance ignores the physical reality that a "test-passing" mental state is a physical result of a causal "studying" process. Furthermore, the fatalist thesis lacks explanatory power regarding the mind-body relationship. If a physical outcome were fixed regardless of prior states, mental effort or deliberation would be reduced to a purposeless shadow, or an epiphenomenon. This contradicts the lived experience of overcoming addiction, where the mental "will" to resist stimuli feels causally significant. While some might introduce theological predestination

into this debate, such a move shifts the discussion into theodicy and remains outside the scope of a philosophy of mind centered on the metaphysics of presence and material processes.

By recognizing that outcomes are not independent of actions, we shift our focus to the tension between Determinism and Indeterminism. Indeterminism, the view that some events are not entirely determined by prior causes and involve an element of chance, initially appears to be an "optimistic" account of agency in the context of phone addiction; it suggests that an individual might "snap out" of a dopamine-driven cycle of doom-scrolling out of the blue not by their past attempts to reduce usage but arbitrarily. The observable implication is that even an observer with complete knowledge of a person's past could fail to predict their next action.

However, I argue that this "uncaused" type of freedom is a misidentification of agency. If a characteristically righteous person suddenly commits a murder with no roots in their previous character, the Indeterminist might label this a "causeless act." The Determinist, however, questions whether a random act, if such a thing even exists, can be attributed to the person. The reasoning holds that if an action lacks a causal link to the person's desires, values, or history, it is not an act of "will," but a biological glitch, be it a random "voltage spike" in the brain for which the person should bear no moral responsibility. The Determinist insists that true freedom requires the action to be "issued" by the person, conforming to their established character. They may go a step further and argue that the "causeless" actions suggested by Indeterminists are a taxonomic misidentification, and that what is perceived as an action unrooted in character is merely a misinterpretation of that character. In the case of this example, the person was simply not righteous in the first place.

This realization shifts the debate toward Compatibilism versus Incompatibilism, where I should argue for the Classical Compatiblist case, who pose the view that "you are free if your actions conform to your character". Succinctly, Incompatibilists argue that free will and determinism cannot coexist, whereas Compatibilists believe they can coexist, redefining freedom as acting on your own desires without external coercion.

To defend a Physicalist account of agency, one must confront the Incompatibilist Argument. This perspective suggests that if Determinism is true, meaning every brain state is caused by a prior physical state, then the "alternate possibilities condition" for free will is violated. The argument follows a strict logic: if I could not have done otherwise in the exact same physical circumstances, I cannot be free.

As a Physicalist, I reject this premise because it inadvertently invites the "logical fallacy" of Indeterminism discussed previously. If my decision to put down my phone was not determined by my internal states (my values, history, and goals), then it would be a random event rather than a "willed" one. For a choice to be mine, it must be determined by who I am. Therefore, rather than viewing Determinism as an impediment, I view it as a necessary framework for authorship. The criteria for free will should not be when they are "uncaused", but when the opposite is the case, **when I am the cause.**

This leads to the Classical Compatibilist view: "You are free if your actions conform to your character". In this framework, freedom is not the ability to break physical laws, but the absence of external constraints. However, applying this to phone addiction reveals a significant limitation. Does the compulsive scrolling of an addict "conform to their character"?

Initially, one might argue that the addict chooses to scroll, and thus it reflects their character. Yet, the lived experience of addiction and the "impediment" it represents suggests a breakdown of this model. If my "character" is to be a productive student, but my "body" is driven by a dopamine-seeking mechanism that overrides my goals, then there seem to be a contradiction, and that the Classical Compatibilist definition fails to account for internal impediments. It fails to distinguish between a healthy "desire" and a "compulsion" that has hijacked the biological machine.

To resolve the tension between biological determinism and personal agency, I turn to Donald Davidson's landmark 1963 and 1970 paper [3, 4]. Davidson argued for the idea that while every mental event is token identical to a physical brain event, they belong to different "explanatory schemas." While the physical schema is governed by the causal laws of chemistry and dopamine release, the mental schema is governed by logic and rationality.

In this framework, an action is explained by a Primary Reason, which consists of a "pro-attitude" (a desire) and a "cognitive link" (a belief that the action satisfies that desire). Applying this to the struggle of phone addiction, I find myself weighing competing reasons:

1. Reason A (Compulsion): A pro-attitude toward dopamine release coupled with the belief that checking a notification will satisfy it
2. Reason B (agency): A pro-attitude toward living a fulfilling life coupled with the belief that putting the phone away leads to better use of time

The action that is executed is caused by the specific belief–attitude pair that, in the context of the agent's deliberation, actually moves the agent to act. While other reasons are considered, only the operative primary reason serves as the cause. This allows us to address the failure of Classical Compatibilism: in Davidsonian terms, a compulsion occurs when a physical cause (the dopamine spike) bypasses the "logical" deliberation of the mental schema.

This allows me to put forward a more precise thesis regarding free will, that **my actions are free when I am able to allow my Primary Reason to determine my physical movements**. When my pro-attitude toward productivity successfully directs my hand to put down the phone, I am exhibiting agency, where high-level logic of my character successfully governs the low-level chemistry of my brain. In effect, it becomes apparent that the optimal framework for explaining phone addiction aligns with that of the non-reductive physicalist. Again, I should restate, to avoid the "ghost in the machine" description, these "reasons" must not be viewed as immaterial substances. Instead, following Davidson, they should be understood as mental descriptions of physical events.

To conclude, While I've failed to come up with an account of the mind completely coherent with the everyday explanations of phone addiction, I have found that a non-reductive framework offers the most robust explanation for the "hijacked" will.

As a final food for thought, as technology evolves, so too must our definitions. If we accept that the "self" is an evolving process rather than a static substance, we must confront the blurring boundaries between the biological brain and the digital tools we use to augment it. If my "Primary Reasons" are formed through the data and memory stored in a note-taking app, effectively a "second brain" (as often referred by Obsidian/Notion (popular note-taking apps) enthusiasts), does my agency extend into the software itself? This poses a new challenge for the future of the field: if the mind is a process, does it matter if that process is partially executed on silicon rather than just neurons?

2 References

[1] Wilder Penfield. *The Mystery of the Mind: A Critical Study of Consciousness and the Human Brain*. Princeton University Press, Princeton, N.J., 1975. With discussions by William Feindel, Charles Hendel, and Charles Symonds.

[2] Hilary Putnam. The nature of mental states. In W. H. Capitan and D. D. Merrill, editors, *Art, Mind, and Religion*, pages 37–48. University of Pittsburgh Press, Pittsburgh, 1968.

[3] Donald Davidson. Actions, reasons, and causes. *The Journal of Philosophy*, 60(23):685–700, 1963.

[4] Donald Davidson. Mental events. In Lawrence Foster and J. W. Swanson, editors, *Experience and Theory*, pages 79–101. University of Massachusetts Press, Amherst, 1970.