Is Addicted Phenomenology Just Human Phenomenology?

Benji Mahaffey

Portland State University

UNST 407

Drs. Angela Coventry, Tom Seppalainen, & Toeutu Faaleava

August 4, 2021

The phenomenon of addiction precedes, by millennia, our scientific inquiries into its psychological manifestations and neural bases. We did not need psychiatrists to 'discover' it; we have long been aware of its dark shadow lurking in our psyches. The discernable, often troubling behaviors of addicts notwithstanding, addiction is not the kind of phenomenon one *observes*; addiction is *experienced*, from the first-person perspective. Its defining features are qualitative: a subjective loss of control, an obsession, a compulsion. The overwhelming phenomenological salience of these features—especially of "compulsion"—has led addicts, philosophers, and psychiatrists alike to imagine that addiction is a discrete (phenomenological, natural, psychological) kind. In this essay, I argue that addicted phenomenology is different only in *degree* from ordinary human experience, and not in kind. This conclusion's major premise derives from our present-best neuropsychological theories of addiction; together, they successfully explain the defining phenomenological features *of* addiction while applying, *mutatis mutandis,* to garden-variety experiences of compulsion. I distinguish between essential and contingent features of addicted phenomenology, which allows for ontological clarity around what it is, exactly, we are talking about when we talk about addiction and demonstrates that the identity-criteria of what we call "addiction" are not phenomenally immanent, so to speak, but are external, contextual and—ultimately—contingent.

I. Introduction. Nearly half a century has elapsed since Thomas Nagel first published his landmark essay "What Is It Like to Be a Bat?" The demarcation criterion contained therein is brilliant in its ontic simplicity: an organism "has conscious mental states" just in case "there is something that it is like to *be* that organism—something it is like *for* the organism."[1] Whether one learned of it first in 1974, reading the *Philosophical Review*, or this side of the millennium in *Mortal Questions*, the conceptual heft of the 'what it is like' criterion strikes the adequately-primed reader as so plausible, so intuitive, that perhaps *we* might have articulated the divide between the mental and the non exactly thus.

In my admiration for Nagel and his chiropteran opus I am in the fine company of the philosopher Owen Flanagan. We have more in common than a refined taste in philosophical literature, however. Pertinently—and to our mutual detriment—we share also a fascination with different

[1] Nagel, "What Is It Like to Be a Bat?," 166.

chemical molecules and their corresponding phenomenological effects on consciousness: why does (say) morphine ingestion produce a sort of euphoric sleepiness, like being in a warm bath, while a hit of crack cocaine superimposes upon consciousness a blisteringly intense, all-salient ecstasy? For Flanagan and myself, the appeal of questions such as these at some point transcended the detached abstraction of the philosopher. Although our drugs of choice differed, we both eventually watched, helpless, as our predilection for the effects of some molecule or other seemed to rob us of our agency. We experienced a fracturing of our very selves across the fault lines of our volitions. To wit: Flanagan is a drug addict. So am I.

What is an addict-cum-philosopher, armed with the Nagelian criterion and some modicum of sobriety, to make of his experience? In Flanagan's case, he chronicles the *phenomenological* character of his addiction. The title of his compelling essay surely wrote itself: "What Is It Like to Be an Addict?" Indeed. Riffing on Nagel, he briefly covers the conceptual 'greatest hits' of "What Is It Like to Be a Bat?" (e.g., the brute impenetrability of the first-person perspective by the third, the fundamental irreducibility of subjectivity, etc.). For Flanagan, however, the Nagelian problem is intractable only in the case of *interspecies* communication. The phenomenological chasm between human and bat is unbridgeable in principle, but "we [humans] are members of the same species, and we are language users, so there are prospects for mutual understanding in the case of addiction. So let me tell you a bit about what it was like [to be an addict]."[2]

Of course, we *can* communicate our subjective experiences—even those concerning chemically-altered mental states—to people who have not, strictly speaking, *experienced* them. This is a

[2] Flanagan, 274.

truism borne out across entire genres of literature. I take no issue with the possibility of memoir *as such* and—where Flanagan functions as memoirist—I bring no criticisms to bear. The fact of this possibility (evidenced by its *actuality*) suggests to me that Flanagan is not puritanical in his Nagelianism. I cannot, no matter how eloquent my prose, communicate the redness of red to someone without sight; if a lion (or a bat) could speak, I would not understand him. How is it that 'what it is like' to be an addict is not similarly ineffable?

Addicted experience is in principle communicable because, I will argue, its constitutive, functional properties *just are* universal features of human phenomenology. I intend to use Flanagan's work as a stalking-horse with which to get my own case off the ground. Thus, I begin with a sketch of his essay, his philosophical conclusions, and their implications for our conceptual models of addiction. Dispensing with Flanagan's ontological excesses, I export only those entities which I deem essential from his account and provide a philosophical articulation of each. I proceed by contrasting their structural relations to those of the relevant neuropsychiatric-theoretical frameworks. In so doing, I find that these entities—which Flanagan supposes to be jointly constitutive of phenomenological addiction—*just are* variations of an ordinary phenomenological theme and attempt to follow the error to its roots. I conclude with the conjecture that our addicted-phenomenological reports are tacitly informed by prevailing 'folk' models of addiction; an addict's introspection is just another form of observation, and as such, it is theory-laden.

II. Phenomenological taxonomy. Do we really need a phenomenology of addiction? Ought we not leave such accounts to literature, to music, or to the addicts themselves, within the rooms of Alcoholics Anonymous? Flanagan believes that we do; he makes the case that even *with* a completed neuroscientific theory of addiction—once we've exhausted all the

neurocomputational facts about brains and different molecules' effects on them—no "full accounting" of addiction can be complete absent the texturing of the phenomenological. [3] To accomplish this texturing, Flanagan recommends his "Natural Method," a holistic, coherentist amendment to scientific inquiry. His holarchy requires a phenomenological 'scaffolding' upon which to attach our scientific theories. Furthermore, it requires a degree of commensurability *between* the entities of our competing conceptual frameworks.

He writes: "Suppose that the neuroscience of addiction revealed that all substance addictions involve exactly the same neural processes. I would say *keep looking* because the recognizable differences in phenomenological feelings produced by different drugs are robust among the people who use them."[4] No conceptual "level" (e.g., neuroscientific, psychological, phenomenological), he writes, is "trump"—each plays a necessary epistemic role—but I take his requirement that neuroscience "keep looking" to mean that our scientific theories of addiction must, in a sense, function as explanations *for* the qualitative hallmarks of addictive phenomenology.

Very well. I am sympathetic to Flanagan's requirement that our different conceptual schemas play nicely with one another. Although his no-trump holarchy turns out to be a hierarchy after all—with phenomenology as its arbitrating strata, no less—it is plausible enough to insist that our scientific theories, if they are to be good for anything, function as explanations for *why* it feels like something to be addicted that is different from what it is like to (say) crave ice cream but refrain from partaking.

[3] Flanagan, 288.
[4] Ibid., 286. Emphasis my own.

Although he offers no explicit model for successful neuropsychological explanations, Flanagan's Natural Method requires of our (scientific) explanantia that they be explanatorily *relevant for* or structurally analogous *to* our explananda (which, in this case, are the puzzling experiences central to addicted phenomenology). I will refer to this as the *commensurability requirement.* The commensurability requirement entails that our explanandum must include addiction's *phenomenological* features; thus, our scientific explanans requires that we sufficiently articulate just what, exactly, said features *are*. Such an articulation is not *assumed* (or given) so much as it is *created*;[5] thus, the (eventual) commensurability between our scientific explanantia hinges on, as Flanagan argues, getting the phenomenology *right.* He believes our explananda are best-developed within a taxonomically-structured articulation of addictive phenomenology. It is here where daylight begins to emerge between his position and my own.

"A complete phenomenology of addiction," he writes, "would need… 'what it is like' reports on… $\text{Addict}_{\text{alcohol}}$ and $\text{Addict}_{\text{benzodiazepines}}$ and $\text{Addict}_{\text{opiate}}$." These taxa can be further subdivided into "$\text{Addict}_{\text{Xanax, lorazepam, clonazepam}}$ and $\text{Addict}_{\text{morphine, codeine, heroin}}$."[6] Flanagan reasons that each of the above-listed molecules has a unique phenomenological feel—each has its own distinct or semi-distinct suite of qualitative effects on consciousness—and thus, it is like something to be addicted to alcohol that is different from what it is like to be addicted to methamphetamine. Is it the case, however, that a complete account of *addiction* must be informed by such a breadth of possible alterations to consciousness? Is the qualitative difference between a heroin and a cocaine high relevant to the experience of being *addicted*, or merely the phenomenology of *getting high*?

[5] Or so the saying goes. See Lakatos, "Falsification and the Methodology of Scientific Research Programmes."
[6] Flanagan, 272.

Flanagan acknowledges this conceptual contrast: "There is something it is like to be an addict, or what is different, to be addicted, to be 'high,'" but makes too little of its importance.[7] His is an ontology of potentially infinite discrimination, and its wealth of possible subjectivity is its undoing. There is, *of course*, something it is like to be (say) a lifelong alcoholic, finding himself yet again incapable of declining the second drink. His experience is (necessarily) phenomenologically unlike that of the novice heroin user who does not understand why or how she pawned a treasured heirloom to afford more junk. The constitutive identity of any chemically-altered mental state is composed of a set of *n* factors, almost all of which (e.g., the personal history of the user, their present environment, arbitrarily minute, quantitative variations of the drug ingested) are nonessential to addiction *as such*. So thorough an ontology cannot make useful distinctions between *types*—addiction, say, versus non-addiction—and swells into a trivial mapping: one token per type. That is not to say that such distinctions are of no importance at all—but they cannot serve as a foundation for or arbiter of scientific theory. In order to tease signal from noise, we must ask: what properties are constitutive of addiction? What properties are essential to it, and which are mere accidentals, contingent and—ultimately—extraneous?

III. Constituents of addiction. Luckily, neither Flanagan nor myself requires the full scope of his unwieldy taxonomy in order to pick out universal features of addiction. Each of the (what I call) *contingent* aspects of addictive phenomenology (e.g., context, personal identity, drug of choice) he cordons off under the heading "addiction-2," which incorporates all those variables that differ across "individuals" and "communities of users."[8] Per Flanagan, every addict-2 has in

[7] Flanagan, 274.
[8] Ibid., 283.

common with every other two "features" that are "invariant."[9] These are (1) a *mental obsession* and (2) a *physical compulsion* (jointly, "addiction-1").[10]

Mental obsession seems to be the kind of experience that most—perhaps even *all*—addicts can testify to having. Flanagan articulates it thus: mental obsession is what it is like when "you are always thinking about where, when and how you will next use."[11] This feature of addictive phenomenology—a near-constant directedness of the mind toward a drug or drug-seeking—seems at least a contender for a constitutive property of addiction. It has, after all, a close counterpart in standard psychiatric practice;[12] although the DSM-5 focuses almost exclusively on *behavioral* symptoms in its diagnostic criteria for substance-use disorders, its single mentalistic criterion can be interpreted as an *obsession*, though the Manual dubs it "craving."[13]

Of course, one can obsess over anything. Perhaps a mental obsession is a necessary property of addicted phenomenology, but it can hardly be sufficient. In his monograph *The Disordered Mind*, the philosopher of psychiatry George Graham mentions "obsessions" nearly fifty times, and in each case is considering a mental phenomenon other than addiction. By itself, "obsession" could just as easily cover a jilted ex-lover whose waking mental life is consumed by thoughts of his former partner. It should be obvious that "addicted," though it may apply to him roughly, is not the most elucidatory label with which to conceptualize his problem. Thus we require a second property—physical compulsion—to round out a set of jointly-sufficient

---

[9] Flanagan, 283.
[10] Ibid.
[11] Ibid.
[12] Here I use the term "practice" instead of "theory." This decision acknowledges the atheoretical pretensions of the DSM (see discussions in Faust and Miner, "The Empiricist and His New Clothes." or Follette and Houts, "Models of Scientific Progress and the Role of Theory in Taxonomy Development.") while still permitting us to use analogous features between phenomenology and diagnostic criteria as, if nothing else, a useful heuristic.
[13] American Psychiatric Association, "Substance-Related and Addictive Disorders." Criteria 4—craving—manifests in "an intense *desire* or urge for the drug that may occur at any time." It is plausible that "craving" and "mental obsession" share a common ostension.

phenomenological features that, taken together, are constitutive of what we are talking about when we talk about "addiction."

Physical compulsion is the experience wherein, Flanagan writes, "the car turns into the liquor store, you do not; the drinking is done… but in some sense *it is not me*."[14] This is the experience of the addict from which emerge the tropes of our folk conceptualization of addiction: powerlessness, neural hijackings,[15] drinks taking drinks, and so on. Much like mental obsession, taken by itself the compulsion criterion could describe any host of ticks, reflexes, and involuntary behaviors. When both criteria are present, however, we are talking about a bona fide *addiction*. For Flanagan, their simultaneous presence in consciousness cumulates in a "P & ~P performative inconsistency"—wanting to do one thing, but finding himself *unable* (in a particular context) to shake his obsession and regain authorial *control*.

Of the two, the phenomenology of the compulsion criterion is the harder to tidily pin down and articulate. ("Mental obsession" is nearer to being an *intuitively* universal experience, so long as the intentional objects "drug" and "drug-seeking" are replaced by ones with more quotidian referents.) Controversial—and harder for the nonaddict to understand—is the claim of the addict that he could not really *help* himself when he decided to use. What do we addicts mean when we speak of experiences like *losing our volition* or *having no control*? A remarkably apt account of an addict's physical compulsion comes from Harry Frankfurt's classic "Freedom of the Will and Concept of a Person."

According to him, as an addict, I have conflicting first-order desires. (First-order desires could be said to provide the immediate causal 'thrust' to my *actions*; per Frankfurt, they constitute my

[14] Flanagan, 280.
[15] Ross et al., *Midbrain Mutiny*, 181.

will). Humans and persons are not unique in having first-order desires; a dog may have a first-order desire to chase a cat, and a laboratory mouse may have a first-order desire for the cocaine-laced water bottle. For the nonhuman animal, the first-order desire merely litigates *what she will next do*. Humans, however, have a brand of 'free will,' which—for Frankfurt—*just is* the ability to *choose* which of their first-order desires becomes, ultimately, their will.[16] The (nonaddict) agent has the first-order desire to *go for a run* but also the conflicting first-order desire to *sit on the couch*, and it is in the case of competing first-order desires that second-order volitions come into play. The nonaddict *identifies* with the desire to *desire* to go on a run; he *wants* to want to run and *does not want* to want to sit on the couch. Volitions situated thusly, his ultimate *behavior*—going on a run—is wholly his own.

Not so for the addict. I may have the first-order desire to obtain and use heroin. I also have a first-order desire (so I report) to *not* obtain or use heroin. (I want to P and to ~P.) Now, I do not *want* to be an addict. Rationally, I understand that the consequences of my use (e.g., imprisonment, interpersonal conflict, death) will directly conflict with others of my long-term desires (e.g., to get married, to become a philosopher). Suppose, then, that I identify with my second-order volition *to want to want to ~P*.[17] I consider myself to be, fundamentally, the kind of person who *does want to get married* and *does want to become a philosopher*, not the kind of person who goes to jail or overdoses. *Who I am* is a person who *wants to want to ~P*. When I find myself P-ing, I am—justifiably—confused and surprised, and if it becomes a pattern in my life, I might become hopeless and utter phrases like "I am powerless to stop using heroin." If I am philosophically inclined, I may seek to develop a phenomenology of addiction for which the

[16] Frankfurt, "Freedom of the Will and the Concept of a Person," 14.
[17] There is a potentially vicious regress lurking in Frankfurt's ever-ascending volitional hierarchy, which I will not discuss here. See Ekstrom, "A Coherence Theory of Autonomy" for one possible solution to it.

development of at least one criterion will come painlessly: addiction must involve an experience of *physical compulsion* in which my autonomy as an agent is frustrated. P & ~P.

With the articulations of the compulsion and obsession criteria sketched above, we arrive at a tentative pair of phenomena that require explanation. Because nothing about addiction hinges on the contingent phenomenological features of a person's identity or situation—addiction-2—we can discard these entries from our explanandum. Thus purged, we say that a person experiences *addiction* if and only if they experience a *mental obsession* coupled with a simultaneous (if intermittent) *physical compulsion*.[18] I have claimed that it is uncontroversial that obsession occurs in nonaddicts; the concept of "obsession" intuitively admits of variations in degree which never bifurcate into (qualitatively) discrete kinds. Thus it remains to be seen only whether relevant scientific theory can account for the salient phenomenological character of physical compulsion.

IV. Neuropsychological theories of compulsion. What is going on in the brain of an addict when she experiences the "me-but-not-me" who takes the drink or drug, despite her identifying with the desire to abstain? In a seminal contribution to the canon of scientific literature on addiction, Robinson and Berridge propose a motivational model with discrete neural mechanisms for two functions that are typically taken to be virtual synonyms: their "incentive-sensitization" theory of addiction distinguishes, pertinently, between 'wanting' and 'liking.'[19]

[18] What about the hypothetical nun, who over the course of many months becomes "addicted" to the painkillers her doctor prescribed? Here it is useful to differentiate between 'addiction' and (physiological) 'dependence.' See O'Brien, "Addiction and Dependence in DSM-V" for a thorough articulation of the two phenomena. Roughly, 'dependence' is a "physiological adaptation" by the body in response to a drug. Dependence and addiction may and often do overlap, but they need not.

[19] Robinson and Berridge, "The Neural Basis of Drug Craving"; Berridge and Robinson, "Liking, Wanting, and the Incentive-Sensitization Theory of Addiction."

Namely, *wanting* and *liking*—though they can be and indeed often are intimately linked as far as their objects go—need not *necessarily* be. Mice, who have been artificially stimulated to produce an excess of dopamine, are motivated to obtain some prize foodstuff (they 'want' it) despite showing none of the usually empirical markers of enjoying ('liking') it once it has been obtained. The converse holds as well. Mice who have been artificially deprived of dopamine may like a prize very much, but remain entirely unmotivated to expend energy in seeking it out.[20] Mesolimbic dopamine functions such that it 'tags' the environmental affordances available to the mouse *as* rewards. Meanwhile, the neural 'liking' function (called "hedonic reaction" in the literature) occurs in different neural regions in the nucleus accumbens and ventral pallidum; rather than responding to dopamine, these "hedonic hotspots" are sensitive to neurotransmitters associated with opioid and endocannabinoid systems.[21] What is good for the goose is good for the gander: the decoupling of the 'wanting' function from the 'liking' observed in mice is evident too during self-administration experiments with cocaine addicts, who continue to 'want' cocaine despite ceasing to 'like' it during a session.[22]

The cocaine, the heroin, the alcohol—these are incentive-salience desires, triggered via conditioned cues in an addict's environment and tagged as rewards by mesolimbic dopamine. They are a- or pre-rational. Dill and Holton contrast these semiautomatic, at least partially subconscious desires (which may or may not be 'liked') with what they call *cognitive desires*. Cognitive desires are phenomenologically distinct from their incentive-salience kin. They are

[20] Dill and Holton, "The Addict in Us All," 2-3.
[21] Anselme and Robinson, "'Wanting,' 'Liking,' and Their Relation to Consciousness." Animal models cannot, of course, tell the whole story of addiction, but the empirical evidence in favor of this partition is nevertheless robust. See Graham, *The Disordered Mind,* pp. 154-56 for a measured review of the strengths and limitations of animal models of addiction.
[22] Fischman and Foltin, "Self-Administration of Cocaine by Humans."

derived rationally from a set of propositions; their objects are generally not the kind of things one craves; one is never 'physically compelled' to obtain one's cognitive desires.

This is because they are "bound up" with *reasons*;[23] I want to become a philosopher because I believe I will become rich and famous. When, to my chagrin, I discover that professional philosophers are categorically poor, unemployed, and unknown, my *reason* evaporates and the desire withers. It requires no strength of will or authorial exertion on my part to abolish the desire. This is not the case for my desire for a cigarette, for heroin, or a drink.[24] Whatever reason motivated *their* initial wanting, the long shadow of desire can persist well after the logical heft of the reason has been extinguished.

The apparent paradox in the addict's experience of physical compulsion is resolved when understood in light of—and, thus *mapped to*—the (hypothetical, functional) structure which emerges from the neuropsychological channels postulated by Robinson's and Berridge's incentive-sensitization theory of addiction. According to the finer-grained, Franfurtean-derived articulation of the phenomenology of *physical compulsion*, the "me-but-not-me who takes a drink" is unable to *choose* which of their first-order desires becomes their will. My account leads to a reframing of Frankfurt's "second-order volitions" as cognitive, reason-bounded desires. I want to want to ~P *because I believe* something or other *about* ~P-ing (e.g., I want to want to ~P because I want to become a philosopher *and* I believe that ~P-ing is a necessary condition for my aspiration's realization).

Unlike my desire to P, my cognitive desire is not mediated by dopaminergic reward-scheduling. Ultimately—contra Frankfurt—I *have* no conflicting first-order desires. I have only the first-

[23] Dill and Holton, "The Addict in Us All." 4.

order desire Σ to P (I 'want' cocaine in the sense that I am extraordinarily motivated to seek and use it) and the cognitive desire (the "second-order volition") Φ to want to want to ~P. There is, however, no first-order desire Ψ "wanting to ~P" upon which my cognitive desire Φ can attach. However—especially when I 'want' to P but do not especially 'like' P-ing—the salience of Φ leads to my mistaking it for Ψ.

V. Is addiction a discrete kind? From Flanagan's taxonomy, we salvaged a skeletal definition of phenomenological addiction. We articulated its most controversial component, *physical compulsion,* and proceeded to posit a commensurable neuroscientific explanans. In so doing, however, have we deprived "addiction" of its uniqueness? Is compulsion, in the sense outlined above, experienced only during addiction proper? Or is it the sort of experience that humans have *all of the time,* in ordinary, nonaddictive contexts?

Consider the following: coffee makes my wife anxious. She tells me—and, by proxy, herself—that she is taking a break from it, starting tomorrow. And yet, when next the sun rises, she is back at the pot, mug in hand. Is her experience best described in terms of *physical compulsion*? Does her behavior constitute an insoluble paradox? Caffeine, of course, is an addictive substance, so perhaps we need a more theoretically-innocent example. I want to write a paper on addiction; instrumentally, I know that the consequences that are likely to follow the publication of a well-argued position are *things that I want*. And yet, I find myself drawn instead to cognitively-undemanding tasks: reading the news, checking my phone, responding to email. How is it that I can—even *in the moment*—consciously desire that some state of affairs come to pass, and yet do the opposite of what will bring it about?

The experience is familiar to everyone. We want to have the salad but reach for the potato chips; we want to enjoy a small piece of chocolate, but devour the bar. We want to read *The Critique of*

Pure Reason but end up on the couch, binge-watching *Game of Thrones*. If we have not been taught to speak of our experience in terms of *compulsion* or *loss of volition*, then we may simply label these discrepancies between our desires and our behavior as lapses in willpower. Qualitatively, however, they *feel* the same (to me) as the 'me-but-not-me' who takes the drink or the drug.

The original research around incentive salience and sensitization may have been initially interested in mechanisms responsible for drug addiction, but its primary discovery was that 'wanting' is mediated entirely via mesolimbic dopamine and is only contingently related to 'liking.'[25] All drugs of addiction act on dopamine production in the brain (though individual mechanisms vary across pharmacological profiles), so incentive-sensitization can account neatly for the phenomenon of *drug* addiction.[26] However, it is uncontroversial that the symptoms which manifest across *behavioral* addictions—like to gambling, to gaming, or even to checking one's social media feed—are virtually indistinguishable from those manifesting across drug addictions.[27] These 'rewards' are similarly mediated by dopamine. There may be and likely is conceptual fine-tuning that remains necessary for our empirical research into how the brain behaves when it is addicted to *World of Warcraft* as opposed to fentanyl, but with functionally similar inputs leading to functionally homogenous symptoms, signs point to the difference being one of degree and not of kind. In both cases, an agent is compelled to want something with a fervor that is out of alignment with her liking of it.

If, however, addiction—or any of our incentive-salience derived, 'wanting' behaviors—is a matter of unconscious, dopamine-fueled animal impulse, how do we get anything *done*? Because

[25] Anselme and Robinson, "'Wanting,' 'Liking,' and Their Relation to Consciousness."
[26] Ibid.
[27] Griffiths, "Is 'Loss of Control' Always a Consequence of Addiction?"

our higher selves do, occasionally, prevail. We pick the salad. We write the paper. Why can we addicts overcome our base desires where they concern the mundane, but experience an overwhelming compulsion when presented with a drink or a drug? Firstly, it is not the case that addicts *always* are at the mercy of their unwanted first-order desires. Even the heroin addict in the grip of her addiction will at times exhibit temporary control, perhaps when her infant daughter is in the car with her or before a crucial interview with her probation officer. Once again, a single explanatory mechanism accounts for the abstaining addict and the successful dieter.

Baumeister et al. posit a functional entity: willpower as a finite cognitive resource.[28] Their *ego-depletion theory* is borne out again and again in experiment. Subjects are less likely to choose a healthy food after having already exerted self-control during a prior task; they are more likely to ignore a stranger in need of help after having behaved altruistically in the recent past.[29] Importantly for the case of the addict, long-term desires—the more cognitive, reason-bounded desires—are also more difficult to pursue for ego-depleted subjects.[30]

The precise neural correlates of the willpower resource require further empirical research. The method in which it functions is also up for debate: Neil Levy, informed by George Ainslie's theory of *hyperbolic discounting* referenced in "What Is It Like to Be an Addict?" argues that the "neural adaptations characteristic of addiction provide a mechanism for [shifts in judgment]."[31] I take Levy to mean that addictions—behavioral or substance—involve a pathological rewiring of an agent's perceived reward schedule.[32] When an addict is triggered by some environmental or

[28] Baumeister et al., "Ego Depletion."
[29] Dill and Holton, "The Addict in Us All," 7–8.
[30] Ibid.
[31] Levy, "Addiction as a Disorder of Belief." 339-340.
[32] In so doing, I follow Dill and Holton, "The Addict In Us All," 9.

physiological cue, the brain attempts to modify the perceived *expected utility* of some affordance. The mirrored medicine cabinet triggers the addict to reevaluate its possible contents as more valuable than they otherwise would be.

Self-control or willpower, then, functions as a kind of firewall against this rescheduling process. With it, the schedule remains unchanged. In the ego-depleted subject, preferences are modified; the agent *prefers* the immediate reward in the medicine cabinet to the more abstract, temporally-distal ends whose realization requires her abstinence.

Wallace's competing account of the willpower mechanism sees it as primarily *motivational*. In it, an agent can "[reason] correctly" but still fail to "comply with the deliberated verdict."[33] Here I side with Wallace. It is my belief that only by conceiving of willpower as a limited *motivational* resource that we can (roughly, metaphorically) map ego-depletion theory onto addicted phenomenology. This is because addicts do not experience temporary *irrationality* whereby we briefly imagine our drug of choice to have higher utility than it does. Likewise, the recalcitrant dieter does not mistakenly evaluate the bag of potato chips as being a healthier option than the salad. In both cases, an agent's reasoning faculties remain intact; it is just that we are unable to pursue what we evaluate as being in our ultimate best interest. Wallace's motivation hypothesis is structurally analogous to the relevant phenomenology of compulsion, where Levy's is not; therefore, only the former fulfills the commensurability requirement.

In this section, I attempted to demonstrate how the most unique of addiction's phenomenologically-salient features—the *physical compulsion*—is distinct from ordinary lapses

[33] Wallace, "Addiction as Defect of the Will," 639.

in willpower in degree only and not different in kind.[34] In the final section of this essay, I propose that the (superficially) apparent distinction between "addictive" compulsion and ordinary phenomenology stems from the theory-laden nature of introspection. The hermeneutical resources available to the addict are derived from the prevailing popular, historical, and medical conceptualizations of addiction.

VI. Theory-ladenness of addictive phenomenology. If addictive phenomenology *just is* garden-variety human phenomenology, why is it that (even, especially) those of us who have *experienced* addiction testify to a fracturing of our agency? Why do we speak so readily in terms that are unlike those used to describe ordinary 'lapses' in willpower? I hold that we speak of addiction in these terms, frankly, because that is how we are *taught* to speak.[35]

Over the past century, there have been two leading conceptual models of addiction. The first is a folk-psychological theory: the moral model of addiction identifies the consumption of a drug as an agent's choice, and an immoral choice to boot.[36] We—addicts and alcoholics—get high because we *want* to, devil take the rest, and the model condemns the bad actions of the weak-willed agent. The theoretical influence of this model is immediately obvious in our social and legal institutions; possession of most drugs is a felony offense, the people who use them are irresponsible and dangerous, and they belong—and usually end up in—jail.

---

[34] Perhaps there is a 'quantity' of the willpower resource required for quotidian tasks that is lesser than the quantity required to, e.g., turn down a drug of addiction. In the explicitly speculatory mode, I suspect something like this to be the case. This might allow us to differentiate between addictive compulsion and non despite their being of a kind.
[35] We need not be card-carrying eliminativists to forward this thesis. When an addict reports her experience of "physical compulsion," we need not insist that "compulsion" lacks a referent or is a folk-psychological, Churchlandian chimera. We need only maintain that the veracity of introspective reports is susceptible—at least in principle—to error. See Schwitzgebel, "Descartes Inverted," for a discussion of introspection and fallibility.
[36] Pickard, "What We're Not Talking about When We Talk about Addiction." 37.

Armed with good intentions, and recognizing the predicament of many addicts to be a far cry from the hedonistic lifestyle of the idealized Dionysian reveler, treatment centers, psychologists, and neuroscientists alike postulated a new, more humane alternative: the diseased-brain model, a common variation of which sees drug addictions as *always* a behavioral manifestation of what is essentially a dysfunction of the brain.[37] In eschewing the moral model, the diseased-brain model allows for a vindication of addicts. Unlike the weak-willed truant who *could* have picked herself up by the bootstraps and gone to work—she *could have done otherwise*—the addict suffers from a medical incapacitation of her agency.[38]

It is natural that, as addicts, we find the diseased-brain model to be the more palatable option; it removes much of the personal culpability involved, which—in many cases—is no small gift. Nevertheless, the diseased-brain model is almost certainly incorrect. The etiology of a malady such as Alzheimer's involves an actual degradation of neural tissues; as an exemplar, it admits of a wholly physical proximal cause. Behavioral manifestations stem *from* myelinic degeneration, but they are not the kind of symptom that can be remedied with behavioral therapies. Constitutively, Alzheimer's *just is* a brain state involving a degree of axonal or myelinic degeneration.[39] Addiction, it must be acknowledged, is not like this. Though it may be physically *instantiated* in the brain, its proximal cause is not best conceptualized in terms of somatic pathology.[40]

Lacking an established phenomenological vocabulary with which to describe our experience, addicts' recourse is to (tacitly) adopt a kind of jury-rigged hodgepodge, borrowing terms from

---

[37] Graham, *The Disordered Mind*, 156. This too is the etiology that the DSM-5 uses to justify its taxonomic schema.
[38] See Levy, "Resisting 'Weakness of the Will.'" for a discussion (and ultimate *dissolution*) of the distinction between weak-willed agents and compulsive agents. The diseased-brain model puts addicts in the latter category.
[39] Salvadores et al., "Axonal Degeneration during Aging and Its Functional Role in Neurodegenerative Disorders."
[40] Graham, *The Disordered Mind*.

both the moral and medical models. Thus we speak of our experience as being like a hijacking, or as being similar to that of the diabetic; paradoxically, the experiences are simultaneously moments of great moral turmoil—we experience shame, guilt, remorse. The prevailing (folk) theoretical framework involved in the *treatment* of addiction—mutual help organizations like Narcotics and Alcoholics Anonymous—is in part responsible for institutionalizing and validating this bizarre 'ontology.'

From the vantage of its theoretical amalgam, addiction is, on the one horn, a spiritual malady emerging from a person's selfish refusal to rely on a higher power.[41] On the other, it is an incurable, progressive, and ultimately fatal somatic illness which can be arrested but never cured.[42] The stockbroker-penned 'Big Book' of Alcoholics Anonymous, and the heady nostrum contained therein, recommends a diversity of treatments varying from the confessing of one's sexual deviancies to another alcoholic, to the more traditional faith-healing palliatives of prayer and supplication.[43]

For my purposes, it is enough to point out the text's relentless emphasis on *powerlessness*. There is little conceptual daylight between the phenomenological reports of powerlessness and those of physical compulsion. Where difference emerges, it is causal; where the latter is agnostic, the former is bolder—powerlessness follows necessarily from godlessness (or self-will). Either, however, can be described in Flanagan's parlance: P & ~P.

---

[41] *Alcoholics Anonymous*. 62.
[42] Ibid., 30-33.
[43] Here I am perhaps too eager to grind my own axe. The (Jamesean-endorsed) 'Big Book' of Alcoholics Anonymous proposes what seems to me a fantastical 'theory' of addiction. Nevertheless, the "Minnesota model," as it is called, seems to work for some addicts. See Schnabel, "Neuroscience: Rethinking Rehab" for a sympathetic (if uncritical) history of Wilson's therapy; see Peele, "AA and Abstinence as Prevention Techniques" for a blunter indictment.

VII. Concluding Remarks. In this essay, I have argued that the phenomenological character of addicted experience exists on a spectrum continuous with ordinary human phenomenology. I have relied extensively upon Owen Flanagan's excellent work as a foil for my own case. In so doing, I first followed his Natural Method, the relevant posits of which I formulated as the *commensurability requirement*. The commensurability requirement stated that scientific explanantia must structurally analogous with corresponding phenomenological explananda.

The requirement proved too keen a razor when applied to the excesses of Flanagan's proposed addictive ontology; thus, in recovering from it a coherent, maximally-simple definition of "addiction" I salvaged just two necessary phenomenological conditions, a *mental obsession* and a *physical compulsion*, which I took to be jointly constitutive of the phenomenology of addiction. Of these, I deemed only *physical compulsion* in need of (theoretical) explanation.

I articulated the compulsion criterion according to Harry Frankfurt's account of first-order desires and second-order volitions, and submitted evidence supporting the incentive-sensitization theory of addiction. I deemed it a satisfactorily-commensurable explanation of the phenomenological character of addicted compulsion; I then made the case that the theory's posited neuropsychological mechanisms explained also ordinary 'lapses' in willpower, an experience common to addicts and nonaddicts alike. Ego-depletion theory and its construal of willpower as limited motivational resource completed the explanation in a manner relevant for the actual phenomenology of compulsion as well as ordinary lapses.

Finally, I acknowledged that my account entails the existence of a series of egregious errors in our phenomenological reports of addicted experience. I proposed a mechanism of (folk) theory-ladenness, in which a vocabulary of addicted introspection is tacitly informed by prevailing and historical models of addiction.

A host of criticisms could be levied against the case I have outlined above. I will address just one. It is the pragmatic concern that—if I am correct and that addiction is best explained in terms of a phenomenology continuous with ordinary human experience—that we risk backsliding, returning to a model of addiction that sees addicts as weak-willed agents and moral failures rather than the foci of complex, intersecting causal factors cultural, neural, and psychological.

This fear is plausible only insofar as we mistake our own ordinary experience of free will and volition for being *more free* than it in fact is. Furthermore, the account of ego-depletion I've given leaves open the possibility that addiction is of a kind with garden-variety lapses in willpower while still admitting of some bifurcation, beyond which it becomes physically impossible to overcome one's cravings, even if the mechanism for so doing is functionally equivalent to ordinary volitional exertions.

Bibliography

Alcoholics Anonymous Big Book. 4th ed. New York City: Alcoholics Anonymous World Services, 2001.

American Psychiatric Association. “Substance-Related and Addictive Disorders.” In *Diagnostic and Statistical Manual of Mental Disorders*. DSM Library. American Psychiatric Association, 2013. https://doi.org/10.1176/appi.books.9780890425596.dsm16.

Anselme, Patrick, and Mike J. F. Robinson. “‘Wanting,’ ‘Liking,’ and Their Relation to Consciousness.” *Journal of Experimental Psychology: Animal Learning and Cognition* 42, no. 2 (2016): 123–40. https://doi.org/10.1037/xan0000090.

Baumeister, Roy F., Ellen Bratslavsky, Mark Muraven, and Dianne M. Tice. “Ego Depletion: Is the Active Self a Limited Resource?” *Journal of Personality and Social Psychology* 74, no. 5 (1998): 1252–65. https://doi.org/10.1037/0022-3514.74.5.1252.

Berridge, Kent C., and Terry E. Robinson. “Liking, Wanting, and the Incentive-Sensitization Theory of Addiction.” *The American Psychologist* 71, no. 8 (2016): 670–79. https://doi.org/10.1037/amp0000059.

Dill, Brendan, and Richard Holton. “The Addict in Us All.” *Frontiers in Psychiatry* 5 (October 9, 2014). https://doi.org/10.3389/fpsyt.2014.00139.

Ekstrom, Laura Waddell. “A Coherence Theory of Autonomy.” *Philosophy and Phenomenological Research* 53, no. 3 (1993): 599–616. https://doi.org/10.2307/2108082.

Faust, David, and Richard Miner. “The Empiricist and His New Clothes: DSM-III in Perspective.” *American Journal of Psychiatry* 143, no. 8 (August 1986): 962–67. https://doi.org/10.1176/ajp.143.8.962.

Fischman, Marian W., and Richard W. Foltin. “Self-Administration of Cocaine by Humans: A Laboratory Perspective.” In *Ciba Foundation Symposium 166 - Cocaine: Scientific and Social Dimensions*, 165–94. John Wiley & Sons, Ltd. Accessed May 22, 2021. https://doi.org/10.1002/9780470514245.ch10.

Flanagan, Owen, Jr. “What Is It Like to Be an Addict?” In *Addiction and Responsibility*, edited by George Graham III. Cambridge, UNITED STATES: MIT Press, 2011. http://ebookcentral.proquest.com/lib/psu/detail.action?docID=3339250.

Follette, William C., and Arthur C. Houts. “Models of Scientific Progress and the Role of Theory in Taxonomy Development: A Case Study of the DSM.” *Journal of Consulting and Clinical Psychology* 64, no. 6 (December 1996): 1120–32. http://dx.doi.org.proxy.lib.pdx.edu/10.1037/0022-006X.64.6.1120.

Frankfurt, Harry G. “Freedom of the Will and the Concept of a Person.” *The Journal of Philosophy* 68, no. 1 (1971): 5–20. https://doi.org/10.2307/2024717.

Graham, George. *The Disordered Mind: An Introduction to Philosophy of Mind and Mental Illness*. 3rd ed. Abingdon, Oxon ; New York, NY: Routledge, 2021.

Griffiths, Mark D. “Is ‘Loss of Control’ Always a Consequence of Addiction?” *Frontiers in Psychiatry* 4 (2013). https://doi.org/10.3389/fpsyt.2013.00036.

Lakatos, Imre. “Falsification and the Methodology of Scientific Research Programmes.” In *Criticism and the Growth of Knowledge*. Cambridge: Cambridge Univ. Press, 1970.

Levy, Neil. “Addiction as a Disorder of Belief.” *Biology & Philosophy* 29, no. 3 (2014): 337–55. https://doi.org/10.1007/s10539-014-9434-2.

———. “Resisting ‘Weakness of the Will’: RESISTING ‘WEAKNESS OF THE WILL.’” *Philosophy and Phenomenological Research* 82, no. 1 (January 2011): 134–55. https://doi.org/10.1111/j.1933-1592.2010.00424.x.

Nagel, Thomas. “What Is It Like to Be a Bat?” In *Mortal Questions*, 165–80. Cambridge University Press, 2012.

O’Brien, Charles. “Addiction and Dependence in DSM-V.” *Addiction* 106, no. 5 (May 2011): 866–67. https://doi.org/10.1111/j.1360-0443.2010.03144.x.

Peele, Stanton. “AA and Abstinence as Prevention Techniques.” *Drugs: Education, Prevention and Policy* 19, no. 4 (August 2012): 284–90. https://doi.org/10.3109/09687637.2012.671861.

Pickard, Hanna. “What We’re Not Talking about When We Talk about Addiction.” *Hastings Center Report* 50, no. 4 (2020): 37–46. https://doi.org/10.1002/hast.1172.

Robinson, Terry E., and Kent C. Berridge. “The Neural Basis of Drug Craving: An Incentive-Sensitization Theory of Addiction.” *Brain Research Reviews* 18, no. 3 (1993): 247–91. https://doi.org/10.1016/0165-0173(93)90013-P.

Ross, Don, Carla Sharp, David Spurrett, and Rudy E. Vuchinich. *Midbrain Mutiny: The Picoeconomics and Neuroeconomics of Disordered Gambling: Economic Theory and Cognitive Science*. 1st ed. Vol. 1. MIT Press Books. Cambridge: MIT Press, 2008.

Salvadores, Natalia, Mario Sanhueza, Patricio Manque, and Felipe A. Court. “Axonal Degeneration during Aging and Its Functional Role in Neurodegenerative Disorders.” *Frontiers in Neuroscience* 11 (2017). https://doi.org/10.3389/fnins.2017.00451.

Schnabel, Jim. “Neuroscience: Rethinking Rehab.” *Nature* 458, no. 7234 (March 2009): 25–27. https://doi.org/10.1038/458025a.

Schwitzgebel, Eric. “Descartes Inverted.” In *Describing Inner Experience?: Proponent Meets Skeptic*. Cambridge, UNITED STATES: MIT Press, 2007. http://ebookcentral.proquest.com/lib/psu/detail.action?docID=3338738.

Wallace, R. Jay. “Addiction as Defect of the Will: Some Philosophical Reflections.” *Law and Philosophy* 18, no. 6 (1999): 621–54. https://doi.org/10.2307/3505095.