

The Emperor's Free Will

by Daniel Lärnhem

1 – Introduction	
1.1 – Thesis	1
1.2 – Literature	1
2 – Definitions (or not)	3
3 – On noticing	
3.1 – What is it to notice?	6
3.2 – The argument from noticing	7
3.3 – Unpacking the premises	8
3.4 – Collapsing the subject-object dichotomy	11
4 – Compatibilism and responsibility	
4.1 – Compatibilism	13
4.2 – Responsibility	17
5 – Conclusion	24
Literature	25

1 – Introduction

1.1 – Thesis

Free will has commonly been used in the philosophical community to justify and ground the concept of moral responsibility. I will attempt to argue that this connection is unjustified. I begin in section 2 by establishing a central postulate that free will needs to fulfill in order to be ethically relevant to this topic: Its freedom needs to be subjectively accessible. Placing my conceptualization of free will purely within the subjective domain puts me at odds both with the compatibilists *and* the incompatibilists, thus staking out a somewhat unusual position within this field. To intuitively demonstrate this position, I introduce the metaphor of the Subjective Prisoner, which will be a central reference point throughout the rest of the paper. In section 3, I explore the structure of our conscious minds through the concept of *noticing* to determine if this postulate can be fulfilled, and show that it, in fact, cannot—thereby establishing the impossibility of any sense of subjective freedom. Section 4 compares the resulting view to the compatibilist position on freedom and responsibility (in particular the so-called *rationalist conception*), and argues that the compatibilist use of the term ‘free will’ is an unhelpful misnomer, and that any practically helpful notion of responsibility is in fact grounded not in anything worth calling free will, but instead the tendency for regularity in the behavior of things in the world. As free will dies—or is revealed to be irrelevant—responsibility lives on just fine, revealing the falsity of their connection, and the most common motivation to advocate for the existence of free will dissipates.

1.2 – Literature

The arguments I present in this paper are largely my own, although I have my influences. In

taking the subjective view of our freedom to be of paramount importance I am, at least nearly, alone. Neuroscientist and philosopher Sam Harris is the only one I have found that seems to share this view, and his book *Free Will*, while sometimes embellished with superfluity to the point of distraction, is the only work cited that overall defends a thesis similar to my own. The conceptualization of free will presented here is therefore somewhat unique in the literature, even to the point of not fitting comfortably in either the compatibilist or the incompatibilist camps, although there are pieces of value to be found in both of these frameworks. I take a largely critical stance toward most other sources cited, the main ones being the work done by the philosopher Daniel Dennett, who, in my view, offers the most respectable defense of compatibilist free will I have yet to come across.

I have, in addition to the works already mentioned, tried my best in the available time to get some sense of the wider literature on free will. I would not feign to have fully run dry the well of material on this topic, as I suspect that endeavor would not only be beyond the scope of this paper, but well and proper beyond the scope of a human lifetime. What I have done instead, as I dare to make somewhat general claims about this field of interest as a whole, is to rummage the shallow waters of this pond, not to know intimately every crevice in its depths, but to acquire a general familiarity with the terrain. I therefore read plenty of philosophical articles on the topic, as well as a number of recommended books I managed to get my hands on. Most of this reading did not make it into this paper, but helped me gain some perspective, and to better streamline my points.

As a result of this wide, rather than deep, exploration, my arguments stand to a large degree on their own, and delve not much into the detailed points of the various cited authors, but relate to them in more general ways, often to provide examples to emphasize my points about the

literature.

2 – Definitions (or not)

Free will is most commonly an issue of ethics. It is meant to inform our thinking on judgment, punishment, blame, forgiveness, justice, and mercy as practiced interpersonally and in wider society. This is therefore a topic that should be rooted in ethical concerns, and in the one condition that allows for such concerns: phenomenal consciousness. While other areas of inquiry, such as politics, psychology, or physics, may inform our ethical efforts, consciousness remains its only conceivable *root*—the part which, were it to be removed, would render any situation completely ethically uninteresting. It is therefore surprising to find that much of the literature on free will ground their arguments not in conscious awareness, but in such things as physics, the presence or absence of determinism, some mysterious ability to “do otherwise,” or the like. These topics are surely interesting on their own, but can in the end, I argue, do nothing to determine whether our wills are free or not. That fact should already be attestable here and now, through the actual experience of our wills.

As the root of all ethical concern, consciousness needs to be intimately involved in the freedom meant to justify the ethically aimed responses to that freedom. What would it be like if this were not the case? Imagine someone locked up in a harsh prison, chained tight to the wall in the dark, unable to move in any way. One day, someone comes into his cell to tell him that “Actually, this prison is all in your head. In reality, you are as free as a bird; walking around the outside world with the rest of us.” Firstly, what reason would this Prisoner have to believe these words, when there is nothing in his experience that could possibly attest to their truth? He would be more reasonable in assuming that this person is simply mocking him. But suppose that what

they said were true, he *is* actually walking around the outside world, while it is only his consciousness that is entirely trapped within this prison. While this would certainly be a different situation, would it grant him any freedom? Would this offer any relief to such a Subjective Prisoner?

Moreover, would you feel justified in basing your concepts of desert and blame on such a notion of freedom? If someone came into the Prisoner's cell every now and again to torture him for the misdeeds of his "free" self out in the world, subjectively inaccessible to him, would this be justice? Perhaps through reason he could come to understand that this torture is the only way to mitigate wider harm in the outside world, but even in this case, should he accept the guilt? Should he suddenly agree that he *deserves* the suffering associated with the punishment? Such a conclusion would seem absurd, as the only thing capable of experiencing the punishment has done nothing to deserve it.

This is the central point. If free will is to be a matter informing ethical concerns, then free will needs to be grounded (or at the very least *groundable*) in the same domain. If it's not, then free will ends up having nothing of interest to say about ethics, and as a result nothing interesting to say about any of the topics we care about that ethics informs. If the only thing capable of experiencing the outcome of our ethical efforts has no free or active role to play, then how could it be responsible? How could it ever deserve? How does one justify its suffering based on events occurring outside its domain of influence?

The point I'm trying to express here is that, for the will to be free in any ethically relevant sense, it needs to *feel* free. Given this postulate, another way to ask the question 'Do we have free will?' emerges: Do our wills feel free? These two questions are, I argue, the *very same* question, and one that necessarily admits of an answer through introspection alone, before any

third-person empirical endeavors we might undertake. All matters of physics, causality, or time can only inform, not answer, this question. Whatever is found in these other domains must be compatible with what is already true in our minds here and now—no result in physics found tomorrow will change the facts of our minds today.

If the emperor in the classic tale, parading naked before his people, were said to be wearing metaphysical clothing, imperceptible to any sense or measurement, rather than normal ones, we would be right to disbelieve. Even if it were somehow *true* that he wore such clothing, they would fill none of the functions of actual clothing, and would therefore be undeserving of that label. They would provide no protection, no cover, confer no status, or whatever else we might use clothing for.

If this emperor is imperceptibly clad, then we have every right to call him naked.

The “free” part of free will is tricky to define.¹ Indeed, it is the center of many of the disagreements in the free will literature. It then does not help matters that I have a conception of this freedom that is subjectively rooted, and therefore rather novel in the field. Thus, the purpose of this section has not so much been to provide the reader with a definition of free will *per se*, but rather to set up a central postulate that the nature of the will must fulfill in order to be properly or in any way usefully called “free.” That is; if the will *is* free, it must *feel* free. Again, what this freedom, whatever it is, would entail is a difficult question (likely even an impossible one), and is not my central issue here. In the next section, I will instead explore the structure of the mind as it is necessarily experienced from the first-person perspective, in order to determine

¹ The ‘will’-part is not entirely simple to manage either, for that matter. When using this term, I will generally be referring to phenomena such as intentions, desires, and wants—the general motivations that guide our actions.

whether the will, or anything else in our perception, actually fulfills this postulate. I hope then to determine, not precisely what this quality of “freedom” *actually* is, but whether it can be rooted in the same domain as ethics, and therefore inform our practices of responsibility in any useful way, as is often claimed.

If this Prisoner is imperceptibly free, then we have every right to call him unfree.

3 – On noticing

3.1 – What is it to notice?

This section of the paper will be based on the concept of *noticing*. I will be using this term to refer to the moment when a mental phenomenon arises in our perception. In other words; to notice something is to begin to be aware of it.

It is quite vital to my point that this be spelled out fully. It is easy, when I speak of “things noticed,” to think of objects of the standard senses. You could notice a sudden smell of gasoline, or a bird taking flight may catch your gaze. A distant melody echoing down the street can suddenly impinge on your awareness, or you could notice an itch on the back of your left leg. These, I would think, are uncontroversial instances of noticing. But my notion of noticing refers to *anything* that can be absent from consciousness in one moment, and present in the next. This appearance is the very definition of the term.

With this definition of noticing in place—the process by which any mental phenomenon enters consciousness—one can now suitably apply it to any object of awareness. This includes not merely the things we perceive with the classical senses (the smell of gasoline; the sound of music; etc.), but to *all* objects of awareness. This means that our thoughts and our emotions, our well-formed intentions and our sudden impulses, our hopes and our desires, our hates and our

passions, our actions and our failures to act, are all phenomena that can be potentially noticed. In fact, if we are aware of them, they *must* have been noticed, otherwise we would not be aware of them at all. There is *no other way* for something to appear in consciousness than for it to be absent in one moment, and present in the next. This is the process of *noticing*.

3.2 – The argument from noticing

In this section I will present a deductive argument meant to establish a fundamental truth about the structure of our conscious minds; that this structure necessarily passive, and that it therefore precludes any first-person sense of freedom—or *subjective freedom*.² I will here describe and sum up the premises and the conclusion of the argument, before delving deeper into my support for each of the premises.

The argument can be summarized as follows:

[Premise 1] *It is the case for all potential mental phenomena that they are either noticed or not noticed, and never both.*

² Several commentators have asked me to define this term more specifically. Although I believe that such a definition is actually *technically impossible* (which only further strengthens my case), I think it may also be usefully formulated as *subjectively active*. While this still does not provide a useful definition (what the hell does “active” mean in this context?), it may be useful as a contrast to its opposite; *subjectively passive*. That total subjective passivity is the necessary and unavoidable nature of consciousness is the point of the argument from noticing.

It should also be noted here that this argument is not necessarily meant to present a direct threat to the compatibilist, who would likely reject this conceptualization of freedom as relevant at all (I will not turn to the compatibilists specifically until section 4). Instead, the objective is to give a description of how consciousness is necessarily structured, and from it to, hopefully, demonstrate intuitively to a wider readership the unreasonableness of referring to anything within this structure as “free.”

[Premise 2] *If a mental phenomenon is not noticed, it is not subjectively free.*

[Premise 3] *If a mental phenomenon is noticed, it is not subjectively free.*

[Conclusion] *No mental phenomena are subjectively free.*

For the reader with some grasp of formal argumentation, this will not only appear valid, but perhaps even trivially so. I have clearly excluded all possibility for any mental phenomena to ever be subjectively free by dividing the set into two subsets that are both incompatible with such a result. I will next offer arguments for each of the premises in turn; in support of the soundness of the argument. I expect premises 1 and 2 to be largely uncontroversial, whereas premise 3 might be in greatest need of justification, and may hold the most unintuitive implications.

3.3 – Unpacking the premises

Premise 1 stated that it is the case for all potential mental phenomena that they are either noticed or not noticed, and never both. This, I imagine, will be readily accepted by most readers.

Philosophers of mind of some sophistication may protest that it may not be quite so simple, and that the question of what we are and are not aware of is more complicated than this. But this makes no difference to the general point. We are either wholly unaware of something, or it is located *at some point* along a hypothetical spectrum of conscious awareness. This is not a dichotomy between 0 and 1, but rather between 0 and any other numerical value. (Of course, in the end, this gradient dichotomy is itself superfluous, as the ultimate point is that all roads lead to the same non-existent Rome, where the Caesar awaits us in the nude.)

Premise 2 stated that if a potential mental phenomenon is not noticed, then it is not subjectively free. I do not expect much protest on this point either, whether the phenomenon in

question is internal or external. An unnoticed external phenomenon, such as a sound or anything else delivered by our senses from without, but which fails to register, will clearly not imply any sense of subjective freedom. Likewise with any internal impressions. One might not notice that one's jaw has been clenched for the past five minutes; or that you swallowed a moment ago, just like you do all throughout the day but have learned to ignore; or that you misspelled a word just now when writing, but failed to notice. These actions do not imply any freedom, because you as the conscious witness did not even know that they happened at all.

Now, premise 3 might be the trickiest. It states that (even) if a mental phenomenon *is* noticed, it is (still) not subjectively free. This premise covers all remaining phenomena left over from premise 2, meaning that if subjective freedom is not to be found here, there is truly nowhere left for it to hide. Once again, the external aspect of this is rather uncontroversial. You may notice a sound, a sensation, a smell, or anything else that might make itself known to you, but you had no free or active role in noticing them. The sound of the song echoing down the street has nothing to do with its hearer—it simply washes over you, and you helplessly notice it whether you choose to or not. There is no sense of freedom in merely noticing sense data.

But there is often an implicit dichotomy erected between such external phenomena and more “internal” phenomena—one that appears clearly false given the view of noticing I presented above. There is only one way for anything to arise in consciousness, and that is by being noticed. This is as true for internal phenomena as for external ones. Emotions are noticed—thoughts are noticed—intentions are noticed—actions are noticed. And as noticing proceeds always in the exact same manner, there is no relevant phenomenological difference between these “different” types of appearances. Internal impressions are no more or less free

than external ones. One moment they are absent from consciousness, and the next they are there, leaving no room for freedom anywhere in the process. Before they are noticed, they cannot be the result of your subjective freedom, as you have not noticed them yet, and there is therefore nothing to deliberate on; and after they are noticed it's too late, they are already noticed, and can only be witnessed. Every subsequent change in the phenomenon, or effects following from it, can likewise only be noticed, thus barring freedom from ever entering the picture.

If you feel yourself tempted to think something along the lines of “But even if there is no freedom in noticing certain impressions, there is a felt sense of freedom in the following choice of whether and how I act on these impressions,” then I urge you to apply my reasoning to the extra mental step you have now inserted. The choice of whether, or in what manner, to act on your intrusive thought is itself another mental phenomenon—something of which you have just become conscious. Your choices, as with the intentions guiding them, can arise by no other means than by being noticed. The choice can do no other than to simply appear—to emerge out of perfect darkness—from your point of view, and no deliberation or any other supposed agentic process can change this fact (as they too can only be noticed).

Ideas are noticed, the desire to act on ideas are noticed, and the resulting motor actions are noticed. *Simply* noticed, from the *subjective* point of view. It should be made clear at this point that this is not a description of some abstract principle to which you have no access, as with many other arguments about free will. What I have presented is a description of *your mind, right now*—the only thing to which you truly do have access. The truth of this argument is ultimately made apparent not by reasoning or speculation (while not running counter to them), but can with some practice be seen directly by empirically examining your own experience.

3.4 – Collapsing the subject-object dichotomy

The argument presented in the previous section has interesting implications that may further offend the intuitions of those who would hold on to free will as a necessary condition for moral responsibility. One such implication is that it collapses the intuition of a, for this topic, relevant dichotomy between ourselves and others. Indeed, the argument would suggest that you stand in the same relation (qualitatively speaking) to your own choices and actions, as the conscious witness of them, as you do to the choices and actions of other people. The only difference is one of quantity; you simply have more *insight* into your own reasons and motives—you *notice* more of them directly.

But there is truly no qualitative difference, from the point of view of consciousness, between the “subject” and what is typically characterized as “objects” *to* that subject. According to the view of noticing as the pathway of all mental phenomena into awareness, you are yourself just as much an object to “your” consciousness as all other people or things in the world that you encounter.

Imagine a game of football, and the case of three different spectators; S1, S2, and S3. S1 is uninterested and unfamiliar with the sport, and knew not that there was a game happening at all. He only hears about it after the game is already over, at which point S1 learns the result. S2 is moderately interested in the sport, enough to watch it on the TV, but has little knowledge about specific details. The eventual result of the game, once the time runs out, will seem less surprising to S2, as he witnessed the process play out. S3, on the other hand, is a proper aficionado. He is present in the arena where the game is played, and knows everything about what players are on both teams, what their strengths and weaknesses are, about the relative effects of altitude and air pressure on the teams’ performances, and so on.

Which spectator had more control over the game?

S1 obviously had no influence on the game. He didn't even know it was happening at all! S2 was in no position to influence the game either, watching it at home on his TV. The result may have seemed less spontaneous to S2 once the game ended, but there was no control exercised on his part; no freedom relevant to the course of the game. S3 was in a much better position to understand—and not be surprised by—the final outcome, or even many of the particular events during the actual game. His quantitative level of insight puts him miles ahead of S1 and S2. But it remains a fact that S3 stands, qualitatively speaking, in the exact same position as the other spectators in relation to the game itself. No matter the quantitative level of insight, which is the *only* difference between them, they could all, in the end, do nothing but witness the events play out before them.

No spectator controls the game, or has any freedom in relation to what happens during its runtime.

In this way, the metaphor of the Subjective Prisoner in section 2 was incomplete. The Prisoner chained to the wall isn't sensorily cut off from his outside "free" self. Let's add a TV-screen in front of him, with a live feed of the actions performed in the outside world. The imprisoned consciousness can then do nothing but witness—*notice*—the actions performed on the screen. Does this witnessing change the predicament of the Prisoner in any relevant way? Has the Prisoner become *more free*? Has he come to *deserve* the suffering that comes with the punishments now that he can see the blameworthy actions on the screen? This conclusion remains absurd. At this point, does it even make any difference *which* person's actions the mind witnesses? It now seems as directly culpable for any action performed by any person, as long as it was there to witness it. The only thing capable of experiencing (which is the source of all

ethical concern) the punishment is still incapable of deserving it as it has no free part to play in the process. The state of freedom for the mind misses the grounds of ethics entirely.

In every possible case it is this Prisoner that needs to be convinced of his own freedom—every discussion on free will begins and ends in his cell, and whatever you want to say about freedom has to be said between the reality of its walls. The rest of the discussion hinges on this particular question: How reasonable is it to call this Subjective Prisoner free, given how we now understand his predicament in light of the argument from noticing? How much do you have to twist your definitions before he agrees to his own freedom? With his immovable background position established, I now turn to a popular modern way of defending free will; compatibilism, and its attempts to explain to the Prisoner, despite his every possible impression, why he is actually free and responsible after all.

4 – Compatibilism and responsibility

4.1 – Compatibilism

The most reasonable way to defend free will these days is through compatibilism, referring to the belief that free will is “compatible” with causal determinism. While the truth or falsity of determinism itself matters little within my framework, it is still among the compatibilists that one finds the most coherent defenses of free will, as it gives a more proper account of the will itself, and about the causes and structure of our minds. In truth, my view of free will could likely technically fit within a compatibilist’s framework, the only issue between us being a mutual accusation of misusing the term ‘free will’ itself.

The picture of the mind as painted by the compatibilist is most in line with scientific

findings and, most importantly, our subjective experience, as much of it relies on our capacity for free reasoning and rationality in service of our goals as agents. “Free” in this context refers to freedom from external (perhaps *rationaly unconvincing*) coercion, and an absence of conditions such as seemingly arbitrary levels of mental illness or emotional immaturity, such as in young children. A choice made, uncoerced, and with a “normal” adult mind is thus made freely. Since it is clearly a fact that most of us fulfill these conditions every day, we thusly achieve, roughly speaking, free will worth wanting.³ The position tends to be more complicated than this brief summary, but it will serve as a decent starting point.⁴

If we simply mean that our conscious choices and decisions arising originally within us as physical organisms, guided by our capacity for rational deliberation, is what constitutes free will, then I have no *technical* issue with this definition. It seems reasonable to assume, even if our thoughts are the latest link in a causal chain stretching back far beyond our own past horizons, that our conscious choices, for example, originated first within our brains, as opposed to being explicitly dictated for us and inserted by some external entity (say, some Cartesian demon, or a neuroscientist with their dubious devices).⁵ If no other agent is in direct control of your mind, then you are the locus of your choice, *qua* choice.

But I do believe this to be an inappropriate use of the term. In response to Harris, Dennett says that “if the events that cause your intentions are thoughts about what the best course of action probably is, and why it is the right thing to do, then that causation strikes me as the very

³ To borrow Daniel Dennett’s phrase from his book *Elbow Room: The Varieties of Free Will Worth Wanting*.

⁴ I recommend, for those who would like to delve deeper into compatibilism, Dennett’s books *Elbow Room* and *Freedom Evolves*, as they offer the most cogent defense of this position that I have yet to come across in my limited survey of the literature.

⁵ The “bugbears” that Dennett cautions us against in chapter one in *Elbow Room*.

epitome of freedom.” (Dennett 2017, p. 12) One could, of course, make this claim. But if my view of the mind through noticing is taken into account, this freedom is simply conjured out of nowhere. Every such thought, along with every consequence, about the “best course of action” and the “right thing to do” arise as purely spontaneous appearances that the Subjective Prisoner can only witness passively on the screen in his cell. The degree to which the choice is predictable to this witness is simply because the mind in question has developed a certain character over time, but it still grants him no more freedom in the process than the knowledgeable football-spectator has in the process of the game.

This is an issue for much of the discussion of compatibilism; freedom of the will is a phantom smuggled into the argument whenever anything desirable about the mind has been reached. In this case it seems to be the process of reason or rationality—what Thomas Pink calls the *rationalist conception of freedom*, which he describes as “simply to identify freedom with reason,” and that “On this view, to be a free agent is simply to be rational in one’s actions.” (Pink 2004, p. 44) This is a difficult position to argue against as there is, again, no *technical* issue with equating freedom and reason. One is free to employ any definition one chooses. In light of this, there are no options left open to me but to simply make an intuitive case against the appropriateness of this proposed equivalence—I am forced into a discussion about mere words, rather than their underlying phenomena. This strikes me as the legacy of compatibilism.

Every discernable fact about the process of rationality must fit atop the foundational view of the mind as expressed in section 3. Rationality can certainly survive this integration, but it should change its character in light of it, no longer being perceived as something done freely or actively by us as conscious witnesses (all that we can ever be from our own perspective). On this view, every step of reasoning must arise through noticing—not being at all available before this,

and being merely witnessed afterward. At no point throughout this process does anything else enter the picture that would allow for any sense of freedom *from your point of view*. You are at the mercy of the process of deliberation in precisely the same way you are at the mercy of, say, a sound suddenly coming over you. Or, as Harris puts it; “you no more decide the next thought you think than the next thought I write.” (Harris 2012, p. 6)

There is no relevant difference between such external appearances and the inner experience of reason and deliberation. To identify one as freedom, but not the other, is to fail to pay sufficient attention to the structure of your own mind, fully available to you in every moment, and erecting a false or arbitrary dichotomy within it. Does hearing Beethoven when his music starts playing on a nearby speaker constitute freedom? If not, then it does not apply to the “sound” of your thoughts either, nor to the eventual effects the thoughts have on your actions. At this point, when there is nowhere left for you to look where you could possibly find your own freedom of will, it should probably be recognized that the term, when applied to mere rationality, is nothing more than a misnomer. The Prisoner is not made free simply because he witnesses more rational actions on his screen.

The freedoms described by compatibilists are not unimportant, even if they are not freedoms of the *will*. They can be vital in that we often notice when we don’t have them, and when we do, we tend not to like it. But, again, these concepts give no subjectively available freedom to the *will*, to the conscious agent in any relevant way. The compatibilist notion of freedom then seems to amount to little more than having experiences that one likes (or perhaps; having experiences one doesn’t *dislike*, in the particular case of feeling compelled), a happy conjunction of intention and action, none of which are distinguishably free.

If free will were little more than “unconstrained rational action,” or most of the time “freedom to be the sort of person you already happen to be,” then yes, we would have free will. But it remains a fact that there is nothing *free* about the *will*, and so why clamor onto this phrase, even if there is no *technical* wrong committed? All this does is to inspire confusion and predictable objections from those who would use the term in more conventional ways, and to thus needlessly prolong a philosophical debate that is far longer-lived than it ever should have been. There is no use in maintaining this specific phrase unless one believes that the phrase itself is inherently valuable. This perceived value is the subject of the next section, where I argue that the actually important concepts underlying the discussion can be usefully maintained without reference to any notion of ‘free will.’

4.2 – Responsibility

Harris declares about compatibilism at one point that “More than in any other area of academic philosophy, the result resembles theology.” (Harris 2012, p. 18) A bold statement, to be sure, and one that Dennett reacts to as a “low blow.” (Dennett 2017, p. 3) Harris does little to elaborate on this point explicitly, but it is my view that there is indeed something to this blow, low or no, in that compatibilist reasoning often seems to have a premeditated aim. Theology has God as its aim, and I have therefore often referred to it as “teleological philosophy.” From what little I have so far read of the compatibilist literature, it reads suspiciously like theology in this way—and Responsibility, Desert, and Blame are Father, Son, and Holy Spirit. Free will rarely comes across as something to be argued for for its own sake, but is simply the stool on which philosophers step in order to reach what they really care about. The definition of free will in many cases may as well just be “whatever justifies moral responsibility.”

This is hardly denied, of course, as this justification is repeatedly either hinted at or admitted outright. Dennett says that we need to understand “what free will really is (and *must be*, to sustain our sense of moral responsibility).” (Dennett 2017, p. 3, emphasis added) In defense of punishment in light of free will he says that it can “be fair...justified, and in fact, our societies could not manage without it.” (Dennett 2017, p. 15) The chapter ‘Why Do We Want Free Will?’ in his book *Elbow Room* is entirely dedicated to defending this connection. He there demonstrates his commitment to Responsibility at the cost of free will by claiming that “having good reasons for wanting free will *is* having good reasons for trying to get oneself to believe one has it,” and that “if [the *illusion* of free will] were the best we could hope for, [it] *might* be worth wanting.” In Dennett’s own words: “The circumstances are,” indeed, “ripe for self-deception.” (Dennett 1984, p. 168) Other compatibilists make this connection as well. John Martin Fischer states that if we do not have compatibilist free will, then “the distinctions we naturally and intuitively draw in common sense and law would be in jeopardy of disappearing,” (Fischer, Kane, Pereboom & Vargas 2007, p. 45) and later that one “salient motivation” for believing in this form of free will is to maintain the “fundamental view of ourselves as free and morally responsible.” (Fischer, Kane, Pereboom & Vargas 2007, p. 71)

The message that emerges from this literature is clear: the *reason* we are even talking about free will in the first place is, for the most part, so that we can maintain Responsibility. Free will, in some form, seems to be solely required for this to be possible—it is both a necessary and sufficient condition. And so, in order to achieve this goal, whatever seems to justify Responsibility duly receives the label ‘free will.’

Responsibility seems to me to have only one (or one *kind* of) non-mystical definition, and that is

to ground it in relation to our ethical goals. It is a practice we employ in response to wrongdoing to achieve desirable results personally, interpersonally, and across wider society. This more grounded form of responsibility is often hinted at in the literature when the word is mentioned in the context of “*holding* people responsible” (Dennett 1984, p. 162, deemphasis added)—when it’s referred to as a *practice* rather than a *thing*. This is the proper conceptualization, as any other would veer into mysticism, as if positing some ethereal responsibility-force permeating the universe. In this way, responsibility in much of the literature positively reeks of mysticism. But responsibility grounded in practical reality can be respectably defended, and, I claim, does not require any sort of free will, even in the compatibilist sense.

This view of responsibility is founded on an essentially consequentialist perspective, although of the broadest kind. This means that, for responsibility to be a good thing, something worth maintaining, it, in some sense, needs to *work*. It needs to achieve, at least more than not, the purpose we employ it to achieve. What, then, do we want to achieve? In the broadest possible sense, we want to mitigate harm and facilitate well-being, and therefore personal or collective practices of responsibility are worthy endeavors only insofar as they successfully guide the behavior of ourselves and others in relation to those goals. In this broadest sense, to hold someone responsible is simply to recognize that they are the entity that immediately manifested the relevant behavior, and that, should this be a behavior we wish to prevent in the future, it is toward this entity that our efforts of prevention ought to be aimed if they are to have any chance of being effective. As soon as the most relevant cause of the action is something other than the entity that directly performed it, our attention should therefore be shifted, or expanded, to include this other cause as well.

Notice that this definition is general enough to be applied to just about anything; as

applicable to coconut trees as to humans (coconut trees do, after all, occasionally commit ruthless and unprovoked attacks against innocent people (Barss 1984)). If there were anything worthwhile we could do to coconut trees that would mitigate the harm done to people having coconuts dropped on their heads, then this would serve the same foundational ethical goals as that of the human practice of responsibility. But, as it turns out, coconut trees are much simpler organisms than homo sapiens, and the same methods of implicit and explicit persuasion and behavioral incentives (the “rational” parts of which compatibilists would be tempted to equate with free will) are not as effective against palm trees. But if they were, would we not be perfectly justified in applying these same methods? Responding usefully to human behavior is complicated enough that it seems to require its own words—among them ‘responsibility’—but it is, in the end, simply another point on the same sliding scale of harm-mitigating practices in response to the behavior of things in the world. To say otherwise is to claim that we are in fact *not* physical systems part of the natural world—a reemergence of the divine nature of mankind, closer to God than to nature.

Of course, compatibilists tend not to believe in such antiquated magical thinking. Their entire view is founded on the claim that our responsibility is compatible with the natural order. In this they are correct, given the view of responsibility sketched above. But at no point does anything worth calling “free will” enter the picture to carry this principle. Dennett himself gives a great explanation of the rationale underlying responsibility, without once invoking any notion of even rationalist free will:

An act has been performed, and we wish to understand how the act came about, why it came about, and what meaning we should attach to it. That is, we want to know what conclusions to draw from it about the future. Does it tell us something about the agent’s character, for instance? Does it

suggest a criticism of the agent that might, if presented properly, lead the agent to improve his ways in some regard? Can we learn from this incident that this is or is not an agent who can be trusted to behave similarly on similar occasions in the future? If one held his character constant, but changed the circumstances in minor—or even major—ways, would he almost always do the same lamentable sort of thing? Was what we have just observed a “fluke,” or was it a manifestation of a “robust” trend—*a trend that persists, or is constant, over an interestingly wide variety of conditions?* (Dennett 1984, p. 142, emphasis added)

As applicable to coconut trees as to humans.

What justifies the practice of responsibility is not freedom, or even rationality, but the fact that there are general and persisting patterns in how objects and entities in the world behave. With coconut trees it is the regularity with which coconuts fall from the tall branches that justifies any countermeasures to this occurrence. No free will need be attributed to the palm tree to require countermeasures rooted in the same ethical concerns. With humans, it is the fact that our personalities tend to develop into (more or less) set patterns consisting of preferences, intentions, and desires that remain at least somewhat similar over time. A harmful deed committed by a human is then indicative of this underlying pattern, meaning that it is at least somewhat likely to manifest again, in some way, in the future if not dealt with—a “trend that persists.” As with the coconut trees, no sense or notion of freedom is needed to ground this practice. Reasoned deliberation itself does not even ground this concept, but only informs it as it reveals something about the underlying patterns in the character that performed that reasoning.

However, if one *wants* to call this regularity freedom, and thus still insist on using this term to ground responsibility, one can. The development of personal character is certainly at least part of what the compatibilist would consider to be your free will. But why? Why feel the need to

bend this far just to keep the phrase ‘free will’ in the picture, when the moral practices we so cherish survive well without it? We are still speaking to the Subjective Prisoner here, and he should rightfully not be convinced of his own freedom simply because the actions he witnesses on his screen follow certain persisting patterns, lest he be mightily confused. He may witness these patterns, and with them the grounds for a practice of personal responsibility, but he will find no freedom in them.

That *regularity* of behavior—not freedom, even in the compatibilist form of reasoned deliberation—is the condition for responsibility can be shown by simply removing it and exploring the consequences. Imagine that human behavior followed no such set patterns—that none of us had any recognizable personalities carried over from day to day. Misdeeds of all kinds, however rationally worked out in advance, would be equally likely to manifest in any person at any time, unpredictable in principle. This, it seems, would entirely eradicate the rationale for holding individuals responsible for their own actions, as the action provided no insight into their overall character, and the punishment no discernable benefit to society. Any person guilty of a crime would be precisely as likely as yourself to commit the same offense again tomorrow, through no fault or credit of their (or your) own. Granted, we would still have serious practical issues to solve if this were the case, but *responsibility* would lose all rational basis. But the moment we reintroduce the set patterns of human behavior, it allows us to identify (at least more credibly) the reliably dangerous, reckless, or incapable individuals. It then becomes much less likely that you or I commit any considerable offenses than, say, the career criminal, and the rationale for holding particular individuals responsible for their own actions reemerges.

Letting psychopathic murderers, or career criminals, roam free among the less guilty population would still be a bad thing, even if there were nothing worth calling free will to be found on either side, simply due to the harm we would risk in allowing it. Responsibility, so central to the wider discussion about free will, over which so many capable thinkers tie themselves in mental knots to justify, is in the end entirely parallel to the free will “problem.”

Of course, this should ultimately be good news for the compatibilist, in that responsibility and other moral considerations do not suffer in the absence of subjective freedom. And so, if responsibility is indeed what really matters—what free will is most commonly employed to achieve—then there is no issue to be solved. If responsibility and the maintenance of a just society is what matters in the end, then let us talk about these concepts using their proper terms. But none of this requires a lateral move to (or an invocation of) freedom of the will at any point, unless one really likes how those particular words sound together. The abandonment of these words, as they do not seem to fit onto anything that we could possibly experience, would serve the compatibilists well, in that there is much in their discussions of other types of freedoms and their effects on moral responsibility that is worth considering. But their mission is greatly hindered by the insistence to stick with a phrase with too much baggage—a phrase that is not needed for the desired results, does not fit comfortably onto any perceptible phenomena, and creates needless contention with others who already utilize it in other more established ways.

While the Subjective Prisoner is free in no discernable way, moral considerations still apply, and there is still value in the practice of responsibility. In this manner, the most common motivation for defending free will collapses.

5 – Conclusion

My overarching claims about free will in this essay have been, essentially, two-fold. One; I put forward a conceptualization of freedom that I claim to be more ethically relevant than the common notions found in the free will-literature, and showed that we do not have this freedom (sections 2 and 3). Two; we don't need free will, or anything worthy of that name, for our practices of responsibility (section 4). The main reason philosophers are so eager to maintain this concept in modern times seems to be to uphold our practices and norms of moral responsibility. But if the nature of the will can be seen to be necessarily unfree in every ethically relevant sense, and if responsibility can survive just fine even in light of this fact, as it rests on entirely separate principles, then what work is there left for free will to do here? While it may be true that, practically speaking, a belief in free will may inspire more moral behavior in some people (Baumeister 2009, Vohs 2008), this does not make it either true or necessary for moral behavior. A belief in the Abrahamic God is claimed to provide similar checks on social behavior, but few would (and none should) use this as a defense for His existence, and changing the definition of 'God' is in the end unhelpful. Free will has become like God in this way.

When the structure of the mind is seen with any clarity, there is no reason to believe that anything worth calling free will exists; and when the nature of responsibility is cleared of its mystical undertones, leaving only the practically useful kind, there is no reason to *want* any sort of free will to exist. With everything of value in human life entirely intact in its absence, why work so hard to preserve it?

Why work so hard to convince ourselves that the emperor is wearing fine robes, when he is simply, and clearly, naked?

Literature

Barss, Peter. "Injuries Due to Falling Coconuts." *The Journal of Trauma* 24.11 (1984): 990-91.

Web.

Baumeister, Roy F, E.J Masicampo, and C. Nathan DeWall. "Prosocial Benefits of Feeling Free:

Disbelief in Free Will Increases Aggression and Reduces Helpfulness." *Personality & Social Psychology Bulletin* 35.2 (2009): 260-68. Web.

Dennett, Daniel C. *Elbow Room: The Varieties of Free Will worth Wanting*. 1984. Print.

———. *Freedom Evolves*. 2003. Print.

———. "Reflections on Sam Harris' "Free Will"." *Rivista Internazionale Di Filosofia E Psicologia* 8.3 (2017): 214-30. Web.

Fischer, John M, Robert Kane, Derk Pereboom, and Manuel Vargas. (2007). *Four Views on Free Will* (1st ed.). Wiley-Blackwell.

Harris, Sam. *Free Will*. 1. Free Press Trade Pbk. ed. 2012. Print.

Pink, Thomas. *Free Will: A Very Short Introduction*. Oxford: Oxford UP, 2004. Print. Very Short Introductions ; 110.

Vohs, Kathleen D, and Jonathan W Schooler. "The Value of Believing in Free Will: Encouraging a Belief in Determinism Increases Cheating." *Psychological Science* 19.1 (2008): 49-54.

Web.