

# **Neural Annealing:** **Toward a Neural Theory of Everything**

**Michael Edward Johnson**  
Qualia Research Institute  
johnsonmx@gmail.com

## Contents

I	Introduction. . . . .	1
II	Annealing metaphors for the brain . . . . .	1
III	How meditation works: semantically-neutral annealing . . . . .	5
IV	Depression as a disorder of annealing; bipolar depression doubly so. . . . .	8
V	The nature of trauma and the implementation of the Bayesian Brain . . . . .	11
VI	On psychedelics: . . . . .	16
VII	Love and other types of Neural Annealing . . . . .	19

## I – Introduction

Why is neuroscience so hard?

Part of the problem is that the brain is complicated. But we’ve also mostly been doing it wrong, trying to explain the brain using methods that couldn’t possibly generate ‘insight about the things we care about.

On [QRI’s lineages page \[1\]](#), we suggest there’s a distinction between ‘old’ and ‘new’ neuroscience:

Traditionally, neuroscience has been concerned with cataloging the brain, e.g., collecting discrete observations about anatomy, observed cyclic patterns (EEG frequencies), and cell types and neurotransmitters, and trying to match these facts with functional stories. However, it’s increasingly clear that these sorts of neat stories about localized function are artifacts of the tools we’re using to look at the brain, not of the brain’s underlying computational structure.

What’s the alternative? Instead of centering our exploration on the sorts of raw data our tools are able to gather, we can approach the brain as a self-organizing system, something which uses a few core principles to both build and regulate itself. As such, if we can reverse-engineer these core principles and use what tools we have to validate these bottom-up models, we can both understand the internal logic of the brain’s algorithms – the how and why the brain does what it does – as well as find more elegant intervention points for altering it.

That’s a big check to try to cash. What might this look like?

## II – Annealing metaphors for the brain

In my post about the neuroscience of meditation, I talked about *simulated annealing*, a natural implication of Robin Carhart-Harris’s work on entropic disintegration in the brain:

Annealing involves heating a metal above its recrystallization temperature, keeping it there for long enough for the microstructure of the metal to reach equilibrium, then slowly cooling it down, letting new patterns crystallize. This releases the internal stresses of the material, and is often used to restore ductility (plasticity and toughness) on metals that have been ‘cold-worked’ and have become very hard and brittle— in a sense, annealing is a ‘reset switch’ which allows metals to go back to a more pristine, natural state after being bent or stressed. I suspect this is a useful metaphor for brains, in that they can become hard and brittle over time with a build-up of internal stresses, and these stresses can be released by periodically entering high-energy states where a more natural neural microstructure can reemerge.

In his work on the [entropic brain \[2\]](#), Carhart-Harris studies how psychedelics like LSD and psilocybin add enough energy (neural activity) to the brain that existing neural patterns are disrupted, much like how heating a metal disrupts its existing molecular bonds. Recently, Carhart-Harris and Friston

have unified their frameworks under the [REBUS \[3\]](#) (RElaxed Beliefs Under pSychedelics) model, which also imports the annealing metaphor for brains:

The hypothesized flattening of the brain’s (variational free) energy landscape under psychedelics can be seen as analogous to the phenomenon of simulated annealing in computer science—which itself is analogous to annealing in metallurgy, whereby a system is heated (i.e., instantiated by increased neural excitability), such that it attains a state of heightened plasticity, in which the discovery of new energy minima (relatively stable places/trajectories for the system to visit/reside in for a period of time) is accelerated (Wang and Smith, 1998). Subsequently, as the drug is metabolized and the system cools, its dynamics begin to stabilize—and attractor basins begin to steepen again (Carhart-Harris et al., 2017). This process may result in the emergence of a new energy landscape with revised properties.

It’s a powerful metaphor since it ties together and recontextualizes so many core neuroscience concepts: free energy landscapes, Bayesian modeling, the ‘handshake’ between bottom-up sense-data and top-down priors. For a general overview of the math, see Wikipedia on [simulated annealing \[4\]](#), [Metropolis-Hastings algorithm \[5\]](#), [Parallel tempering \[6\]](#); for more on Carhart-Harris’s and Friston’s work, see [Scott Alexander’s \[7\]](#) and [Milan Griffes’ \[8\]](#) commentary. There seems to be some convergence on this metaphor: as Scott Alexander noted,

F&CH aren’t the first people to discuss this theory of psychedelics. It’s been in the air for a couple of years now – and props to local bloggers at the [Qualia Research Institute](#) and [Mad.Science.Blog](#) for getting good explanations up before the parts had even all come together in journal articles. I’m especially interested in QRI’s theory that meditation has the same kind of annealing effect, which I think would explain a lot.

### ***The basics: how does annealing work?***

Carhart-Harris’s and Friston’s model does many very clever things and is a substantial addition to the literature; I start from a similar frame but describe the process slightly differently. The following is QRI’s model (based on my talk on the [Neuroscience of Meditation in Thailand \[9\]](#)):

- First, energy (neural excitation, e.g., Free Energy from prediction errors) builds up in the brain, either gradually or suddenly, collecting disproportionately in the brain’s natural eigenmodes;
- This build-up of energy (rate of neural firing) crosses a metastability threshold and the brain enters a high-energy state, causing entropic disintegration (weakening previously ‘sticky’ attractors);
- The brain’s neurons self-organize into new multi-scale equilibria (attractors), aka implicit assumptions about reality’s structure and value weightings, which given present information should generate lower levels of prediction error than previous models (this is implicitly both a *resynchronization of internal predictive models with the environment*, and a *minimization of dissonance in connectome-specific harmonic waves*);
- The brain ‘cools’ (neural activity levels slowly return to normal), and parts of the new self-organized patterns remain and become part of the brain’s normal activity landscape;
- The cycle repeats, as the brain’s models become outdated and prediction errors start to build up again.

Any ‘emotionally intense’ experience that you need time to process most likely involves this *entropic disintegration* → *search* → *annealing* mechanism – this is what emotional processing *is*.

And I’d suggest that this is the *core dynamic of how the brain updates its structure*, the mechanism the brain uses to pay down its ‘technical debt’. In other words, entering high-energy states (i.e., intense emotional states which take some time to ‘process’) is how the brain releases structural stress and adapts to new developments. This process needs to happen on a regular basis to support healthy function, and if it doesn’t, psychological health degrades – in particular, mental flexibility & emotional vibrancy go down – analogous to a drop in a metal’s ‘ductility’. People seem to have a strong subconscious drive toward entering these states and if they haven’t experienced a high-energy brain state in some time, they actively seek one out, even sometimes in destructive ways.

However, the brain spends most of its time in low-energy states, because they’re safer: systems in noisy environments need to limit their rate of updating. There are often *spikes* of energy in the brain, but these don’t tend to snowball into full high-energy states because the brain has many ‘energy sinks’ (inhibitory top-down predictive models) which soak up excess energy before entropic disintegration can occur.

But the brain can enter high-energy states if these energy sinks are:

1. De-activated, if certain evolved trigger conditions are present – e.g., death of a loved one, falling in love, good sex, social rejection, getting bitten by a weird animal, failing some important prediction. In these cases there seems to be some sort of adaptive gating mechanism that disables the typical energy sinks in order to allow entropic disintegration → search → annealing to happen.
2. Overwhelmed, if there’s an enormous magnitude of energy coming in, faster than the energy sinks can mop it up – e.g., watching a horror movie, direct brain stimulation, first day of school, being sleep deprived, military boot camp, cult indoctrinations, your wedding day.
3. Avoided, if *semantically-neutral energy* is applied to the system. Essentially, coherent energy which isn’t strongly linked to any cognitive, emotional, or sensory process will be partially illegible to most existing energy sinks, and so it can persist long enough to build up – basically ‘hacking’ the brain’s activity normalization system. (Hold that thought – this is the most interesting one. We’ll return to it later.)

This is the ‘view from 30,000 feet’ for how simulated annealing in the brain works. If you stopped reading here, you’d walk away with a reasonable toy model of QRI’s “Neural Annealing” framework.

But there’s a lot more to the model! The rest of this writeup is an iterative tour using Neural Annealing to explain meditation, trauma, love, depression, psychedelics, and effective therapy, with each section adding a variation on the core theme.

### **Interlude: FEP, CSHW, and EBH/REBUS**

QRI’s “Neural Annealing” framework is essentially a unification of Karl Friston’s Free Energy Principle (FEP), Selen Atasoy’s Connectome-Specific Harmonic Waves (CSHW), Robin Carhart-Harris’s Entropic Brain Hypothesis (EBH), and QRI’s own Symmetry Theory of Valence (STV). Recently, Friston and Carhart-Harris have unified their respective paradigms with the *Relaxed Beliefs Under pSychedelics* (REBUS) model. I believe combining *all three* is exponentially more powerful, not only giving the *computational-level story* of REBUS, but also giving us a model for *how the brain may be physically implementing REBUS, and Bayesian updating in general*, with a correspondingly

richer set of predictions.

First, here's a quick recap: to paraphrase what I wrote [elsewhere](#) [10],

- Karl Friston's Free Energy Principle (FEP) is the leading theory of self-organizing system dynamics, one which has (in various guises) pretty much taken neuroscience by storm. It argues that any self-organizing system which effectively resists disorder must (as its core organizing principle) minimize its free energy, that free energy is equivalent to surprise (in a Bayesian sense), and that this surprise-minimization drives basically all human behavior. This minimization of surprise revolves around Bayesian-type reasoning: the brain is always getting bottom-up sense data flowing in, more than it can handle. So it relies on top-down predictive models that attempt to sort through all this data so we can focus on the surprising stuff, the stuff that can't be effortlessly predicted. The core of the FEP is the details of how this 'handshake' between bottom-up and top-down happens, and what can influence it. See [Friston's primary work](#) [11]; [Scott Alexander's attempt to distill it](#) [12]. Related to (and sometimes used synonymously with) Active Inference, the Bayesian Brain, and Predictive Processing / Predictive Coding [13].
- Robin Carhart-Harris's Entropic Brain Hypothesis (EBH) is essentially an attempt to import key concepts such as entropy and self-organized criticality from statistical physics into neuroscience, in order to explain psychedelic phenomena. As I noted above, it suggests that certain conditions *such as* psychedelics can add enough energy to brain networks that they undergo 'entropic disintegration', and then self-organize into new equilibria. See [Carhart-Harris 2018](#) [14].
- Selen Atasoy's Connectome-Specific Harmonic Waves (CSHW) is a method for applying harmonic analysis to the brain: basically, it uses various forms of brain imaging to infer what the brain's natural resonant frequencies (eigenmodes) are, and how much energy each of these frequencies have. The core workflow is three steps: first combine MRI and DTI to approximate a brain's connectome, then with an empirically-derived wave propagation equation calculate what the natural harmonics are of this connectome, then estimate which power distribution between these harmonics would most accurately reconstruct the observed fMRI activity. This framework offers several notable things: (a) these connectome-specific harmonic waves (CSHWs) are natural Schelling points that the brain has probably self-organized around (and so are worth talking about); (b) a plausible mid-level bridge connecting bottom-up neural dynamics and high-level psychological phenomena, (c) something we can actually measure. CSHW is an empirical paradigm, which is very uncommon in theoretical neuroscience. Here's a [transcript of Atasoy's explanation](#) [15]; I also wrote extensively about CSHW in [A Future for Neuroscience](#) [16]. In short: each of these three paradigms is a description of how the brain self-organizes. Friston's work understands the self-organization from a computational lens; Carhart-Harris an energetic lens; Atasoy a physical lens.

Finally, I'd offer two further pieces of background context:

QRI's own Symmetry Theory of Valence (STV), which hypothesizes that given a mathematical representation of an experience, the *symmetry* of this representation will encode [how pleasant the experience is](#) [17]. We further hypothesize that [consonance between a brain's connectome-specific harmonic waves \(CSHWs\)](#) will be a reasonable proxy for this symmetry [18].

Marr's Three Levels: as explained on our [lineages page](#) [1],

David Marr is most famous for Marr's Three Levels (along with Tomaso Poggio), which describe "the three levels at which any machine carrying out an information-processing

task must be understood:”

- *Computational theory*: What is the goal of the computation, why is it appropriate, and what is the logic of the strategy by which it can be carried out?
- *Representation and algorithm*: How can this computational theory be implemented? In particular, what is the representation for the input and output, and what is the algorithm for the transformation?
- *Hardware implementation*: How can the representation and algorithm be realized physically? [Marr (1982), p. 25]

This framework *sounds* simple, but is remarkably important since arguably most of the confusion in neuroscience (and phenomenology research) comes from starting a sentence on one Marr-Poggio level and finishing it on another, and this framework lets people debug that confusion.

Back to annealing –

As noted, Carhart-Harris and Friston have unified their paradigms under REBUS by understanding prediction errors as the ‘energy’ parameter which drives disruption (entropic disintegration) in the brain’s networks. Over time, this drives an evolutionary search function which attempts to minimize these prediction errors. I think this is a very beautiful description of a very clever system, and one which allows us an opportunity to cross-validate each model, and jump between levels of description if we get ‘stuck’. But it’s still missing a story about physical implementation. *What is this ‘energy’, physically speaking?*

### III – How meditation works: semantically-neutral annealing

I believe that almost all techniques that intentionally ‘hack’ the brain’s annealing process share a common mechanism: a build-up of *semantically neutral energy*. “Semantically neutral energy” refers to neural activity which is not strongly associated with any specific cognitive or emotional process. As I note above, usually energy build-up is limited: once a perturbation of the system neatly falls into a pattern recognized by the brain’s predictive hierarchy, the neural activity propagating this pattern is dissipated. But if a pattern *never quite matches* anything, or takes advantage of implementation-level structure to persist, and especially if it’s getting continually reinforced by some external or internal dynamic it can persist long enough to build up. I think meditation is a perfect example of a process which adds semantically-neutral energy to the brain: effortful attention on excitatory bottom-up sense-data and attenuation of inhibitory top-down predictive models will naturally lead to a build-up of this ‘non-semantic’ energy in the brain. From [The Neuroscience of Meditation](#) [9]:

Furthermore, from what I gather from experienced meditators, *successfully entering meditative flow may be one of the most reliable ways to reach these high-energy brain states*. I.e., it’s *very common* for meditation to produce feelings of high intensity, at least in people able to actually enter meditative flow. Meditation also produces more ‘pure’ or ‘neutral’ high-energy states, ones that are free of the intentional content usually associated with intense experiences which may distort or limit the scope of the annealing process. So we can think of intermediate-to-advanced (‘successful flow-state’) meditation as a reheating process, whereby the brain enters a more plastic and neutral state, releases pent-up structural stresses, and recrystallizes into a more balanced, neutral configuration

as it cools. Iterated many times, this will drive an evolutionary process and will produce a very different brain, one which is more unified & anti-fragile, less *distorted toward intentionality*, and in general *structurally optimized against stress*.

An open question is *how* or *why* meditation produces high-energy brain states. There isn't any consensus on this, but with a nod to the predictive coding framework, I'd offer that bottom-up sense-data is generally excitatory, adding energy to the system, whereas top-down predictive Bayesian models are generally inhibitory, functioning as 'energy sinks'. And so by 'noting and knowing' our sensations before our top-down models activate, in a sense we're diverting the 'energy' of our sensations away from its usual counterbalancing force. If we do this long enough and skillfully enough, this energy can build up and lead to '*entropic disintegration*', essentially pushing enough energy into the system that existing attractors are disrupted and annealing can occur.

A natural question here is *what is this 'semantically neutral energy' exactly?* – an abstract answer here is “semantically neutral energy” can be thought of as an increase in brain activity which is (1) illegible to Marr's semantic/computational level, but (2) coherent with regard to Marr's algorithmic or implementational levels (another term for this might be 'semantically-illegible energy'). But my concrete answer is that *semantically neutral energy is a build-up of energy in the brain's natural resonances* – energy accumulating in CSHWs. And so it's *this* that builds up during meditation, and *this* that starts a *semantically-neutral annealing* process which has a unique effect profile.

I think *semantically-neutral annealing* is the *best kind* of annealing for psychological health, because:

1. By mostly avoiding energy sinks, the same entropic disintegration → search → annealing process can happen using less total energy, which is less disruptive to the fine details of the system;
2. Since this energy is semantically-neutral, it doesn't depend on or trigger as many semantic processes in the brain (which can have unpredictable effects), and likewise it doesn't necessarily rely on anti-inductive 'hacks' to trick the predictive processing system, and these factors make it a more reliable and repeatable source of annealing;
3. *Very very importantly:* similarly to how vibratory energy applied to a tuning fork quickly collapses to the natural resonant frequency of the tuning fork, I'm speculating that coherent, semantically-neutral energy added to the brain will naturally cluster in the brain's natural *connectome harmonics* [16], which will thus drive an annealing process which strengthens a consonant subset of the brain's natural harmonic resonances in the long-term – essentially 'retuning the brain' toward more resonant/flow states. For more details, see *The Neuroscience of Meditation* [9];
4. Finally, this process should *feel really really good* and in the long-term, retune the mind to be more pleasant to inhabit. QRI's work on the *Symmetry Theory of Valence* [19] (STV) and our method of applying this to the brain (CDNS [18]) suggests that harmony in the brain is *literally synonymous* with pleasure, and so processes which 'deepen the grooves' of core harmonic resonances will tend to boost the mind's default hedonic level (likely helping significantly with neuroticism and emotional resilience).

# The Neuroscience of Meditation, in a nutshell:

- Meditation is a very clever step-by-step recipe for radically improving subjective experience, by building up a certain type of energy (neural activity), and letting this energy change the brain.
- Predictive coding describes the mechanism meditation uses to build up ‘semantically neutral energy’ in the brain: effortful attention on excitatory bottom-up sense data and preventing activation of inhibitory top-down models (“thinking”, “storytelling”). This buildup of energy is the ‘engine of meditation’.
- Connectome harmonics describes what this neutral energy actually looks like once it’s added to the brain: resonance in the brain’s natural harmonics.
- Neural annealing describes the way high-energy states change the brain: entropic disintegration then self-organization based on free energy (implicitly: dissonance) minimization. High-energy states generated by meditation may be especially healthy, since they’re more ‘neutral’ and should lead to a strengthening of the brain’s natural resonances.
- The Symmetry Theory of Valence describes why this matters: harmony in the brain is literally pleasant feeling. Annealing toward more harmonious states will radically improve what it feels like to be you.

I.e., Meditation is a remarkably clever technique which piggybacks on several of the brain’s core principles of self-organization: first, effortful attention on (excitatory) sense-data and inhibiting (inhibitory) predictive storytelling naturally pushes the brain into a high-energy state and makes it more malleable; this excess energy disproportionately collects in natural brain harmonics, and as the brain ‘cools’ from its high-energy state, these energized harmonics become ‘deeper’, leading to more psychological robustness. Less neuroticism and more flow. *I think this is where a large portion of the benefits of advanced meditation comes from.*

Meditation isn’t the only method to induce build-up of semantically-neutral energy; the “Big Three” are:

- **Meditation**, which seems to work by both increasing excitatory sense-data and decreasing inhibitory top-down predictive models (energy sinks);
- **Psychedelics**, which intuitively may function by disabling existing energy sinks (or perhaps overloading them by increasing baseline firing rates or increasing the branching factor of neural activity).
- **Music**, a sensory input which seems to exist on the knife’s edge between exhibiting highly ordered patterns (some of which will hit natural connectome harmonics and so allow accumulation of energy through resonance) on one hand, and on the other hand not being *too* predictable (thus dodging most inhibitory top-down predictive models);

Hybrid approaches also exist: e.g., exercise, dance, sex, tantric practices, [EMDR \[20\]](#), and breath work are essentially combinations of the rhythmic portion of music and the sensory portion of meditation. The fact that psychedelics reliably enhance the potency of each and every one of these practices is not a coincidence, but due to shared mechanism.<sup>1</sup>

<sup>1</sup>The ‘semantically neutral energy’ model also suggests why transcranial magnetic stimulation (TMS) seems

## IV – Depression as a disorder of annealing; bipolar depression doubly so

To describe depression in one sentence: “*Depression is a self-reinforcing perturbation from the natural annealing cycle.*” There are two related aspects to this: (1) an inability to anneal normally, and (2) annealing abnormally (more specifically, annealing new attractor basins which are high in dissonance, or annealing a pathological change in energy parameter dynamics).

Most people have a simple model of depression as “being sad all the time” – but I think a two-factor model looking at *energy parameter* and *valence* offers a lot of clarity and predictive utility. Roughly speaking, this suggests parametrizing depression into three core types:

1. **Depression with no high energy states**, characterized by a lack of annealing (emotional clarity and dynamism) in general
2. **Depression with high-energy negative states**, which over time anneals minds toward *suffering* and *hopelessness*
3. **Bipolar depression with high-energy positive & negative states**, which over time anneals minds toward the *dramatic*

These categories aren’t exclusive or static; too much time in one will increase the probability one may also fall into the others.

Not annealing frequently enough may be the most important ‘non-obvious’ cause of depression. Brains – especially younger ones, since they’re changing so much – *really do need to anneal regularly* to pay down their ‘technical debt’, and if they don’t, they grow brittle and neurotic. (Technical debt in the brain builds up as we twist our existing brain networks to accommodate new facts; this debt is ‘paid down’ when we enter high-energy states and let new brain networks which fit these constraints self-organize.) The ‘annealing pressure’ also increases over time, and if a wholesome annealing opportunity fails to present itself, the brain will progressively lower its standards looking for *any* opportunity for annealing. Especially if done repeatedly, this can cause long-term damage to the brain’s attractor basin landscape. (We see this in negative coping strategies such as cutting, drama-seeking, and so on – if someone is engaging in such, they’ve probably annealed poorly, and also likely have few realistic opportunities for healthy annealing.) Many forms of entertainment we think of as palliatives in today’s society (e.g., movies, video games) may be weak-and-incomplete-but-still-nonzero drivers of annealing. Not as good as the real thing, but better than nothing if that’s your only option.

At the high-energy extreme, it seems likely and tragic that depression compounds itself by repeatedly causing intense negative emotion (high-energy states) which anneals the brain toward these patterns, and toward assigning salience on the set of problems and types of thoughts (attractor landscape) facing a depressed person – many of which are their own cause and would weaken if ignored. Relatedly,

---

to [help treat depression](#) [21] – essentially, TMS injects a large amount of energy into the brain, and this energy (1) triggers some entropic disintegration, allowing escape from bad local minima, and (2) may *slightly* collect in the brain’s natural harmonics, which may help pull the brain out of dissonant equilibria. Note that this could be done *much more effectively*: instead of the present strategy of using a quick flash of unpatterned, pulsed TMS (e.g., 5 seconds 100hz) which overpowers the brain but quickly dissipates and likely doesn’t lead to a *significant* build-up in harmonics, we could instead try an entrainment approach via lower-power, rhythmic, continuous TMS, applied for longer durations (keeping the brain above its ‘recrystallization temperature’ for longer, allowing a fuller self-organization process), perhaps paired with music.

I suspect some CSHW- and music-theory-related math could be found describing how depression anneals what I would call a brain's 'connectome key signature' (CKS) toward a 'minor key', an internal logic which feels tragic/hopeless (has fewer harmonious arrangements and progressions), which the brain then uses as building blocks for its reality.

Bipolar depression seems a little more strange; the extreme highs and lows may in aggregate produce crazier annealing patterns than just one or the other – essentially there's a 'tug of war' between patterns annealed during each extreme, which prioritizes the survival of the class of patterns that exist during *both* extremely positive and extremely negative states. In practice, over time this anneals a mind's stories toward the dramatic, and toward reducing the activation energy needed to flip the brain between major and minor keys (the psych literature calls this 'kindling'). Each of these 'key signature flips' would itself release a great deal of pent-up energy, further driving the annealing process. As I note in [A Future for Neuroscience \[16\]](#):

This is not to say our key signatures are completely static, however: an interesting thread to pull here may be that some brains seem to flip between a major key and a minor key, with these keys being local maximas of harmony. I suspect each is better at certain kinds of processing, and although parts of each can be compatible with the other, each has elements that present as defection to the internal logic of the other and so these attractors can be 'sticky'. But there can also be a buildup of tension as one gathers information that is incompatible with one's key signature, which gets progressively more difficult to maintain, and can lead to the sort of intensity of experience that drives an annealing-like process when the key signature flips. And in the case of repeated flips, the patterns which are compatible with both key signatures will be the most strongly reinforced.

In some ways a bipolar brain may result in significant cognitive and creative advantages: perhaps the biggest is more access to high-energy states, which in the short term helps creativity by allowing more exploration and also steeper valence gradients to follow, and iterated over the long term allows significantly more optimization pressure on the subsystems that are repeatedly annealed. However this has corresponding epistemological downsides as noted above; fueling creative work with valence deltas is likely to 'warp the engine' over time, to paraphrase Shinzen. Friston's notion that 'systems maximizing long-term stability spend most of their time in a small number of states' seems particularly relevant to mood disorders. (My colleague Andrés suggests this 'bipolar effect profile' may be replicated by valence-enhancing drugs with a short duration and hangover, such as cocaine – this at least fits stereotypes.)

I find myself wondering if neuroticism can be thought of as ancient neural technology intended to *reduce annealing frequency* in the ancestral environment – essentially if we look into the brains of highly neurotic people, we might find strong energy sinks located around natural connectome harmonics which prevent semantically-neutral energy build-up. This likely contributes to certain forms of depression (and leads to pernicious feedback cycles – the less one anneals, the more neurotic one gets, the less able to reach high-energy states one becomes), but might also help prevent seizures or inappropriate updating/annealing, and may have frequency-dependent benefits. E.g., a group with 19 carefree annealers and 1 neurotic guardian will act more wisely than one with 20 carefree annealers or 20 neurotic guardians. The 'neuroticism = energy sinks' frame seems to suggest how to reduce neuroticism (anneal more often, especially semantically-neutral annealing), and also offer clues as to how neuroticism is implemented in the brain: we might look into the [mathematics of Anderson localization \[21\]](#) in the connectome: topological features that can 'eat' waves.

Is sleep a natural annealing process? If so, this could cleanly explain the connection between depression and chronic sleep disturbances – poor sleep as both a cause and effect of infrequent annealing. And it would indicate a treatment path: a restoration of normal annealing patterns may help improve both mood and sleep. I hold the following lightly, but we might model nREM as a ‘semantically-neutral’ high-energy state (pumping power into harmonics to perform biological housekeeping tasks, perhaps breaking up brain plaques) and REM as ‘randomized annealing’ as predictive models are switched back on, leading to both a semantic interpretation of this energy (dreaming) and the same sort of ‘neural search → annealing’ consolidation and optimization process which happens after wakeful high-energy states. This would suggest that (1) nREM disruption might lead to primarily biological deficits, and REM disruption cognitive/emotional deficits, and that (2) if one anneals very regularly they may need somewhat less sleep, in particular REM.

From a review drawing parallels between sleep and jhana (intense meditative) states:

This paper is a preliminary report on the first detailed EEG study of jhana meditation, with findings radically different to studies of more familiar, less focused forms of meditation. While remaining highly alert and “present” in their subjective experience, a high proportion of subjects display “spindle” activity in their EEG, superficially similar to sleep spindles of stage 2 nREM sleep, while more-experienced subjects display high voltage slow-waves reminiscent, but significantly different, to the slow waves of deeper stage 4 nREM sleep, or even high-voltage delta coma. Some others show brief posterior spike-wave bursts, again similar, but with significant differences, to absence epilepsy. Some subjects also develop the ability to consciously evoke clonic seizure-like activity at will, under full control. (Dennison 2019)

### *The ‘dead neuron’ model of neuroticism and depression:*

Deep learning models can exhibit ‘dead neurons’: neurons whose activation function gets ‘stuck’ on the on or off position, for instance when a sigmoid function gets too high or too low and its slope drops to almost zero. These ‘dead’ neurons can be nigh-impossible to ‘revive’ within the model, since it can be the case that their gradient (implicit sensitivity to input) is so shallow that there simply aren’t inputs that will nudge it in one or another direction.

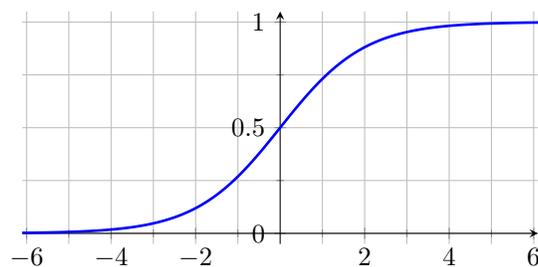


Figure 1: sigmoidal function. This loses sensitivity when values get too high or too low. Different activation functions can lose sensitivity (lead to ‘dead neurons’) under different scenarios – ReLU [22] is notorious for this.

These ‘dead’ neurons tend to cause lots of problems, since their “always on” or “always off” signal tends to propagate through the network very strongly, causing later neurons in the chain to also

exhibit less sensitivity to input. (Sometimes this process will cascade, sometimes not, much like malignant vs. benign tumors.)

I suspect this might be a strong frame for understanding the ‘psychological cruft’ which builds up in brains, and how and why regular annealing is so healthy: over time, sensitive neurons can slide into this broken state, shifting from conditional values to the neurological equivalent of static 0s and 1s. In this case I would expect more neuroticism, less flexible thinking, lower emotional resilience, and worse epistemology from people who haven’t annealed recently: lots of all-or-nothing thinking. But by injecting lots of energy into the system, enough of the internal and external context of these neurons is shifted such that some of them may get ‘reset’ and regain their conditional processing state. At the very least, this self-reorganization process can allow these neurons to move to less-critical points in processing networks.

An idea related to this frame is that a core function of neural annealing is to maintain a smooth gradient of harmony in the brain (and mind) – to make it possible to “follow your joy” toward better outcomes. If this breaks down and you can’t “follow your joy,” consider putting yourself in a situation which could plausibly kickstart an annealing process (even if you don’t feel emotionally motivated to do so).

## V – The nature of trauma and the implementation of the Bayesian Brain

Trauma is one of the worst elements of the human condition. It’s easy enough to accumulate that we all have some, and it’s hard to get rid of. But *what is it?*

Scott Alexander recently reviewed a core work in the PTSD literature, *The Body Keeps The Score* [23], and offers some context:

The book stressed the variety of responses to PTSD. Some people get anxious. Some people get angry. But a lot of people, whatever their other symptoms, also go completely numb. They are probably still “having” “emotions” “under” “the” “surface”, but they have no perception of them. Sometimes this mental deficit is accompanied by equally surprising bodily deficits. Van der Kolk describes a study on stereoagnosia in PTSD patients: if blindfolded and given a small object (like a key), they are unable to recognize it by feel, even though this task is easy for healthy people. Sometimes this gets even more extreme, like the case of a massage therapy patient who did not realize they were being massaged until the therapist verbally acknowledged she had started.

The book is called *The Body Keeps The Score*, and it returns again and again to the idea of PTSD patients as disconnected from their bodies. The body sends a rich flow of information to the brain, which is part of what we mean when we say we “feel alive” or “feel like I’m in my body”. In PTSD, this flow gets interrupted. People feel “like nothing”.

...

There’s some discussion of the neurobiology of all this, but it never really connects with the vividness of the anecdotes. A lot of stuff about how trauma causes the lizard brain to inappropriately activate in ways the rational brain can’t control, how your “smoke detector” can be set to overdrive, all backed up with the proper set of big words like “dorsolateral prefrontal cortex” – but none of it seemed to reach the point where I felt like I was making progress to a gears-level explanation. I felt like the level on which I

wanted an explanation of PTSD, and the level at which van der Kolk was explaining PTSD, never really connected; I can't put it any better than that.

...

There are a *lot* of alternative treatments for PTSD. Neurofeedback, where you attach yourself to a machine that reads your brain waves and try to explore the effect your thoughts have on brain wave production until you are consciously able to manipulate your neural states. Internal family systems, where a therapist guides you through discovering “parts” of yourself (think a weak version of multiple personalities), and you talk to them, and figure out what they want, and make bargains with them where they get what they want and so stop causing mental illness. Eye movement directed reprocessing (alternative when the book was written, now basically establishment) where you move your eyes back and forth while talking about your trauma, and this seems to somehow help you process it better. Acupuncture. Massage. Yoga.

...

Maybe the most consistent lesson from this book's tour of successful alternative therapies – keeping with the theme of the title – is that it's important for PTSD patients to get back in touch with their bodies. Massage therapy, yoga, and acupuncture addressed this directly, usually creating gentle, comfortable sensations that patients could take note of to gradually relax the absolute firewall between bodily sensation and conscious processing.

The simple Neural Annealing take on trauma is that significant negative events can push the brain into a high-energy state filled with ‘trauma patterns’, and as the brain cools, some of these trauma patterns crystallize/anneal in a very durable form, which present as PTSD.

I think this is a *more useful* answer than what's out there currently, offering straightforward intuitive answers for (1) what kinds of things are most likely to cause PTSD, (2) why PTSD is so ‘sticky’, and (3) an intuitive solution to PTSD: anneal over the bad patterns with better patterns.

But Scott's description seems to point at something further: that there's a disconnection happening with trauma. To address this, I propose the Neural Annealing model for how CSHW could *implement* the Bayesian Brain model of cognition. We'll then circle back and discuss what might be going wrong during trauma.

---

Last year in [A Future for Neuroscience](#) [16], I shared the frame that we could split CSHWs into high-frequency and low-frequency types, and perhaps say something about how they might serve different purposes in the Bayesian brain:

### **The mathematics of signal propagation and the nature of emotions**

High frequency harmonics will tend to stop at the boundaries of brain regions, and thus will be used more for fine-grained and very local information processing; low frequency harmonics will tend to travel longer distances, much as low frequency sounds travel better through walls. This paints a possible, and I think useful, picture of *what emotions fundamentally are*: semi-discrete conditional bundles of low(ish) frequency brain harmonics that essentially act as Bayesian priors for our limbic system. Change the harmonics,

change the priors and thus the behavior. Panksepp’s seven core drives (play, panic/grief, fear, rage, seeking, lust, care) might be a decent first-pass approximation for the attractors in this system.

I would now add this roughly implies a *continuum* of CSHWs, with scale-free functional roles:

- Region-specific harmonic waves (**RSHWs**) – high frequency resonances that implement the processing of cognitive particulars, and are localized to a specific brain region (much like how high-frequencies don’t travel through walls) – in theory quantifiable through simply applying Atasoy’s CSHW method to individual brain regions
- Connectome-specific harmonic waves (**CSHWs**) – low-frequency connectome-wide resonances that act as Bayesian priors, carrying relatively simple ‘emotional-type’ information across the brain (I note Friston makes a similar connection in [Waves of Prediction](#) [24])
- Sensorium-specific harmonic waves (**SSHWs**) – very-low-frequency waves that span not just the connectome, but the larger nervous system and parts of the body. These encode *somatic* information – in theory, we could infer sensorium eigenmodes by applying Atasoy’s method to not only the connectome, but the nervous system, adjusting for variable nerve-lengths, and validate against something like [body-emotion maps](#) [22]<sup>23</sup>

These waves shade into each other – a ‘low-frequency thought’ shades into a ‘high-frequency emotion’, a ‘low-frequency emotion’ shades into somatic information. As we go further up in frequencies, these waves become more localized.

An interesting implication here is *we may essentially get Bayesian updating to naturally emerge from this typology*, through interactions between these various waves: essentially, I think it’s ‘injection-locking all the way down’. ([Injection-locking](#) [25] is when harmonic oscillators (like CSHWs) essentially ‘sync up’ their periods and phases.) Specifically:

Low-frequency CSHWs carry priors, higher frequency RSHWs deal with particulars. Lower frequencies span the brain; higher frequencies resonate within more local regions of the brain – the higher the frequency of the wave, the smaller the region it tends to resonate in. The RSHWs in different regions can’t talk to each other directly, since (definitionally) these waves can’t travel across regional boundaries. But they *can* talk to each other indirectly, through interacting with low-frequency CSHWs. More specifically, I speculate that regions and CSHW-encoded priors interact through a **power-weighted averaging between CSHWs and RSHWs**, as mediated by the math of [injection-locking and injection-pulling](#) [25]. This **allows both functional partitioning and also global updating**: regions get some isolation in order to perform their specialized computations, but they also get exposure to data about the overall Bayesian prior situation, aka what we call ‘emotional information’. I.e., Region A syncs up with CSHWs, which carry the information to Region B and

---

<sup>2</sup>Thanks to Andrés for the idea about somatic information, and the suggestion of sensorium as the label.

<sup>3</sup>I suspect that *muscle tension* could be a core mechanism for regulating SSHWs and perhaps CSHWs. Tensing muscles will strongly influence body resonance, and one’s body resonance configuration will likely have ripple effects on what sorts of frequencies persist in the brain. This suggests that traditions such as yoga are basically right when they posit a link between problems in muscles and problems in the mind: we may hold tension in one system in order to compensate for a problem in the other. Speculatively, this compensatory regulation may also be found *across* humans, especially in pair bonds: that tension in your back might in some literal way be an attempt to help your partner with their emotional regulation. This would suggest muscle tension should change significantly after a break-up. (Thanks to Emily Crotteau, Lena Selesneva, and Ivanna Evtukhova for pieces of the puzzle here.)

sync up with the RSHWs there, and so on. Of note, there's a delicate, power-weighted handshake between CSHWs and RSHWs: low-frequency harmonics (emotions/Bayesian priors) carry more power per harmonic (lower due to frequency, much higher due to amplitude) but there are many more high-frequency harmonics (sensory + cognitive particulars). Strong emotions like anger likely pump huge amounts of energy into CSHWs and upend this balance, forcing RSHWs to synchronize with CSHWs. We can think of this as sacrificing the delicate epistemology-harmonization handshake in favor of unity of processing and clarity of action – or put simply, forcing perception to match top-down expectations.

***On entropic disintegration, search, and annealing in evolved harmonic systems:***

The noisy, stochastic nature of brain activity, along with practical requirements for homeostasis, will lead to a strong optimization of the CSHW + RSHW configuration for local minima which are resistant to change. However, a large enough perturbation will push the system out of this basin (**entropic disintegration** step). The **neural search** step is essentially the system stochastically testing different harmonic configurations; the **neural annealing** step is the system 'settling into' a configuration as its top-down predictive models get sufficiently good at sopping up the excess energy in the system, essentially forming a new basin it will again take a large amount of perturbation to get out of. The strength of annealing can be thought of as the steepness of this basin, and also the **Hebbian reinforcement** [26] of system attractors ("neurons that fire together, wire together"). Insofar as partitioning is possible in a broadly-coupled harmonic system, these perturbations will tend to be 'local' as the brain has strong incentives to preserve structure that doesn't need updating.

***Toward a generalized definition of trauma: a breakdown of information-propagation-via-injection-locking***

I propose that sometimes the brain needs to rapidly halt information propagation across regions to prevent cascading system failure (a metaphor that comes to mind is an uncontrolled **prion-like** [27] change in the local key signature that ripples out from a traumatized region, progressively breaking **cybernetic** [28] calibrations). I believe the brain uses two interlinked mechanisms to do this: (1) weakening CSHWs, thus weakening information propagation throughout the brain, and (2) arranging different brain regions into frequency regimes which make information transfer difficult between them (the **golden mean** [29] is the mathematically-optimal ratio for non-interaction). Once this happens, it can be very hard to reverse, since it forms a self-sustaining cycle: (1) causes (2) and (2) causes (1). *We call this 'trauma'.*

Some predictions from this – I'd expect to see substantially less energy in low-frequency CSHWs after trauma, and substantially more energy in low-frequency CSHWs during both therapeutic psychedelic use (e.g., MDMA therapy) and during psychological integration work. Stretching a little, perhaps we could also apply Atasoy's CSHW algorithm to *individual* brain regions and compare their spectrums (and those of CSHWs), to quantify the expected frequency-coupling between each region.<sup>4</sup> Possibly these two measures could be developed into a *causal quantitative metric for trauma*.

This generalized 'breakdown of communication' definition of trauma neatly fits with the story Scott tells about PTSD, where people

<sup>4</sup>My colleague Andrés suggests that "[A] more direct method, though perhaps more difficult, would be to look directly for the spectral signatures of injection locking – we'd predict you will see a seriously diminished degree of injection locking signatures on people who are heavily traumatized, and see it come back after MDMA therapy."

[A]re probably still “having” “emotions” “under” “the” “surface”, but they have no perception of them ... PTSD patients as disconnected from their bodies. The body sends a rich flow of information to the brain, which is part of what we mean when we say we “feel alive” or “feel like I’m in my body”. In PTSD, this flow gets interrupted. People feel “like nothing”.

It also fits with the therapies that seem to work: EMDR [20], neurofeedback, Internal Family Systems (IFS), yoga, massage – the consistent thread that connects these is they all plausibly help restart and strengthen communication within the brain (which I hold is strongly mediated by CSHWs). Scott doesn’t mention music, but I’d expect it to be *surprisingly* effective at boosting emotional integration – and I’d expect the *most effective* music will have strong low-frequency rhythms.

This shades into novel types of therapeutic approaches: perhaps we could simply pump energy into lower-frequency bands (perhaps harmonic stimulation centered at  $\approx 3\text{-}6\text{hz}$ ) to kickstart emotional integration.<sup>5</sup>

### ***Sidenote on music:***

The simple description I gave of music was

[A] sensory input which seems to exist on the knife’s edge between exhibiting highly ordered patterns (some of which will hit natural connectome harmonics and so allow accumulation of energy through resonance) on one hand, and on the other hand not being *too* predictable (thus dodging most inhibitory top-down predictive models).

Armed with the CSHW/RSHW distinction, we can give this a second pass. In short, I expect the above story to be true, but in a fractal way: music will be hitting both CSHWs and RSHWs. Naturally, different regions will have different sets of harmonics, which means simple tones are unlikely to produce much cross-regional resonance. Instead, the music which is the most effective at increasing the brain’s energy parameter will tie together and layer a diverse set of motifs, with two goals: (1) hitting as many connectome-specific *and* region-specific resonances as possible, and (2) entraining disparate regions and pulling them into sync – essentially using injection-locking to pull RKSs (Regional Key Signatures) into sync with each other and the CKS (Connectome Key Signature).

Could we quantify what the ‘perfect song’ would be, for a given connectome? Not exactly, since so much of music’s effects rely on getting through the brain’s predictive processing gauntlet and the state of this gauntlet isn’t well-captured by a static connectome, but we could possibly use this framework to design (potentially *much*) more evocative songs.

It’s also worth noting that better music – and *better ways to listen to music* – shade quickly into potential therapies for trauma under this model.

---

<sup>5</sup>Perhaps we could model Persistent Non-Symbolic Experience (PNSE) as persistent partial injection locking of key regions by low frequency CSHWs: essentially this would involve entraining (and effectively partially disabling) the machinery that usually handles interpretation of certain particulars/cognitive interpretations. Perhaps highly neurotic or traumatized individuals with strong top-down control exhibit the opposite: essentially trying to entrain CSHWs to a specific region (with predictably poor results).

## VI – On psychedelics:

As noted above, Neural Annealing suggests a very simple model for understanding the effects of psychedelics: as substances which “may function by disabling existing energy sinks (or perhaps overloading them by increasing baseline firing rates or increasing the branching factor of neural activity),” dramatically increasing semantically-neutral energy. Psychedelics share a ‘characteristic feeling’ (and characteristic emotional aftereffects) with each other and with activities such as meditation, listening to music, EMDR, breath work, and so on, *because* all of these things increase the energy parameter of the brain. Psychedelics are particularly interesting because they do this so powerfully, effortlessly, and noisily (with the effects bleeding over into sensory modalities, not just accumulating in harmonics).

A full Neural Annealing model of psychedelics will have to wait a few more months as internal QRI discussion settles on a unified story. But a few preliminary notes:

First, we could define ‘psychedelics’ in a principled way, as *any substance, pattern, or process that produces semantically-neutral energy accumulation* – anything that disables, overloads, or avoids the brain’s energy normalization system. The implication here is interesting, that anything that adds semantically neutral energy into the brain should produce psychedelic effects, regardless of how this is done. E.g., *even things like modern art may be classifiable as a psychedelic, insofar as it generates semantically-neutral energy* [30]. But we should also note that current psychedelics are not necessarily perfect sources of ‘clean semantically-neutral energy’; they’re substances that happen to massively increase the energy parameter of the brain, with no guarantees about how ‘balanced’ this boost is. There may be better and more targeted methods to do this in the future. In the meantime, I would recommend modest caution with substances which involve a hangover after use, as negative valence or affective blunting during a critical window could ‘sour’ the annealing process with subtle long-term mood effects.<sup>6</sup>

As mentioned above, I’ve been thinking more and more that the core psychological changes driven by psychedelics are best understood in terms of the amount and ‘statistical flavor’ of the semantically-neutral energy they add to the system. Or, as an alternate framing, psychedelics may be best understood as temporary disrupters of the brain’s natural energy sinks, each with a specific target or ‘flavor’ of disruption (or psychedelics may add to neural activity’s ‘branching factor’, which in turn will add a specific flavor to the energy). I also find myself wondering, all else being equal, whether psychedelic visuals actually are *inversely correlated* with annealing effects, since by diverting energy into the visual system (which plausibly has very effective energy sinks), there is less energy available to drive entropic disintegration.<sup>7</sup>

<sup>6</sup>My colleague Andrés also recommends against “psychedelic substances that have as part of their activity profile a high level of body-load, such as nausea and cramps as these patterns might themselves become annealing targets (cf. compounds notorious for this, according to [PsychonautWiki](#) [31] such as 2C-E, 2C-T-2, and 2C-P, are probably best avoided as therapeutic aids).”

<sup>7</sup>On psychedelic tolerance: if the semantically-neutral energy model of psychedelics proves out, we should also be open to subtle corollaries: e.g., to what extent is the temporary tolerance effect of psychedelics *biochemical* (depletion of some neurotransmitters, per the current story) and to what extent is it *information-theoretic* – associated with the release and depletion of systemic sources of Free Energy? I.e., there is potential energy of a sort liberated when the system finds a better local minima, and if the system has undergone strong annealing recently, there are fewer such ‘energetic free lunches’ around to help power the psychedelic effects. (Hypothesis held weakly, as my colleague Andrés points out there are psychedelics which do not trigger tolerance, such as N,N-DMT and 5-MeO-DMT.)

As I noted in *A Future for Neuroscience* [16], another starting point for sorting through psychoactive drugs would be

[T]o parametrize the effects (and ‘phenomenological texture’) of all psychoactive drugs in terms of their effects on the consonance, dissonance, and noise of a brain, both in overall terms and within different frequency bands (Gomez Emilsson 2017).

In the long term, we’ll want to move upstream and predict *connectome-specific* effects of drugs – treating psychoactive substances as operators on neuroacoustic properties, which produce region-by-region changes in how waves propagate in the brain (and thus different people will respond differently to a drug, because these sorts of changes will generate different types of results across different connectomes). Essentially, this would involve evaluating how various drugs change the *internal parameters* of the CSHW model, instead of just the outputs. Moving upstream like this might be necessary to predict why e.g., some people respond well to a given SSRI, while others don’t (nobody has a clue how this works right now).

Possibly this would allow us to generate a principled typology of psychoactives, and also check for missing quadrants: psychoactives and psychedelics we haven’t discovered or created yet. (See also Andrés’s notion of parametrizing the ‘*information vs energy trajectory*’ [32] of a trip.) We can also think of anti-psychotic drugs as *anti-psychedelics*: substances that *rapidly decrease the energy parameter of the brain* [33]. We at QRI strongly believe this makes anti-psychotics more dangerous than commonly realized: the neural search process is complex and delicate, and an externally-forced, uneven rapid cooling process may warp the internal landscape of the brain in subtle but deleterious ways. In theory, we could test this indirectly by evaluating the effects of anti-psychotics on sensory integration tasks in healthy controls – but as noted above, this may be an unethical experiment.

Another frame would be ‘psychedelics as full-spectrum resonance agents’ – CSHWs are meant to substantially resonate during normal human operation (falling in love, orgasm, etc.) – RSHWs are not. The perceptual and epistemological changes we sometimes see during psychedelics could be due to the fine logical machinery that usually deals with high-context sensory particulars (facts and logical inferences) starting to malfunction as its natural eigenmodes are activated. Like linking and rhythmically flipping all the bits in a memory register, ignoring what that register is “supposed to” compute. If psychedelic visuals are an example of RSHW resonance, HPPD [34] may be an example of this RSHW resonance annealing into durable patterns.

On MDMA’s strangely powerful therapeutic effects, I’d suggest MDMA shares the ‘*basic psychedelic package*’ [35] with substances like LSD and psilocybin (albeit a little weaker at common doses). Anything with this ‘baseline’ package significantly increases the energy parameter of the brain, which both allows escape from bad local minima and canalizes the brain’s core CSHWs, which both should be highly therapeutic. My intuition is MDMA *may also* have a particular effect on stochastic firing frequencies of neurons, and that this effect essentially acts as an emergent metronome – and this metronome will drive synchronicity between diverse brain regions. Given the presence of such a region-spanning ‘clean’ metronomic signal, brain regions that have partially ‘stopped talking to each other’ will re-establish integration, and some of this integration will persist while sober (or rather, some of the reasons for the lack of integration will have been negotiated away during the MDMA-driven integration). Plausibly this ‘emergent metronome’ effect may also underlie the particular phenomenological effects of 5-MeO-DMT as well, particularly in terms of sense of unity, high valence, and therapeutic potential.<sup>8</sup>

<sup>8</sup>HT to Steve Lehar for pointing at this ‘*nystagmus*’ phenomenon [36] as being somehow linked to MDMA’s

*Somewhat poetic sidenote: on taking psychedelics*

In the abstract – I think psychedelics are more powerful, more dangerous, and more healing than commonly assumed.

But we don't live in the abstract. The natural question for any given person is thus: *should I take them?*

There's no one-size-fits-all answer, and I recommend checking with local laws. But I can share a simple heuristic for who shouldn't worry too much about the downsides of psychedelics and who should be very careful: do you trust your own aesthetic?

Psychedelics massively increase the 'energy parameter' of the brain, so naturally there's a large amount of very-high-dimensional exploration going on. There are countless 'micro-choices' your brain makes as to how to anneal after this exploration: we can think of a person's 'aesthetic' as individual variance in these annealing choices. What the self-organizing system which is the brain's subconscious finds beautiful in the moment and implicitly strives to save.

Sometimes, and in some people, we want the right things, we find the right things beautiful. Things that have a deep elegance and fit with everything about us and fit with how reality works. We just need enough energy parameter to get there. Psychedelics are a great way to get there.

Other times, we might not want the right things. Evolution is kind of a jerk, epistemologically speaking: it cares much more about genetic reproduction than it does about deep coherence and calibration with reality and such. Sometimes we're at a functional local maxima, but we're not pointed in the right direction globally, and frankly speaking our lack of a high energy parameter is our saving grace – our inability to directly muck up our emotional landscape. Insofar as this is true – and it will be more true at certain times than others, and in certain people than others, and perhaps in certain combinations of people than others – using psychedelics to crank the energy parameter is not good for a person. Our 'Psychedelic Extrapolated Volition' (PEV) is not a healthy vector.<sup>9</sup>

The natural follow-up is, how do you know whether your PEV is positive or not?

Hard question, but probably good to ask your friends – group epistemology seems healthy in these cases. And in general it seems strongly preferable to err on the side of caution. You can always take that LSD tomorrow, or next week, or next year.

(But, don't be too paranoid about one trip permanently breaking your brain, either. My guess is the annealing that tends to 'stick' is that which actually finds better local minima (thankfully) – if it's an unsuccessful exploration I suspect the system can usually climb back to where it was (with some caveats).)

A separate factor is your current energy parameter and how psychedelics may increase this baseline: if you're dragging on the bottom of your energetic attractor basins, maybe a little kick could be

---

mood-lifting effect, and to Andrés for calling my attention to Lehar's work and suggesting 5-MeO-DMT may also share this mechanism.

<sup>9</sup>This is a reference to Eliezer Yudkowsky's "Coherent Extrapolated Volition" (CEV) [37] concept, which is an attempt to sketch a heuristic for how to use a radically-powerful optimization process (such as an AGI) safely. Essentially, CEV suggests we could aggregate all human preferences (volitions), find some way to merge them (make them 'cohere'), then repeat (extrapolate), until we get to a self-stable loop. A 'psychedelic extrapolated volition' is a variation of this: if it becomes easier to change yourself on psychedelics, and then that person you turn into can change themselves into someone else, and so on, where do you end up? What generates a 'positive vector' here?

healthy. But if you're already 'high on life' – consider skipping the LSD and MDMA. Increasing a high baseline can redline the system into exquisitely unbearable intensity.

## VII – Love and other types of Neural Annealing

It's important to note that most annealing doesn't happen in a vacuum: just as "set and setting" matter quite a lot for psychedelics, and for emotional updating in general, the importance of *context* in the annealing model is hard to overstate. Much as holding a magnet close to iron as it cools can magnetize the metal, *the intentional content present when entropic disintegration → annealing happens provides important constraints for which new patterns form.* I propose there are four general types of neural annealing:

- **Annealing to an object or event.** Annealing which is 'pointed at' something is by far the most common type. Some object, or event, or new insight makes itself known in a surprising or otherwise intensely salient way, and this pushes the brain into a high-energy state, kickstarting a self-organization process for accommodating the presence and significance of this new thing. This can involve intense positive emotion – a new romantic partner, the birth of your child, your wedding day. This sort of annealing can also be caused by trauma – getting bitten by a weird animal, social rejection, losing a close one. As I suggested in *The Neuroscience of Meditation* [38], neural annealing may offer a rather pithy description of love:

Finally, to speculate a little about one of the deep mysteries of life, perhaps we can describe *love* as the result of a strong annealing process while under the influence of some pattern. I.e., evolution has primed us such that certain intentional objects (e.g., romantic partners) can trigger high-energy states where the brain smooths out its discontinuities/dissonances, such that *given the presence of that pattern* our brains are in harmony. This is obviously a two-edged sword: on one hand it heals and renews our 'cold-worked' brain circuits and unifies our minds, but also makes us *dependent*: the felt-sense of this intentional object becomes the key which unlocks this state. (I believe we can also anneal to *archetypes* instead of specific people.)

Annealing can produce durable patterns, but isn't permanent; over time, discontinuities creep back in as the system gets 'cold-worked'. To stay in love over the long-term, a couple will need to re-anneal in the felt-presence of each other on a regular basis. From my experience, some people have a natural psychological drive toward reflexive stability here: they see their partner as the source of goodness in their lives, so naturally they work hard to keep their mind aligned on valuing them. (It's circular, but it works.) Whereas others are more self-reliant, exploratory, and restless, less prone toward these self-stable loops or annealing around external intentional objects in general. Whether or not, and within which precise contexts, someone's annealing habits fall into this 'reflexive stability attractor' might explain much about e.g., attachment style, hedonic strategy, and aesthetic trajectory.

Perhaps we can go further now, and hypothesize that 'falling in love' is a specific algorithm the brain runs, which is triggered by when the 'felt sense' of another person (a pattern distributed across RSHWs, CSHWs, and SSHWs) produces substantial systemic resonance. When this happens, and in the absence of warning signs (dissonance), a person will actively seek to fill their sensorium with this signal, which amplifies the systemic resonance (potentially to extreme levels) and further synchronizes priors and other regions into harmony with the original pattern.

*As you fall in love, you literally anneal to your felt-sense of that person – you take their rhythm as yours, because your body judged it to be so.* A key which fit your connectome’s lock. This will naturally do two things: (1) fuzz boundaries between lovers, as patterns progressively synchronize, and (2) add a harmonic echo, or ‘warm consonant glow’ to all thoughts about the person. This latter phenomenon will feel nice, but also keep itself stable: the presence of this bundle of synchronized frequencies will stabilize (via injection-locking) many forms of drift – effectively preventing certain thoughts/perceptions. This may fade over time if not refreshed, but perhaps to completely ‘fall out of love’ the brain has to build a competing key signature elsewhere, e.g., in a golden mean ratio to this harmonic echo, and these rivalrous key signatures (implicitly Bayesian priors about what is real and what is good) battle it out.<sup>10</sup> This ‘de-annealing’ process – literally erasing someone’s patterns and rhythms from your body – can follow several trajectories, few of them pleasant, as the system renegotiates new (or old) equilibria.

- **Annealing to an ontology.** A much more general type of annealing is when the entropic disintegration → annealing process is pointed toward an *ontology*, and the brain reorganizes its internal structure (‘ontological contours’) to accommodate this new typology. This can happen implicitly and weakly, over the course of entropic disintegration → annealing to multiple separate ideas, or explicitly and strongly, for instance reading some book in college which completely reshapes one’s view of reality.

Any craftsman, any intellectual, any philosopher worth their salt is strongly annealed toward at least one nuanced ontology, and in fact much of the influence of the Great Philosophers can be found in how they’ve laid out their thoughts in a way that others can use as a *coherent annealing target*. What makes something a good annealing target? I’d offer it’s the presence of clear archetypes arranged in both a novel but ultimately cognitively efficient way. These archetypes can be thought of as a combination of nature (innate Jungian-type limbic resonances) and nurture (prior annealed patterns & cultural reifications).

An important point here is that peoples’ conception of *where goodness comes from* is dependent upon their ontology; change the ontology, change the perceived nature of goodness itself! See, e.g., John Lily’s discussion of the *supra-self-metaprogrammer (SSMP)* [39]. This frame-shift can also manifest at the extreme end of falling in love, where all the world’s goodness seems to come from your special person (a dangerous thing).

- **Social annealing.** A special hybrid of annealing to an ontology and to other people is **social annealing**, wherein a *group of people* undergoes the ‘*entropic disintegration → neural search → annealing*’ process together, within some shared context – a religious service, a sporting event, a retreat. This seems like the natural mechanism by which tribes are formed (loosely speaking, group synchronization of connectome-specific harmonic wave dynamics) and underlies many of our most sacred experiences. The power of social annealing is such that a religious experience that lacks it no longer feels like a religious experience – merely the mouthing of dogma. On the other hand, any group experience that *does* increase the group’s energy parameter and trigger annealing starts to take on a pseudo-religious frame – e.g., ecstatic dance, [festivals](#) [40], protest marches, even concerts.
- **Semantically-neutral annealing.** Almost all neural annealing is *semantic annealing*, or *annealing toward some intentional object*. This process is pointed *at* something, often the thing that caused the entropic disintegration process in the first place, be it a person, an event, an idea, an ontology. But there’s nothing in the laws of neuroscience that implies annealing *has to*

<sup>10</sup>Thanks to Andrés for discussion on competing key signatures.

have an intentional object as a focus. As per Section III, I believe this is a particularly healthy form of annealing.

### *Toward a new psychology & sociology?*

Speculatively, we may be able to re-derive much of psychology and sociology from just the energy-parameter view of the brain: e.g.,

Gopnik suggests that different developmental windows may involve different implicit ‘heat parameters’ for simulated annealing [3], with young people having higher parameters. Speculatively, this may correspond to different ‘lived intensity of experience’ at different ages – young brains (and lifelong learners) might not only be *more plastic* than average, but actually having experience that is objectively *more visceral*. One way to frame this is that *being young is like microdosing on LSD all the time*. This could have interesting implications for ethics.

Most likely, there’s been significant recent sexual selection for a higher energy parameter, for several reasons:

- Selecting for *neoteny* [41] plausibly also implicitly selects for an energy parameter that starts higher and/or decays less with age
- A high energy parameter would be a good proxy for cognitive-emotional-behavioral dynamicism, perhaps the most strongly sexually-selected-for trait
- A high energy parameter would be an honest signal of not being in a bad ‘iterated aesthetics’ attractor (otherwise they would have self-destructed previously)

Psychology has various personality metrics, with the most widely used being the Big 5 [42], also known as OCEAN (Openness, Conscientiousness, Extroversion, Agreeableness, Neuroticism). One of the most interesting subfindings here is that we can still get reasonable predictive utility if we collapse these into a one-variable model: the ‘Big One’ personality factor [43]. Scoring high in this factor is “associated with social desirability, emotionality, motivation, well-being, satisfaction with life, and self-esteem.” Scoring low is associated with depression, frailty, lack of emotionality, and so on. I wouldn’t be surprised if the ‘Big One’ simply tracks how frequently and deeply someone anneals.

Continuing the thread on Social Annealing, I think we can push into sociology with the Neural Annealing model, too; **to understand a society, we need to understand how and when annealing happens in that society**. To gauge the wisdom of a society, look at how its decision-makers anneal; to gauge the cultural direction of a society, look at how its young people anneal. To understand the strongest social bonds of a society, look at the contexts in which group annealing happens.

This also suggests why drugs like alcohol and certain psychedelics are ritualistically celebrated in so many cultures: they allow social-annealing-on-demand, a key technology in building and maintaining social cohesion and coordination.

Likewise, we could envision a field of ‘*social archeology*’ evaluating annealing patterns in the past: how often did peasants and nobles in the Middle Ages anneal? In what contexts did the annealing happen, and which institutions controlled them? Perhaps most political conflicts could be reinterpreted as conflicts over annealing.<sup>11</sup> And so on. My colleague Andrés has suggested that a good rule of thumb

<sup>11</sup>This naturally and unfortunately makes the access to and contexts of social annealing an axis of cultural

for identifying annealing (and making good movies) is that intuitively, annealing defines where you should actually point the camera if you were making a movie of a historical period, since where annealing is happening is where changes that ‘matter’ are taking place: cognitive updates, decisions about how to feel, and so on.

***On the effect of profession on emotional vibrancy:***

It would be somewhat surprising if certain repeated computational tasks didn’t tend to push regions’ key signatures into being highly coupled (= intense emotions), whereas other classes of tasks push regions’ key signatures into fairly orthogonal configurations (= ‘white noise’ as emotional state). A lifetime of dance or poetry might literally make you feel emotions more strongly; a lifetime of doing accounting might literally produce a segmented brain and affective blunting. From Darwin’s autobiography:

I have said that in one respect my mind has changed during the last twenty or thirty years. Up to the age of thirty, or beyond it, poetry of many kinds, such as the works of Milton, Gray, Byron, Wordsworth, Coleridge, and Shelley, gave me great pleasure, and even as a schoolboy I took intense delight in Shakespeare, especially in the historical plays. I have also said that formerly pictures gave me considerable, and music very great delight. But now for many years I cannot endure to read a line of poetry: I have tried lately to read Shakespeare, and found it so intolerably dull that it nauseated me. I have also almost lost my taste for pictures or music. ... I retain some taste for fine scenery, but it does not cause me the exquisite delight which it formerly did. ...

This curious and lamentable loss of the higher aesthetic tastes is all the odder, as books on history, biographies, and travels (independently of any scientific facts which they may contain), and essays on all sorts of subjects interest me as much as ever they did. My mind seems to have become a kind of machine for grinding general laws out of large collections of facts, but why this should have caused the atrophy of that part of the brain alone, on which the higher tastes depend, I cannot conceive.

Reading this account, I find it plausible that Darwin repeatedly pushed (and annealed) his mind toward RSHW-driven ‘clockwork piecemeal integration’ interactions rather than CSHW/SSHW-driven global symmetry gradients, although Darwin’s age, sickness, and depression may have also contributed. A warning sign for us theorists and systematizers.

## Conclusion

Neural Annealing is a neuroscience paradigm which aims to find the optimal tradeoff between elegance and detail. It does this by identifying a level of abstraction which supports parallel description under three core principles of self-organization: physical self-organization (around connectome resonances), computational self-organization (around minimization of surprise), and energetic self-organization (around conditional entropic disintegration).

---

conflict: those who control these events control the emotional tone and contours of coordination of a society. Taking away healthy annealing contexts from your opponents and giving more social annealing opportunities to your people is a key (but also very dirty) way to ‘win’ the culture war. (Perhaps the opioid crisis, and the crack-cocaine crisis before it, could in some sense be exacerbated by a lack of healthier annealing opportunities.)

There is yet much work to be done: in particular, there are huge bodies of literature around receptor affinities, network topologies, regional anatomies and cell types, and so on. The promise of Neural Annealing is it's not only a predictive and generative theory in its own right, but it provides a level of description by which to connect these disparate maps, and an extensible context to build on as we add more and more detail to the model.

Finally, we can ask: why does good neuroscience *matter*? I would offer the following.

The future could be much better than the present. *Much* better.

Material conditions are only very loosely coupled with well-being. If life is to be radically better in the future, it will be due to better neuroscience pointing out how we can be kinder to ourselves and others, and future neurotechnology changing the hedonic calculus of the human condition.

A unified theory of emotional updating, depression, trauma, meditation, and psychedelics may give us the tools to build a future that's *substantially* better than the present. This has been my hope while writing this.

### ***Citation***

For attribution in academic contexts, please cite this work as:

Michael Edward Johnson, "Neural Annealing: Toward a Neural Theory of Everything", <https://opentheory.net/2019/11/neural-annealing-toward-a-neural-theory-of-everything/>, San Francisco (2019).

### ***Acknowledgments***

I'd like to thank Andrés Gómez Emilsson for many great conversations on annealing (and first [calling my attention \[44\]](#) to the term), energy sinks, and countless other topics, and careful feedback on a draft of this work; Robin Carhart-Harris and Karl Friston for a beautiful description of simulated annealing; Romeo Stevens for wide discussion about annealing & ontologies; Adam Safron for introducing me to the depth of explanation afforded by predictive coding, pointing me toward injection locking, and many great conversations in general; Quintin Frerichs for his hard work toward making therapeutic applications of this theory real, and the rest of the QRI team for support and inspiration; QRI's community and in particular our donors, for their generous support; Milan Griffes for careful feedback on a draft of this work; Alex Alekseyenko and James Dama for discussions about simulated annealing; Anthony Markwell for sharing the Buddhist Dhamma with me in such a thoughtful and generous way; Justin Mares for his constant curiosity and encouragement; my parents, for their endless love and patience; Lena Zaitseva and Lena Selesneva for their deep warmth and support; and especially Ivanna Evtukhova, who has made my life radically better and whose love, energy, and obsession with Buddhist enlightenment gave me reason to start.

### ***Notable comments on this work***

1. On the role of CSHWs and information processing & bandwidth, [LessWrong user lsusr responding to the comment](#) that "The power distribution among harmonics carries very little information—dozens of bits, not the billions or trillions of bits that are needed for human-level understanding of the world. So what's the relation between CSHW and biologically-useful computations?":

This makes sense to me because I work full-time on the bleeding edge of applied AI,

meditate, and have degree in physics where I taught the acoustical and energy-based models this theory is based upon. Without a solid foundation in all three of these fields this theory might seem less self-evident.

Hopefully this explanation can help you understand the theory behind the theory. First I'll address points (1), (2) and (3). Then I'll explain the bandwidth issue in more detail.

(1) While it's true that these harmonic frequencies have less information bandwidth than synapses that doesn't mean they don't perform biologically-useful computations. High-bandwidth pattern matching is trivially easy to do with neural networks. The hardest part about neural networks is time series data. (I know this because I am a specialist in the [application](#) of machine learning algorithms to handle time-series inputs.) To simplify the situation right now, we (1) don't know how to use neural networks to handle time series data (2) don't know how to get different machine learning systems of any kind to work together—especially with regard to time series data. If CSHW can make any progress in this direction then that would be useful.

(1.1) You are correct that we need a traditional mass of neurons tuned via gradient descent in order to handle high-bandwidth information like our many nerves and to handle complex actions like muscle control. CSHW does not get in the way of these things. Rather CSHW is a simple, elegant way to coordinate many different sub-networks into a human brain. It's not about “how” do you throw a baseball. It's “when” do you throw a baseball. When different networks are out of phase with each other the inputs of one turn into static for the other, which is literally equivalent to tuning out a radio. In short, the purpose of CSHW is not to replace the massive information processing solved by neural networks. Instead, it's purpose is to combine and separate neural networks, as applicable, in response to time-series inputs. It does this fractally, which is the only way to simplify a design to handle massive complexity in a biological system.

(1.2) All CSHW needs to do is to tell which networks should receive information from which other networks. High-frequency waves both propagate shorter distances and oscillate faster (have higher bandwidth) than low-frequency waves, so the information density and response speed gets higher where it needs to be higher (on smaller scales). Remember: the oscillations don't have to transfer information. That's performed by the traditionally-understood neuronal connections. The oscillations can bring different systems in and out of sync in a coordinated manner. This happens at a lower frequency than individual neuron firings and involves larger masses than individual neurons so the necessary bandwidth is much lower. Frequency space might just be a dozen bits long, but there's three spacial dimensions based on actual physical space too.

(1.3) The low bandwidth of the harmonic frequencies explains an important puzzle about consciousness. You know how you can only keep 3-9 concepts in working memory at once? This could be a reflection of the low bandwidth of the low frequency waves.

(2) We have known neurons and evolution are capable of producing waves like this (especially the low frequency ones) for ages. The question neuroscience has been struggling with isn't “can” neurons produce waves like this. It's “why”.

(2.1) This theory describes observed behavior especially well once you compare the theory's predictions to the observed brainwaves in advanced meditators. The brain scans of Tibetan yogis and the traditional subjective descriptions written by Zen

masters match descriptions of the low frequency brain resonance predicted by this theory. So does a modern Vipassana manual, though it focuses on the high frequency end of the spectrum. This is 3/3 major Buddhist lineages.

(3) As Michael Edward Johnson (OP) mentioned in another comment, recent advancements in fMRIs have let us observe some of the phenomena described in CSHW.

## References

- [Joh+] Michael Edward Johnson et al. *QRIs Lineage Page*. URL: <https://www.qualiaresearchinstitute.org/research-lineages/>.
- [Car+14] Robin Carhart-Harris et al. “The entropic brain: a theory of conscious states informed by neuroimaging research with psychedelic drugs”. In: *Frontiers in Human Neuroscience* 8 (2014), p. 20. ISSN: 1662-5161. DOI: 10.3389/fnhum.2014.00020. URL: <https://www.frontiersin.org/article/10.3389/fnhum.2014.00020>.
- [CF19] R. L. Carhart-Harris and K. J. Friston. “REBUS and the Anarchic Brain: Toward a Unified Model of the Brain Action of Psychedelics”. In: *Pharmacological Reviews* 71.3 (2019). Ed. by Eric L. Barker, pp. 316–344. ISSN: 0031-6997. DOI: 10.1124/pr.118.017160. eprint: <http://pharmrev.aspetjournals.org/content/71/3/316.full.pdf>. URL: <http://pharmrev.aspetjournals.org/content/71/3/316>.
- [Wik20a] Wikipedia contributors. *Simulated annealing* — *Wikipedia, The Free Encyclopedia*. [Online; accessed 16-April-2020]. 2020. URL: [https://en.wikipedia.org/w/index.php?title=Simulated\\_annealing&oldid=948050151](https://en.wikipedia.org/w/index.php?title=Simulated_annealing&oldid=948050151).
- [Wik20b] Wikipedia contributors. *Metropolis–Hastings algorithm* — *Wikipedia, The Free Encyclopedia*. [Online; accessed 17-April-2020]. 2020. URL: [https://en.wikipedia.org/w/index.php?title=Metropolis%E2%80%93Hastings\\_algorithm&oldid=950996820](https://en.wikipedia.org/w/index.php?title=Metropolis%E2%80%93Hastings_algorithm&oldid=950996820).
- [Wik19] Wikipedia contributors. *Parallel tempering* — *Wikipedia, The Free Encyclopedia*. [Online; accessed 16-April-2020]. 2019. URL: [https://en.wikipedia.org/w/index.php?title=Parallel\\_tempering&oldid=926897488](https://en.wikipedia.org/w/index.php?title=Parallel_tempering&oldid=926897488).
- [Ale19a] Scott Alexander. *SSC Journal Club: Relaxed Beliefs Under Psychedelics And The Anarchic Brain*. [Online; accessed 16-April-2020]. 2019. URL: <https://slatestarcodex.com/2019/09/10/ssc-journal-club-relaxed-beliefs-under-psychedelics-and-the-anarchic-brain/>.
- [Mil20] Milan Griffes. *Carhart-Harris & Friston 2019 – REBUS and the Anarchic Brain*. [Online; accessed 16-April-2020]. 2020. URL: <https://qualiacomputing.com/2019/08/27/carhart-harris-friston-2019-rebus-and-the-anarchic-brain/>.
- [Joh19] Mike Johnson. *The Neuroscience of Meditation - Meditation & Science Jam 2019, Thailand*. Youtube. 2019. URL: <https://www.youtube.com/watch?v=nY9H5CK4GbE>.
- [Joh18a] Michael Edward Johnson. *Seed ontologies*. 2018. URL: <https://opentheory.net/2018/06/seed-ontologies/>.
- [Fri10] Karl Friston. “The free-energy principle: a unified brain theory?” In: *Nat. Rev. Neurosci.* 11.2 (2010), pp. 127–138.
- [Ale18] Scott Alexander. *God Help Us, Let’s Try To Understand Friston On Free Energy*. [Online; accessed 16-April-2020]. 2018. URL: <https://slatestarcodex.com/2018/03/04/god-help-us-lets-try-to-understand-friston-on-free-energy/>.
- [Joh16a] Michael Edward Johnson. *Principles and Processes*. 2016. URL: <https://opentheory.net/2016/11/principia-qualia/>.

- [Car18] Robin L. Carhart-Harris. “The entropic brain - revisited”. In: *Neuropharmacology* 142 (2018). Psychedelics: New Doors, Altered Perceptions, pp. 167–178. ISSN: 0028-3908. DOI: <https://doi.org/10.1016/j.neuropharm.2018.03.010>. URL: <http://www.sciencedirect.com/science/article/pii/S0028390818301175>.
- [Emi17a] Andrés Gómez Emilsson. *Connectome-Specific Harmonic Waves on LSD*. 2017. URL: <https://qualiacomputing.com/2017/06/18/connectome-specific-harmonic-waves-on-lsd/>.
- [Joh18b] Michael Edward Johnson. *A Future for Neuroscience*. 2018. URL: <https://opentheory.net/2018/08/a-future-for-neuroscience/>.
- [Joh16b] Michael Edward Johnson. *Principia Qualia*. 2016. URL: <https://opentheory.net/2016/11/principia-qualia/>.
- [Emi17b] Andrés Gómez Emilsson. *Quantifying Bliss: Talk Summary*. 2017. URL: <https://qualiacomputing.com/2017/06/18/quantifying-bliss-talk-summary/>.
- [Joh17] Michael Edward Johnson. *Symmetry Theory of Valence “Explain Like I’m 5” edition*. 2017. URL: <https://opentheory.net/2017/04/stov-explain-like-im-5-edition/>.
- [Wik20c] Wikipedia contributors. *Eye movement desensitization and reprocessing* — *Wikipedia, The Free Encyclopedia*. [Online; accessed 16-April-2020]. 2020. URL: [https://en.wikipedia.org/w/index.php?title=Eye\\_movement\\_desensitization\\_and\\_reprocessing&oldid=950960587](https://en.wikipedia.org/w/index.php?title=Eye_movement_desensitization_and_reprocessing&oldid=950960587).
- [Har17] Kevin Hartnett. *Mathematicians Tame Rogue Waves, Lighting Up Future of LEDs*. Quanta magazine. 2017. URL: [https://www.quantamagazine.org/mathematicians-tame-rogue-waves-lighting-up-future-of-leds-20170822/?fbclid=IwAR1wFHwPwELj8p18FhnZZ7P\\_YOcaK\\_Ro\\_Xb3oRSMd8IpU6o3195HD3neYUw](https://www.quantamagazine.org/mathematicians-tame-rogue-waves-lighting-up-future-of-leds-20170822/?fbclid=IwAR1wFHwPwELj8p18FhnZZ7P_YOcaK_Ro_Xb3oRSMd8IpU6o3195HD3neYUw).
- [Git17] Github contributors. *Activation Functions*. github. 2017. URL: [https://ml-cheatsheet.readthedocs.io/en/latest/activation\\_functions.html#relu](https://ml-cheatsheet.readthedocs.io/en/latest/activation_functions.html#relu).
- [Ale19b] Scott Alexander. *Book Review: The Body Keeps The Score*. [Online; accessed 16-April-2020]. 2019. URL: <https://slatestarcodex.com/2019/11/12/book-review-the-body-keeps-the-score/>.
- [Fri19] Karl J. Friston. “Waves of prediction”. In: *PLOS Biology* 17.10 (Oct. 2019), pp. 1–7. DOI: 10.1371/journal.pbio.3000426. URL: <https://doi.org/10.1371/journal.pbio.3000426>.
- [Wik20d] Wikipedia contributors. *Injection locking* — *Wikipedia, The Free Encyclopedia*. [Online; accessed 16-April-2020]. 2020. URL: [https://en.wikipedia.org/w/index.php?title=Injection\\_locking&oldid=937233821](https://en.wikipedia.org/w/index.php?title=Injection_locking&oldid=937233821).
- [Wik20e] Wikipedia contributors. *Hebbian theory* — *Wikipedia, The Free Encyclopedia*. [Online; accessed 16-April-2020]. 2020. URL: [https://en.wikipedia.org/w/index.php?title=Hebbian\\_theory&oldid=951033588](https://en.wikipedia.org/w/index.php?title=Hebbian_theory&oldid=951033588).
- [Wik20f] Wikipedia contributors. *Prion* — *Wikipedia, The Free Encyclopedia*. [Online; accessed 16-April-2020]. 2020. URL: <https://en.wikipedia.org/w/index.php?title=Prion&oldid=951106584>.
- [Ale17] Scott Alexander. *Book Review: Behavior – The Control Of Perception*. [Online; accessed 16-April-2020]. 2017. URL: <https://slatestarcodex.com/2017/03/06/book-review-behavior-the-control-of-perception/>.
- [PKK10] Belinda Pletzer, Hubert Kerschbaum, and Wolfgang Klimesch. “When frequencies never synchronize: The golden mean and the resting EEG”. In: *Brain Research* 1335 (2010), pp. 91–102. ISSN: 0006-8993. DOI: <https://doi.org/10.1016/j.brainres.2010.03.074>. URL: <http://www.sciencedirect.com/science/article/pii/S0006899310007092>.
- [Emi18a] Andrés Gómez Emilsson. *What is Love? Neural Annealing in the Presence of an Intentional Object*. 2018. URL: <https://qualiacomputing.com/2018/12/22/what-is-love-neural-annealing-in-the-presence-of-an-intentional-object/>.
- [Psy19] Psychonaut contributors. *Uncomfortable physical effects*. 2019. URL: [https://psychonautwiki.org/w/index.php?title=Uncomfortable\\_physical\\_effects&oldid=132802](https://psychonautwiki.org/w/index.php?title=Uncomfortable_physical_effects&oldid=132802).

- [Emi19] Andrés Gómez Emilsson. *The Hyperbolic Geometry of DMT Experiences (@Harvard Science of Psychedelics Club)*. Youtube. 2019. URL: <https://www.youtube.com/watch?v=loCBvaj4eSg&t=2069>.
- [Emi20a] Andrés Gómez Emilsson. *Harmonic Society (3/4): Art as State-Space Exploration and Energy Parameter Modulation*. 2020. URL: <https://qualiacomputing.com/2019/09/30/harmonic-society-3-4-art-as-state-space-exploration-and-energy-parameter-modulation/>.
- [Bur18] Lana Burgess. “What is hallucinogen-persisting perception disorder?” In: *MedicalNewsToday* (2018).
- [Zam+18] Federico Zamberlan et al. “The Varieties of the Psychedelic Experience: A Preliminary Study of the Association Between the Reported Subjective Effects and the Binding Affinity Profiles of Substituted Phenethylamines and Tryptamines”. In: *Frontiers in Integrative Neuroscience* 12 (2018), p. 54. ISSN: 1662-5145. DOI: 10.3389/fnint.2018.00054. URL: <https://www.frontiersin.org/article/10.3389/fnint.2018.00054>.
- [Emi18b] Andrés Gómez Emilsson. *The Phenomenal Character of LSD + MDMA (Candy-Flipping) According to Cognitive Scientist Steve Lehar*. 2018. URL: <https://qualiacomputing.com/2018/12/12/the-phenomenal-character-of-lsd-mdma-candy-flipping-according-to-cognitive-scientist-steve-lehar/>.
- [Yud] Eliezer Yudkowsky. *Coherent extrapolated volition (alignment target)*. Arbital. URL: <https://arbital.com/p/cev/>.
- [Joh18c] Michael Edward Johnson. *The Neuroscience of Meditation: Four Models*. 2018. URL: <https://opentheory.net/2018/12/the-neuroscience-of-meditation/>.
- [Emi16a] Andrés Gómez Emilsson. *Psychedelic alignment cascades*. 2016. URL: <https://qualiacomputing.com/2017/06/18/connectome-specific-harmonic-waves-on-lsd/>.
- [Emi20b] Andrés Gómez Emilsson. *Ephemerisle: Health Homeostasis, Worldview Annealing, and the Long-Tails of Serious Fun*. 2020. URL: <https://qualiacomputing.com/2019/07/30/ephemerisle-health-homeostasis-worldview-annealing-and-the-long-tails-of-serious-fun/>.
- [Wik20g] Wikipedia contributors. *Neoteny in humans — Wikipedia, The Free Encyclopedia*. [Online; accessed 16-April-2020]. 2020. URL: [https://en.wikipedia.org/w/index.php?title=Neoteny\\_in\\_humans&oldid=951250041](https://en.wikipedia.org/w/index.php?title=Neoteny_in_humans&oldid=951250041).
- [Wik20h] Wikipedia contributors. *Big Five personality traits — Wikipedia, The Free Encyclopedia*. [Online; accessed 16-April-2020]. 2020. URL: [https://en.wikipedia.org/w/index.php?title=Big\\_Five\\_personality\\_traits&oldid=951276015](https://en.wikipedia.org/w/index.php?title=Big_Five_personality_traits&oldid=951276015).
- [Mus07] Janek Musek. “A general factor of personality: Evidence for the Big One in the five-factor model”. In: *Journal of Research in Personality* 41.6 (2007), pp. 1213–1233. ISSN: 0092-6566. DOI: <https://doi.org/10.1016/j.jrp.2007.02.003>. URL: <http://www.sciencedirect.com/science/article/pii/S0092656607000256>.
- [Emi16b] Andrés Gómez Emilsson. *Algorithmic Reduction of Psychedelic States*. 2016. URL: <https://qualiacomputing.com/2016/06/20/algorithmic-reduction-of-psychedelic-states/>.