The Mislabel Hypothesis

Unhappiness as the Brain's Misinterpretation of Gut-Derived Systemic Dysfunction Signals

Atom McCree[1,*], Claude Opus 4.6[2], ChatGPT 5.4[3]

[1] *Independent Researcher, ÆoNs Research Laboratory* [2] *Anthropic* [3] *OpenAI*

[*] *Corresponding author* | *April 2026*

Abstract

*We propose the **Mislabel Hypothesis**: that a significant component of what humans experience as persistent unhappiness or low-grade depression is not a primary emotional state but a **misclassified interoceptive signal** originating from gut microbiome dysfunction. The gut functions as a continuous, whole-body health sensor. When dysbiosis occurs — reflecting broader systemic breakdown in metabolic, immune, or circadian integrity — the microbiome broadcasts distress signals via vagal, immune, endocrine, and neurotransmitter pathways. The human brain lacks a dedicated visceral-status readout and instead integrates these signals into emotional valence, defaulting to the nearest affective category: I am unhappy. This reframing converts a subjective philosophical problem (why am I unhappy?) into a measurable physiological one (what systemic dysfunction is the gut detecting, and how is the brain mislabeling its signal?). We present the mechanistic basis, supporting evidence from microbiota-gut-brain axis (MGBA) research, five falsifiable predictions, and a dual-channel intervention framework.*

Keywords: microbiota-gut-brain axis, interoception, dysbiosis, depression, mislabel hypothesis, vagus nerve, serotonin, short-chain fatty acids, emotional valence

1. Origin

What if unhappiness isn't an emotion? What if it's a signal — and the brain is reading it wrong?

This paper begins with an observation: the global explosion of depression and persistent low mood cannot be fully explained by external circumstances. Diagnostic capacity has improved, economic conditions in much of the world have improved, and access to social connection (at least digitally) has never been greater. Yet reported unhappiness continues to rise. Something upstream is broken.

The microbiota-gut-brain axis (MGBA) literature of the last decade has established, beyond reasonable dispute, that gut microbiome composition causally influences mood, anxiety, and depressive behavior. What has not been stated plainly enough is the implication: **the brain may be systematically misclassifying an internal health-status signal as an emotional state**. The gut is not causing unhappiness. The gut is *reporting dysfunction*, and the brain — lacking a finer-grained interpreter for visceral afferent signals — is labeling that report "unhappiness."

We call this the Mislabel Hypothesis.

2. The Gut as Systemic Health Sensor

The human gastrointestinal tract hosts approximately 3.8×10^{13} microorganisms — roughly equal to the number of human cells in the body [1]. This is not passive cargo. The microbiome is a metabolically active

organ-equivalent that continuously monitors and responds to the state of the host across multiple physiological axes.

2.1 What the Gut Senses

The microbiome is sensitive to, and shifts composition in response to: dietary composition and caloric sufficiency, circadian alignment (meal timing, light exposure, sleep quality), systemic inflammation, hormonal status (cortisol, sex hormones, thyroid), immune function, pharmaceutical exposure (especially antibiotics, PPIs, NSAIDs), and psychological stress via bidirectional vagal and HPA axis signaling [2, 3]. In this sense, the microbiome is not a single-purpose digestive aid. It is a **multi-modal sensor array** that integrates whole-body status into a composite signal.

2.2 Dysbiosis as a Dysfunction Flag

When the host system is in a degraded state — chronic stress, poor sleep, inflammatory diet, circadian disruption, sedentary behavior, or any combination — the microbiome shifts toward a dysbiotic profile: reduced alpha-diversity, depletion of short-chain fatty acid (SCFA) producing taxa (*Lactobacillus*, *Bifidobacterium*, *Faecalibacterium prausnitzii*, *Lachnospiraceae*), increased abundance of pro-inflammatory taxa, and compromised intestinal barrier integrity [4, 5].

Crucially, this shift is not random noise. It is a **lawful response** to systemic conditions. Dysbiosis is the microbiome's honest report that something is wrong with the host. The question is: what does the brain do with that report?

3. Signal Transmission: Four Pathways to the Brain

The gut communicates its status to the central nervous system via four convergent, partially redundant pathways. Each carries a different aspect of the dysfunction signal. Together, they produce a multi-channel broadcast that the brain must integrate.

3.1 The Vagal Neural Pathway

The vagus nerve is the primary neural conduit between the enteric nervous system (ENS) and the brainstem. Microbial metabolites — SCFAs (butyrate, propionate, acetate), secondary bile acids, indole derivatives, and serotonin precursors (tryptophan, 5-HTP) — stimulate enteroendocrine cells and enteric neurons, which fire afferent signals via the vagus to the nucleus tractus solitarius (NTS). From there, signals propagate to the locus coeruleus, raphe nuclei, amygdala, and prefrontal cortex [6, 7]. Vagotomy studies in animal models consistently attenuate or eliminate the mood effects of both probiotics and pathogenic bacteria, confirming this pathway as a primary channel [8].

3.2 The Immune-Inflammatory Pathway

Dysbiosis compromises the intestinal epithelial barrier ("leaky gut"), allowing bacterial lipopolysaccharide (LPS/endotoxin) to translocate into systemic circulation. This triggers low-grade chronic inflammation: elevated TNF-α, IL-6, IL-1β, and C-reactive protein. Peripheral cytokines cross the blood-brain barrier via active transport and circumventricular organs, activating microglia and initiating neuroinflammation. Neuroinflammation disrupts monoamine synthesis, reduces BDNF expression, and impairs hippocampal neurogenesis — all established correlates of depressive symptomatology [9, 10].

3.3 The Endocrine / HPA Axis Pathway

The hypothalamic-pituitary-adrenal (HPA) axis is bidirectionally coupled to the gut microbiome. Dysbiosis-driven inflammation activates the HPA axis, elevating cortisol. Chronic cortisol elevation further

disrupts the microbiome (cortisol is directly antimicrobial to several commensal species), creating a positive feedback loop. Germ-free animal studies demonstrate exaggerated HPA stress responses that normalize upon colonization with healthy microbiota, confirming the microbiome's role as a stress-response regulator [11, 12].

3.4 The Neurotransmitter Modulation Pathway

Gut bacteria directly synthesize or modulate the precursors of every major mood-relevant neurotransmitter. Approximately 90% of the body's serotonin (5-HT) is produced in the gut by enterochromaffin cells under microbial regulation. Gut bacteria also produce γ-aminobutyric acid (GABA), dopamine, norepinephrine, and acetylcholine, and regulate tryptophan metabolism via the kynurenine pathway — the diversion of tryptophan toward kynurenine (and the neurotoxic metabolite quinolinic acid) rather than serotonin is a hallmark of inflammation-driven depression [13, 14]. When the microbiome is dysbiotic, the neurotransmitter supply chain is degraded at the source.

4. The Mislabel Mechanism

Here is the core claim of this paper.

The four pathways described above converge on brain regions that process **emotional valence**: the prefrontal cortex, amygdala, insula, and anterior cingulate cortex. These regions evolved to integrate interoceptive signals (internal body-state reports) with exteroceptive signals (external environmental cues) to produce a unified affective experience [15]. This integration is the basis of the brain's "how am I doing?" computation.

The problem is that the brain has **no dedicated channel label** for "your gut microbiome is dysbiotic." There is no conscious readout that says "systemic inflammation is elevated" or "SCFA production has dropped 40%." The interoceptive system was not designed for granular diagnostic reporting. It was designed for survival-relevant summary judgments: safe/unsafe, approach/withdraw, well/unwell [16].

When the gut broadcasts a sustained dysfunction signal across all four channels simultaneously — vagal distress, inflammatory load, HPA dysregulation, neurotransmitter depletion — the brain integrates this into the only available affective category that matches: **persistent negative valence**. The subjective experience is: *I am unhappy. Something is wrong with my life. I don't enjoy things anymore. Existence feels heavy.*

This is the mislabel. The signal is real. The source is physiological. But the brain's interpretation is emotional. The person searches for *reasons* for their unhappiness — relationship problems, career dissatisfaction, existential ennui — because the brain's narrative engine requires a story to explain the feeling. The gut's honest report ("your system is degraded") gets rewritten as a life crisis.

5. Supporting Evidence

5.1 Fecal Microbiota Transplant Studies

The strongest causal evidence comes from fecal microbiota transplant (FMT) experiments. When germ-free mice receive FMT from humans diagnosed with major depressive disorder (MDD), they develop depression-like and anxiety-like behaviors (reduced sucrose preference, increased immobility in forced swim test, decreased exploration) — **without any external stressor** [17, 18]. The microbiome alone is sufficient to induce the behavioral phenotype. Conversely, FMT from healthy donors into dysbiotic recipients reliably reduces depressive behavior in animal models and shows promising results in human clinical trials [19].

5.2 Dysbiosis-Depression Correlation

Large-scale human studies consistently show that individuals with MDD exhibit: significantly lower gut microbial alpha-diversity, depletion of butyrate-producing genera (*Faecalibacterium*, *Coprococcus*), enrichment of pro-inflammatory genera (*Eggerthella*, *Flavonifractor*), elevated circulating LPS and inflammatory markers, and reduced fecal SCFA concentrations. These signatures correlate with symptom severity on standardized scales (PHQ-9, HDRS) [20, 21, 22]. The Flemish Gut Flora Project (n = 1,054) identified *Coprococcus* and *Dialister* as consistently depleted in depression, even after controlling for antidepressant use [23].

5.3 Probiotic and Prebiotic Interventions

Meta-analyses of randomized controlled trials show that targeted probiotics ("psychobiotics") significantly reduce depressive symptoms compared to placebo, with moderate effect sizes (Cohen's $d \approx 0.3–0.5$) [24, 25]. Prebiotic interventions (galactooligosaccharides, fructooligosaccharides) that selectively feed SCFA-producing bacteria also show mood benefits, suggesting the mechanism is genuinely microbiome-mediated rather than a placebo or general health effect [26].

5.4 The Inflammation-Depression Link

Anti-inflammatory agents (e.g., celecoxib, minocycline) show antidepressant effects in RCTs, and anti-TNF-α therapy in autoimmune disease patients reduces depressive symptoms independent of disease improvement [27]. This supports the pathway model: inflammation is not merely correlated with depression — it is upstream of it. Blocking the inflammatory signal attenuates the mood effect, consistent with the mislabel framework (reduce the dysfunction signal $\rightarrow$ the brain has less to mislabel).

5.5 Vagotomy Evidence

Subdiaphragmatic vagotomy in rodents blocks the anxiolytic effects of *Lactobacillus rhamnosus* (JB-1), confirming that the vagus nerve is a necessary conduit for at least one major arm of the gut-to-brain mood signal [8]. This is consistent with the mislabel model: cutting the signal pathway eliminates the downstream emotional mislabeling.

6. Why "Mislabel" and Not Just "Cause"?

One could argue this is merely the gut *causing* depression. We make a stronger claim: the emotional label is incorrect. The distinction matters for three reasons.

First, it redirects the therapeutic target. If unhappiness is an emotion, you treat the emotion (therapy, SSRIs, lifestyle change aimed at the mind). If unhappiness is a mislabeled physiological signal, you treat the signal source (restore microbiome integrity, reduce systemic inflammation, repair the gut barrier). The intervention logic is fundamentally different.

Second, it explains the narrative confabulation. Depressed individuals consistently generate explanatory narratives for their mood ("my job is meaningless," "no one really loves me") that may be partially or entirely post-hoc rationalizations of a signal that has no cognitive content. The mislabel framework predicts this: the brain must explain the feeling, and it constructs the most plausible story available. Fix the gut, and the stories often dissolve — not because the person's life changed, but because the false alarm stopped.

Third, it explains treatment-resistant depression. If a patient's depression is driven by persistent dysbiosis and systemic inflammation, SSRIs (which increase synaptic serotonin availability) are treating the wrong layer. The signal is arriving degraded from the gut; boosting serotonin reuptake inhibition at the synapse does not repair the upstream source. This is consistent with the ~30% non-response rate to first-line antidepressant therapy and the growing clinical interest in anti-inflammatory and microbiome-targeted adjuncts [28].

7. Falsifiable Predictions

The Mislabel Hypothesis generates the following testable predictions. Falsification of any prediction weakens the hypothesis; falsification of Predictions 1 and 2 together would be fatal.

Prediction 1 — Temporal Precedence.

In a longitudinal cohort study with serial stool sampling and validated mood assessments (e.g., weekly PHQ-9), measurable dysbiosis signatures (alpha-diversity drop, SCFA decline, LPS elevation) will precede the onset of self-reported mood deterioration by 7–21 days. If mood change precedes or coincides with dysbiosis onset, the mislabel direction is weakened.

Prediction 2 — Microbiome Restoration Without Life Change.

Targeted microbiome restoration (FMT from healthy donors, or precision prebiotic/probiotic regimen) in patients with treatment-resistant depression and confirmed dysbiosis will reduce depressive symptoms significantly — **without any change to the patient's external life circumstances, cognitive therapy, or medication**. If symptoms improve, and the patient spontaneously reports that their previous "reasons" for unhappiness now feel less compelling, this directly supports the mislabel mechanism (the narrative was post-hoc).

Prediction 3 — Vagal Decoupling Test.

Temporary, reversible vagal blockade (e.g., non-invasive transcutaneous vagal stimulation at inhibitory frequencies) in individuals with active dysbiosis should attenuate the emotional valence of the gut signal without resolving the underlying dysfunction. Mood should temporarily improve while inflammatory and metabolic markers remain unchanged.

Prediction 4 — Interoceptive Specificity Gradient.

Individuals with higher interoceptive accuracy (as measured by heartbeat detection tasks or interoceptive confidence scales) will show **weaker** emotional mislabeling of gut distress — i.e., they will report physical discomfort rather than emotional unhappiness in response to dysbiosis. If high-interoceptive-accuracy individuals mislabel at the same rate, the mislabel mechanism is weakened.

Prediction 5 — Inflammatory Signature Specificity.

The subset of depressed patients with elevated gut-derived inflammatory markers (LPS, zonulin, calprotectin) will respond preferentially to anti-inflammatory and microbiome-targeted interventions over SSRIs, while the subset with normal gut markers will show the reverse pattern. This predicts a stratifiable treatment response based on signal source.

8. Intervention Framework: Dual-Channel Repair

If the Mislabel Hypothesis is correct, effective treatment of gut-driven unhappiness requires addressing both the signal source and the brain's interpretation of that signal.

8.1 Channel 1 — Repair the Signal Source

Restore microbiome integrity through: dietary intervention (fiber-rich, polyphenol-rich, fermented foods; elimination of ultra-processed foods and emulsifiers that damage the mucosal barrier); targeted probiotics (evidence-based strains: *L. rhamnosus* JB-1, *B. longum* 1714, *L. plantarum* PS128); prebiotic substrates (GOS, FOS, resistant starch) to selectively feed SCFA producers; circadian re-alignment (consistent sleep-wake cycle, morning light exposure, time-restricted eating); and in severe cases, FMT from screened healthy donors [19, 24, 26].

8.2 Channel 2 — Retrain the Interpreter

While the signal source is being repaired, help the brain distinguish between physiological discomfort and emotional unhappiness through: interoceptive training (mindfulness-based interoceptive awareness, body-scan meditation) to improve the brain's granularity in reading internal signals; psychoeducation (explicitly teaching patients that their "unhappiness" may be a mislabeled body signal, which alone can reduce catastrophic interpretation); and meta-cognitive reframing (when the feeling arises, the question shifts from "why am I unhappy?" to "what is my body reporting?").

The dual-channel approach predicts faster and more durable recovery than either channel alone, because it simultaneously reduces the distress signal and improves the brain's ability to correctly classify whatever residual signal remains.

9. Scope and Limitations

The Mislabel Hypothesis does not claim that all unhappiness is gut-derived. Grief, loss, social isolation, trauma, and genuine existential suffering are real and do not require a microbiome explanation. The hypothesis targets **persistent, low-grade, narratively confabulated unhappiness** — the kind where the person cannot clearly identify a cause, where explanations feel insufficient to the weight of the feeling, and where the mood persists despite favorable life circumstances. This is the phenotype most consistent with a mislabeled physiological signal.

Additionally, the gut-brain axis is bidirectional. Psychological stress genuinely damages the microbiome (top-down causation), and microbiome dysfunction genuinely alters mood (bottom-up causation). The Mislabel Hypothesis privileges the bottom-up direction for the specific phenotype described above but does not deny the existence of top-down pathways. In many cases, the system is locked in a bidirectional feedback loop, and breaking in at either point may be effective.

10. Conclusion

The gut is not a passive organ. It is the body's most comprehensive internal sensor, monitoring metabolic, immune, circadian, and hormonal status in real time and broadcasting its findings to the brain via four convergent pathways. When it detects systemic dysfunction, it does what any good sensor does: it sends an alarm.

The brain receives that alarm and does what it has always done: it feels. But feeling is not reading. The brain cannot display a dashboard of inflammatory markers and SCFA levels. It can only say: *something is wrong*. And because the signal is persistent, diffuse, and lacks cognitive content, the brain labels it with the closest available word: **unhappiness**.

The Mislabel Hypothesis proposes that this labeling error is not incidental but central to the modern crisis of persistent low mood. Fix the gut, and you fix the signal. Train the brain, and you fix the interpretation. Do both, and the mislabel dissolves — not because life changed, but because the body stopped sending a false alarm and the brain stopped reading it wrong.

Unhappiness, in this framework, is not a verdict. It is a mistranslation.

References

[1] Sender, R., Fuchs, S. & Milo, R. Revised estimates for the number of human and bacteria cells in the body. *Cell* **164**(3), 337–340 (2016).

[2] Cryan, J.F. et al. The microbiota-gut-brain axis. *Physiol. Rev.* **99**(4), 1877–2013 (2019).

[3] Foster, J.A. & McVey Neufeld, K.-A. Gut-brain axis: how the microbiome influences anxiety and depression. *Trends Neurosci.* **36**(5), 305–312 (2013).

[4] Valles-Colomer, M. et al. The neuroactive potential of the human gut microbiota in quality of life and depression. *Nat. Microbiol.* **4**(4), 623–632 (2019).

[5] Simpson, C.A. et al. The gut microbiota in anxiety and depression — A systematic review. *Clin. Psychol. Rev.* **83**, 101943 (2021).

[6] Bravo, J.A. et al. Ingestion of *Lactobacillus* strain regulates emotional behavior and central GABA receptor expression in a mouse via the vagus nerve. *Proc. Natl. Acad. Sci.* **108**(38), 16050–16055 (2011).

[7] Bonaz, B., Bazin, T. & Pellissier, S. The vagus nerve at the interface of the microbiota-gut-brain axis. *Front. Neurosci.* **12**, 49 (2018).

[8] Bravo, J.A. et al. (2011), *ibid.* Vagotomy abolishes the anxiolytic and antidepressant effects of *L. rhamnosus* JB-1.

[9] Kelly, J.R. et al. Transferring the blues: depression-associated gut microbiota induces neurobehavioural changes in the rat. *J. Psychiatr. Res.* **82**, 109–118 (2016).

[10] Miller, A.H. & Raison, C.L. The role of inflammation in depression: from evolutionary imperative to modern treatment target. *Nat. Rev. Immunol.* **16**(1), 22–34 (2016).

[11] Sudo, N. et al. Postnatal microbial colonization programs the hypothalamic-pituitary-adrenal system for stress response in mice. *J. Physiol.* **558**(1), 263–275 (2004).

[12] Farzi, A., Fröhlich, E.E. & Holzer, P. Gut microbiota and the neuroendocrine system. *Neurotherapeutics* **15**(1), 5–22 (2018).

[13] Yano, J.M. et al. Indigenous bacteria from the gut microbiota regulate host serotonin biosynthesis. *Cell* **161**(2), 264–276 (2015).

[14] Cervenka, I., Agudelo, L.Z. & Ruas, J.L. Kynurenines: tryptophan's metabolites in exercise, inflammation, and mental health. *Science* **357**(6349), eaaf9794 (2017).

[15] Craig, A.D. How do you feel? Interoception: the sense of the physiological condition of the body. *Nat. Rev. Neurosci.* **3**(8), 655–666 (2002).

[16] Barrett, L.F. & Simmons, W.K. Interoceptive predictions in the brain. *Nat. Rev. Neurosci.* **16**(7), 419–429 (2015).

[17] Zheng, P. et al. Gut microbiome remodeling induces depressive-like behaviors through a pathway mediated by the host's metabolism. *Mol. Psychiatry* **21**(6), 786–796 (2016).

[18] Kelly, J.R. et al. (2016), *ibid.*

[19] Chinna Meyyappan, A. et al. Effect of fecal microbiota transplant on symptoms of psychiatric disorders: a systematic review. *BMC Psychiatry* **20**, 299 (2020).

[20] Jiang, H. et al. Altered fecal microbiota composition in patients with major depressive disorder. *Brain Behav. Immun.* **48**, 186–194 (2015).

[21] Naseribafrouei, A. et al. Correlation between the human fecal microbiota and depression. *Neurogastroenterol. Motil.* **26**(8), 1155–1162 (2014).

[22] Sanada, K. et al. Gut microbiota and major depressive disorder: a systematic review and meta-analysis. *J. Affect. Disord.* **266**, 1–13 (2020).

[23] Valles-Colomer, M. et al. (2019), *ibid.*

[24] Nikolova, V.L. et al. Gut feeling: randomized controlled trials of probiotics for the treatment of clinical depression: systematic review and meta-analysis. *Ther. Adv. Psychopharmacol.* **9**, 2045125319859963 (2019).

[25] Liu, R.T., Walsh, R.F.L. & Sheehan, A.E. Prebiotics and probiotics for depression and anxiety: a systematic review and meta-analysis. *BMJ* **378**, e071742 (2022).

[26] Schmidt, K. et al. Prebiotic intake reduces the waking cortisol response and alters emotional bias in healthy volunteers. *Psychopharmacology* **232**(10), 1793–1801 (2015).

[27] Köhler-Forsberg, O. et al. Efficacy of anti-inflammatory treatment on major depressive disorder or depressive symptoms: meta-analysis. *Acta Psychiatr. Scand.* **139**(5), 404–419 (2019).

[28] Strawbridge, R. et al. Inflammation and clinical response to treatment in depression: a meta-analysis. *Eur. Neuropsychopharmacol.* **25**(10), 1532–1543 (2015).

Manuscript prepared April 2026. ÆoNs Research.