OPEN ACCESS Freely available online

Theoretical Article (Hypothesis & Theory) Vol. 2, Issue 2, July-December 2026 | ISSN: 3107-9024

Journal of Psychiatry and Psychological Sciences

What Do Theories of Depression Actually Explain? A Systematic Symptom-Level Audit of 25 Etiological Frameworks

Reuben Vaisman Tzachor

Department of Clinical and Forensic Psychology, Counseling Center of Santa Monica, Los Angeles, California, USA

***Correspondence to:** Reuben Vaisman Tzachor, Department of Clinical and Forensic Psychology, Counseling Center of Santa Monica, Los Angeles, California, USA, E-mail: vavnik@hotmail.com

Received: September 29, 2026; **Manuscript No:** JPPC-26-7011; **Editor Assigned:** October 01, 2026; **PreQc No:** JPPC-26-7011 (PQ); **Reviewed:** October 16, 2026; **Revised:** November 10, 2026; **Accepted:** November 12, 2026; **Published:** November 23, 2026

Citation: Tzachor RV (2026). What Do Theories of Depression Actually Explain? A Systematic Symptom-Level Audit of 25 Etiological Frameworks. J. Psychiatr. Psychol. Sci. Vol.2 Iss.2, November (2026), pp:292-307.

Copyright: © 2026 Reuben Vaisman Tzachor. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

Theories of depression are conventionally appraised by asking whether they explain depression. If major depressive disorder is a heterogeneous collection of symptoms with partly independent causes, that question is malformed: what matters is which of the symptoms each theory explains. The author audited 25 etiological frameworks neurochemical, neurobiological, cognitive–behavioral, interpersonal–developmental, vulnerability–stress, and evolutionary/structural against a 16-column symptom space comprising the nine DSM-5 criterion symptoms, with the three bidirectional criteria split into their poles, plus four melancholic, atypical and anxious-distress specifier symptoms. Each of the 400 cells was coded DIRECT, SECONDARY, SILENT or CONTRADICTED against primary sources, with a citation and written justification for every non-silent cell. Coverage was thin: 48.3% of cells were silent and 21.0% direct. Depressed mood was the only explanatory hub, directly modelled by 19 of 25 theories. Anhedonia and fatigue attracted claims from 23 and 21 theories but direct mechanisms from only 7 and 6; their apparent consensus is produced by secondary inheritance rather than prediction. Four symptoms were theoretical orphans: psychomotor agitation (0 direct codings, 4 theories contradicting it), hyperphagia/weight gain (0), leaden paralysis (1) and diurnal mood variation (2). Only one framework, network theory, addressed both poles of all three bidirectional criteria. Hierarchical clustering of explanatory profiles did not recover disciplinary families (adjusted Rand index = 0.04); the alliances that emerged crossed them. The author argues that symptom-level coverage should become a routine appraisal criterion, that the bidirectional DSM criteria currently lack etiological warrant, and that the orphaned symptoms mark where RDoC has no construct.

Keywords: Depression; Theory Appraisal; Symptom-Level Analysis; Heterogeneity; DSM-5; Research Domain Criteria

Galley Proof

INTRODUCTION

The Appraisal Question the Field Asks, And Why It Is the Wrong One

The standard way to evaluate a theory of depression is to ask whether the evidence supports it as an account of the disorder. Reviews are organized around that question; so are textbook chapters, funding panels, and the recurrent public arguments about whether the serotonin hypothesis or the inflammatory hypothesis or the cognitive model has been vindicated. The question presupposes that "the disorder" is one thing, with one explanandum, against which competing accounts can be scored.

Two decades of work on the structure of the depressive phenotype suggests it is not. The DSM-5 criteria for major depressive disorder are polythetic: five of nine symptoms, one of which must be depressed mood or anhedonia, with three criteria specified bidirectionally. That combinatorial structure admits hundreds of distinct qualifying profiles, many of which share no symptom beyond the gate criterion [1]. Two patients can both carry the diagnosis while overlapping on a single feature. Aggregation into a sum score conceals this: the score is a count of heterogeneous events treated as interchangeable indicators of one latent severity.

That treatment is not merely inelegant; it is empirically wrong in a way that matters for theory. Individual criterion symptoms have different risk factors within the same cohort, different biological correlates, different associations with impairment, and different courses [2-5]. Depression sum scores, as Fried and Nesse, put it, do not add up: insomnia and suicidal ideation contribute equally to a total while behaving as different phenomena under every analysis that separates them [6]. Factor-analytic work in large general-population samples recovers multiple symptom dimensions rather than a single liability [7]. Depression is not a consistent syndrome, and the categorical diagnosis is a poor proxy for whatever the causal processes are [5].

Kendler's, argument for empirically based pluralism supplies the conceptual frame. Psychiatric causation, on his account, is "dappled": many causes of small effect, operating at several

levels, interacting, partly idiosyncratic to individuals, with no expectation that one mechanism maps onto one diagnostic category [8]. If that picture is right, then a theory that accounts for depressed mood and a theory that accounts for hypersomnia are not competitors at all. They have different explananda. Evaluating both against "depression" obscures what each has actually achieved and lets each borrow credibility from the other's coverage.

Category-Level Validation Versus Symptom-Level Appraisal

The distinction the author wants to draw is between two appraisal regimes. Under category-level validation, a theory earns support when patients with the diagnosis differ from controls on the variable the theory names: cerebrospinal 5-HIAA, dexamethasone non-suppression, C-reactive protein, dysfunctional attitude scores, ventral striatal reward-anticipation signal. The dependent variable is group membership or a sum score, and the theory is credited with explaining depression as a whole. Under symptom-level appraisal, a theory earns support for the specific symptoms whose occurrence, form, or polarity it predicts, and earns nothing for the rest. The two regimes can deliver opposite verdicts. A theory with an impeccable category-level record can turn out, on inspection, to name a mechanism for one symptom and be silent about the other eight. A theory with a thin category-level record can turn out to make sharp, falsifiable claims across the phenotype.

Network theory made symptom-level structure central to psychopathology research by treating disorders as systems of causally interacting symptoms rather than as latent entities that produce symptoms [9-10]. Applied to depression, it yields dynamic-system models of episode onset and maintenance and an empirical program on which symptoms are central, well-connected, or predictive [11-13]. The Research Domain Criteria initiative reorganized research funding around dimensional constructs negative and positive valence, arousal and regulatory systems, cognitive and social processes explicitly to escape categorical diagnosis [14]. Both moves pushed the field toward finer grain. Neither produced a stocktake of the *existing theoretical corpus* at that grain.

The gap is specific. We have symptom-level empirical work and symptom-level measurement models, and we have a century of etiological theory appraised almost entirely at the category level. Nobody has systematically asked, of each major theory, which symptoms it commits to. As a result, the field cannot presently say which depressive symptoms are over-explained, which are explained only by inheritance from an adjacent claim, and which have no dedicated etiological account at all. Nor can it say whether the familiar taxonomy of theories biological versus psychological, neurochemical versus cognitive carries any information about explanatory content, or whether it merely records which department the theorist worked in.

Two Failure Modes That Symptom-Level Appraisal Exposes

Two problems become visible only at symptom grain, and both are consequential for how the field allocates credit.

The first is Secondary Inheritance. A theory that models low mood can describe anhedonia and fatigue as downstream consequences of low mood without predicting anything about them: not their severity relative to other symptoms, not their time course, not which patients show them, not what would remove them. In review prose this reads as coverage. Under appraisal it functions as an unfalsifiable extension, because no symptom-level observation can embarrass it. When many theories inherit the same symptom, the result is the appearance of explanatory consensus about a symptom that almost nobody has modelled a convergence artefact that can be mistaken for accumulated knowledge.

The second is Polarity Asymmetry. Three DSM criteria are bidirectional: sleep may be reduced or increased, appetite and weight may fall or rise, psychomotor activity may slow or accelerate. The diagnostic system treats each pole as an equivalent route to the same criterion. A causal theory has no such license. A mechanism that explains why a depressed person cannot sleep does not thereby explain why a different depressed person sleeps fourteen hours, and a mechanism that explains anorexia and weight loss is under no obligation to explain hyperphagia. Asking which poles each theory covers converts a classificatory convenience into an empirical question about theoretical reach and the answer bears directly on whether the melancholic/atypical distinction, the subtyping with the longest history of claimed treatment-moderating value, has any etiological underwriting.

The Present Study

In this review the author conducts the missing audit. The author treats 25 etiological frameworks as objects of study, specify a 16-column symptom space that separates the poles of the bidirectional criteria and adds four specifier symptoms, and code each of the resulting 400 theory-by-symptom cells for the kind of claim the theory makes, with a citation and a written justification attached to every non-silent cell. The aims were to (1 construct an exhaustive theory-by-symptom matrix with per-cell source justification; (2 identify explanatory hubs and theoretical orphans, distinguishing direct mechanism claims from secondary inheritance; (3 quantify how the bidirectional DSM criteria are covered, separately for biological and for psychological/social frameworks; and (4 test whether theories grouped together by discipline are in fact similar in what they explain. Section Methods describes the taxonomy, the symptom space, and the coding rules. Section Results reports coverage, the hub-and-orphan structure, polarity asymmetry, and the cluster analysis. Section Discussion draws out three sets of implications for how theories are appraised, for how the diagnostic criteria are revised, and for how RDoC's construct set might be extended and states the limits of what an audit of claims can establish.

Throughout, one caveat governs interpretation. This is an audit of what theories claim, not of what is true. Breadth of symptom-level commitment is a property of how a theory was formulated. It is not evidence that the theory is correct, and a narrow theory that is right is worth more than a broad theory that is wrong. What the audit can establish is where the corpus has committed itself, where it has quietly declined to, and where it has committed to the opposite of what patients present with.

Galley Proof

Methods

Design

This was a structured appraisal of a fixed theory set against a fixed symptom set, producing a fully enumerated matrix with per-cell codes, citations and justifications. No human participants or patient data were involved. The coding frame, the coded matrix and the analysis script are provided as supplementary materials so that every reported quantity can be recomputed and every cell inspected against its stated source.

Theory Set and Taxonomy Construction

The sampling frame was a curated map of the principal theoretical accounts of unipolar depression, organized into six families and 26 numbered entries, each entry carrying its founding primary sources and its principal contemporary statements. The six families were: (1) neurochemical, (2) neurobiological, (3) cognitive–behavioral, (4) interpersonal/psychodynamic/developmental, (5) vulnerability–stress, and (6) evolutionary/structural.

One entry was excluded before coding. Heterogeneity arguments and critiques of unitary models the position that "major depression" does not pick out a single condition make no symptom-level mechanism claim and would be SILENT in all 16 cells by construction, contributing 16 uninformative cells and depressing every aggregate. Excluding it yields the 25 frameworks audited here, listed with their coverage counts in Table 1.

Two boundary decisions shaped the row set. First, Granularity: an entry was retained as one row where its variants share a mechanism and differ in emphasis and split where variants make incompatible symptom predictions. In practice this meant the four evolutionary accounts (analytical rumination, social risk, social competition, bargaining) were coded as a single row, as were the anaclitic and introjective elaborations of object-loss theory, while the monoamine-deficiency, serotonin-specific and receptor-adaptation accounts were kept separate because they differ in what they predict about symptom polarity and treatment latency. Second, Credit Assignment Within Rows: where a row aggregates variant, any variant's claim was credited to the row. Both decisions inflate rather than deflate coverage for the aggregated rows, so the headline result that coverage is thin is conservative with respect to them.

Each framework was assigned to exactly one family, and separately to a binary theory group (biological versus psychological/social) for the polarity analysis. The biological group comprised the four neurochemical and seven neurobiological frameworks (n = 11); the psychological/social group comprised the seven cognitive–behavioral, three interpersonal/developmental, two vulnerability–stress and two evolutionary/structural frameworks (n = 14).

Symptom Space

The symptom space has 16 columns. Twelve are criterion columns derived from the nine DSM-5 A-criteria for major depressive disorder, with the three bidirectional criteria split into their two poles:

A1 depressed mood

A2 anhedonia / loss of interest

A3a appetite or weight Loss / A3b hyperphagia or weight Gain

A4a insomnia / A4b hypersomnia

A5a psychomotor Retardation / A5b psychomotor Agitation

A6 fatigue / loss of energy

A7 worthlessness / guilt

A8 concentration difficulty / indecisiveness

A9 suicidal ideation / death wishes

Splitting the bidirectional criteria is the central design decision. Under the DSM's own logic, the two poles are interchangeable routes to one criterion; under any causal logic they are distinct explananda, and a theory that derives one has not thereby derived the other. Collapsing them would make polarity asymmetry invisible and would let a theory that explains insomnia be scored as explaining "sleep disturbance".

Four further columns capture specifier symptoms that carry subtyping and treatment-selection signal but sit outside the A-criteria:

- Leaden paralysis (atypical features specifier) — heavy, leaden sensations in the limbs
- Rejection sensitivity (atypical features specifier) — long-standing interpersonal-rejection sensitivity
- Diurnal mood variation (melancholic features specifier) — regular morning worsening of mood
- Somatic anxiety / autonomic arousal (anxious distress specifier) — restlessness, keyed-up tension, autonomic symptoms

Including these four is what allows the audit to ask whether the theorized part of the phenotype coincides with the clinically discriminating part. The atypical and melancholic specifiers are precisely where the reverse-neurovegetative pattern lives, and that pattern has the strongest claim to biological distinctiveness in the current literature [15-16].

The resulting frame is 25 theories × 16 symptoms = 400 cells, all of which were coded.

Coding Rules

Each cell received exactly one of four codes. The rubric, with its decision rule and analysis weight, is given in Table 5.

DIRECT

The theory explicitly models the symptom: it names a mechanism whose output is that symptom, or it predicts the symptom's presence, form, or polarity as a consequence of the modelled process. The test applied was whether the primary source licenses a symptom-specific prediction that could be wrong. Example: the glucocorticoid-cascade account of melancholia predicts early-morning awakening and altered cortisol diurnal rhythm as consequences of CRH hypersecretion and impaired feedback, so both insomnia and diurnal mood variation were coded DIRECT [17-18].

SECONDARY

Galley Proof

The symptom is presented or implied as a non-specific downstream consequence of the modelled process, without a symptom-specific prediction. The test applied was whether removing the symptom from the theory's stated scope would require any revision of the theory. If not, the code is SECONDARY. Example: monoamine-deficiency accounts describe fatigue and anhedonia as features of the depressive state they explain, but predict nothing distinctive about either, so both were coded SECONDARY rather than DIRECT.

SILENT

No mechanism claim, direct or downstream, could be located in the primary or contemporary sources. Silence was coded positively rather than treated as missing data: a theory that does not address hyperphagia has taken a position on its own scope, and that position is the object of interest.

CONTRADICTED

The theory predicts the opposite polarity, or predicts the symptom's absence, or explicitly excludes it. Example: learned-helplessness accounts have behavioral passivity as their signature output, so psychomotor agitation was coded CONTRADICTED, not SILENT [19-20]. This code is what distinguishes a theory that has nothing to say about a symptom from a theory for which the symptom is anomalous.

The DIRECT/SECONDARY boundary is the most consequential judgement in the scheme and the one most exposed to rater disagreement. The author operationalized it as the falsifiability test stated above does the source license a claim that symptom-level data could contradict and recorded the reasoning for every non-silent cell so that readers can apply a different threshold and recompute. Every non-SILENT cell carries a citation and a one- to three-sentence justification naming the specific claim in the cited work that licenses the code. Codes were assigned against primary sources and canonical contemporary statements. Locators are argument- or section-level rather than verified page images; coding was interpretive, single-rater and unblinded (see Limitations Section).

Analysis

Codes were weighted DIRECT = 2, SECONDARY = 1, SILENT = 0, CONTRADICTED = −1 to produce a 25 × 16 numeric coverage matrix. The weighting treats a contradicting claim as informative in the opposite direction from a direct claim, which is the desirable property for profile comparison: two theories that both contradict agitation are similar in that respect, and both differ from a theory that models it.

Three sets of quantities were derived. Theory-level: counts of each code, breadth (number of symptoms receiving any non-silent claim and a summed coverage score (Table 1. Symptom-level: counts of each code across theories, a silence rate and a direct rate (Table 2. Two pre-specified thresholds were used: a symptom was an *explanatory hub* if more than 10 theories coded it DIRECT, and a *theoretical orphan* if at most 2 theories coded it DIRECT. Polarity-level: for each bidirectional criterion, within each theory group, the number of theories covering each pole directly, covering each pole at all, contradicting either pole, or covering both poles (Table 3.

Theories were then clustered by Euclidean distance on their 16-element coverage vectors using Ward linkage, and cut at k = 5 to match the number of families with non-zero direct coverage. Correspondence between disciplinary family and profile cluster was assessed with the adjusted Rand index (ARI, by comparing mean within-family with mean between-family profile distance, and by computing each theory's nearest neighbor and recording whether it lay inside or outside its own family. Analyses used Python 3 with pandas, numpy, scipy.cluster.hierarchy and scikit-learn; the full script is supplied.

RESULTS

Coverage Is Thin, And Thinner for The Specifiers

The author first asked how much of the symptom space the theoretical corpus addresses at all. Of the 400 cells, 193 (48.3% were SILENT, 115 (28.8% SECONDARY, 84 (21.0% DIRECT and 8 (2.0% CONTRADICTED (Figure 1. The modal relationship between an etiological theory of depression and a symptom of depression is therefore *no claim at all*, and fewer than one cell in four carries a mechanism the theory could be held to.

Silence is distributed unevenly across the phenotype, and the direction of the unevenness is unfavorable. For the 12 criterion columns the silence rate is 42.7% (128 of 300 cells; for the four specifier columns it is 65.0% (65 of 100. The clinical qualifiers that carry most of the treatment-selection signal in current practice are the least theorized part of the phenotype. Within the specifiers, diurnal mood variation is silent for 22 of 25 theories (88.0% and rejection sensitivity for 16 (64.0%.

The biological/psychological contrast is smaller than the field's rhetoric implies and runs against the usual assumption. Among the 11 biological frameworks, 52.8% of cells are silent and 18.2% direct; among the 14 psychological/social frameworks, 44.6% are silent and 23.2% direct. Mean DIRECT codings per theory range from 0.0 in the vulnerability–stress family and 1.75 in the neurochemical family, through 3.57 (neurobiological and 4.00 (cognitive–behavioral, to 4.67 (interpersonal/developmental and 5.00 (evolutionary/structural (Table 1. The families with the highest symptom-level commitment are the two usually treated as the softest.

The biological/psychological contrast is smaller than the field's rhetoric implies and runs against the usual assumption. Among the 11 biological frameworks, 52.8% of cells are silent and 18.2% direct; among the 14 psychological/social frameworks, 44.6% are silent and 23.2% direct. Mean DIRECT codings per theory range from 0.0 in the vulnerability–stress family and 1.75 in the neurochemical family, through 3.57 (neurobiological and 4.00 (cognitive–behavioral, to 4.67 (interpersonal/developmental and 5.00 (evolutionary/structural (Table 1. The families with the highest symptom-level commitment are the two usually treated as the softest.

Galley Proof

ID	§	Framework	Family	DIRECT	SEC.	SILENT	CONTRA.	Breadth	Score
T07	2.3	Inflammatory / cytokine hypothesis	Neurobio-logical		5	2	1	13	21.0
T19	4.1	Psychodynamic /object-loss accounts	Interperson-al/ Devel.		3	5	0	11	19.0
T24	6.1	Evolutionary theories (analytical rumination, social risk, bargaining)	Evolution-ary/ Struct.		3	5	1	10	17.0
T06	2.2	HPA-axis / glucocorticoid cascade	Neurobio-logical		6	1	3	12	18.0
T14	3.3	Hopelessness theory	Cognitive/ Behavioral	6	4	5	1	10	16.0
T13	3.2	Beck's cognitive theory	Cognitive/ Behavioral	5	5	6	0	10	15.0
T09	2.5	Circadian / social-zeitgeber theory	Neurobio-logical		5	6	0	10	15.0
T18	3.7	Computational /predictive-processing (allostatic)	Cognitive/ Behavioral	5	6	5	0	11	16.0
T21	4.3	Attachment / early adversity / developmental programming	Interperson-al/ Devel.		5	7	0	9	13.0
T16	3.5	Behavioral theory / behavioral activation	Cognitive/ Behavioral	4	5	7	0	9	13.0
T17	3.6	Self-regulation / self-discrepancy accounts	Cognitive/ Behavioral	3	7	6	0	10	13.0
T25	6.2	Network theory	Evolution-ary/ Struct.		12	1	0	15	18.0
T15	3.4	Response styles theory (rumination)	Cognitive/ Behavioral	3	6	7	0	9	12.0
T04	1.4	Glutamatergic /synaptic-plasticity hypothesis	Neurochem-ical		2	11	0	5	8.0

Galley Proof

T01	1.1	Catecholami monoamine-deficiency hypothesis	Neurochem-ical		5	8	1	7	9.0
T05	2.1	Neurotrophic hypothesis	Neurobio-logical		2	12	0	4	6.0
T02	1.2	Serotonin hypothesis	Neurochem-ical		6	8	0	8	10.0
T12	3.1	Learned helplessness (attributional reformulation)	Cognitive/ Behavioral	2	5	8	1	7	9.0
T08	2.4	Reward-circuit / anhedonia account	Neurobio-logical		4	10	0	6	8.0
T20	4.2	Interpersonal theories (incl. social signal transduction)	Interpersonal/ Devel.		6	8	0	8	10.0
T11	2.7	Microbiome–gut–brain axis	Neurobio-logical		4	11	0	5	6.0
T10	2.6	Mitochondrial /bioenergetic hypothesis	Neurobio-logical		5	10	0	6	7.0
T03	1.3	Receptor-adaptation / post-receptor signaling	Neurochem-ical		2	14	0	2	2.0
T23	5.2	Kindling / stress sensitization	Vulnerability–Stress		1	15	0	1	1.0
T22	5.1	Diathesis–stress and gene × environment interaction	Vulnerability–Stress		1	15	0	1	1.0

Table 1: Symptom-level coverage by theory, ordered by number of DIRECT codings. Breadth = number of the 16 symptoms receiving any non-silent claim. Coverage score = Σ(DIRECT × 2 + SECONDARY × 1 + CONTRADICTED × −1).

Three frameworks have zero DIRECT codings. Receptor-adaptation and post-receptor signaling accounts make claims about the *latency* of antidepressant response rather than about symptoms; diathesis–stress/G×E and kindling make claims about *who* becomes depressed and *when recurrence occurs*, taking the diagnostic category as the outcome. These are coherent divisions of labor, and the author does not read the zeros as defects. Rather, they mean that three frameworks in wide contemporary use constrain no symptom-level prediction whatever and cannot be confirmed or embarrassed by any observation about the form a depressive episode takes.

At the other extreme, the widest DIRECT coverage belongs to the inflammatory/cytokine hypothesis and to psychodynamic object-loss accounts (8 each), then evolutionary theories (7), then the HPA-axis and hopelessness accounts (6 each). That a 1917 account and an immunological account tie for breadth is not a claim about their truth. It reflects a difference in how theories were formulated: both were built to explain a *described*

Galley Proof

clinical picture, and so made claims about sleep, appetite, activity and self-reproach, whereas several modern frameworks were built to explain a mechanism and are silent about most of the phenotypes. Breadth here indexes the ambition of the original formulation, not its warrant.

Network theory is the instructive outlier on the score column. It has only 3 DIRECT codings but the highest breadth in the matrix (15 of 16 symptoms receive some claim) and 12 SECONDARY codings, giving it a coverage score (18.0) equal to the HPA axis's. This is a direct consequence of its architecture: because symptoms are nodes whose explanation is relational, every symptom is inside the theory's scope and almost none is derived from a mechanism. Whether that counts as broad coverage or as no coverage is exactly the question Section For theory appraisal: report coverage, and stop crediting inheritance takes up.

One Hub, Not Three and The Secondary-Inheritance Illusion

The author next asked which symptoms are the corpus's shared explanatory targets. Under the pre-specified criterion (more than 10 theories coding DIRECT), exactly one symptom qualifies: depressed mood, directly modelled by 19 of 25 theories and receiving some claim from all 25, with a silence rate of zero (Table 2). It is the only point in the phenotype on which the corpus converges.

Symptom	Source	DIRECT	SEC.	SILENT	CONTRA.	Any claim	Silence rate	Direct rate
Depressed mood	DSM-5 A1	19	6	0	0	25	0.00	0.76
Worthlessness/ guilt	DSM-5 A7	8	4	13	0	12	0.52	0.32
Psychomotor retardation	DSM-5 A5a	7	9	9	0	16	0.36	0.28
Anhedonia / loss of interest	DSM-5 A2	7	16	2	0	23	0.08	0.28
Concentration/ indecisiveness	DSM-5 A8	7	10	7	1	17	0.28	0.28
Fatigue / loss of energy	DSM-5 A6	6	15	4	0	21	0.16	0.24
Suicidal ideation / death wishes	DSM-5 A9	6	7	12	0	13	0.48	0.24
Insomnia	DSM-5 A4a	6	7	12	0	13	0.48	0.24
Rejection sensitivity	Atypical specifier	5	4	16	0	9	0.64	0.20
Somatic anxiety / autonomic arousal	Anxious-distress specifier	4	6	15	0	10	0.60	0.16
Appetite / weight loss	DSM-5 A3a	3	4	18	0	7	0.72	0.12
Hypersomnia	DSM-5 A4b	3	3	18	1	6	0.72	0.12
Diurnal mood variation	Melancholic specifier	2	1	22	0	3	0.88	0.08
Leaden paralysis	Atypical specifier	1	11	12	1	12	0.48	0.04
Psychomotor agitation	DSM-5 A5b	0	5	16	4	5	0.64	0.00
Hyperphagia / weight gain	DSM-5 A3b	0	7	17	1	7	0.68	0.00

Table 2: Coverage by symptom, ordered by number of DIRECT codings. Any claim = DIRECT + SECONDARY + CONTRADICTED. Rates are proportions of the 25 theories.

Galley Proof

The two symptoms usually assumed to be equally over-determined are not. Anhedonia receives a claim from 23 of 25 theories a silence rate of 8%, the lowest in the matrix after depressed mood but a DIRECT mechanism from only 7; the remaining 16 are SECONDARY. Fatigue receives a claim from 21 theories but a DIRECT mechanism from only 6, with 15 SECONDARY. The ratio is the finding: for anhedonia, secondary claims outnumber direct ones better than two to one; for fatigue, two and a half to one.

This is the audit's most consequential descriptive result, and it is a result about the structure of apparent consensus rather than about any single theory. If one reads the literature at the level of "does this theory address anhedonia," almost every framework does, and anhedonia looks like the best-explained feature of depression after low mood. If one asks instead which frameworks predict something about anhedonia that symptom-level data could contradict, the number collapses to seven. The high any-claim count is produced by theories that inherit anhedonia as a downstream consequence of a mood or activity mechanism. Inheritance costs nothing and buys the appearance of coverage; because it generates no symptom-specific prediction, no observation about anhedonia can disconfirm it.

Dedicated mechanisms do exist for both symptoms, and they are a minority position rather than a consensus. For anhedonia: mesolimbic reward-circuit accounts specify blunted reward anticipation and reduced effort-based responding, with reinforcement-learning formalizations that make quantitative predictions (Huys et al., 2013) [21-24]. For fatigue: allostatic self-efficacy models derive fatigue from a metacognitive assessment of failing interoceptive control, and inflammatory accounts derive it from cytokine-induced sickness behavior and dopaminergic effects on effort [25-28]. Seven theories and six theories, respectively, out of 25.

The same logic sharpens the reading of the rest of Table 2. Psychomotor retardation has 7 DIRECT and 9 SECONDARY codings, so roughly half its apparent coverage is inheritance. Worthlessness/guilt has an unusual profile 8 DIRECT against only 4 SECONDARY because it is hard to inherit: a theory either has an account of self-devaluation or it does not, which is why 13 theories are simply silent about it. Leaden paralysis is the extreme case of the pattern, with 1 DIRECT against 11 SECONDARY (Section Four theoretical orphans).

Four Theoretical Orphans

The author then asked which symptoms fall below the orphan threshold (at most 2 theories coding DIRECT), and examined each against what is empirically known about its prevalence and correlates. Four symptoms qualify: psychomotor agitation, hyperphagia/weight gain, leaden paralysis and diurnal mood variation (Table 2, bottom four rows). Their common feature is that they are exactly the features that carry the melancholic/atypical distinction.

Psychomotor Agitation is the starkest case in the matrix: 0 DIRECT codings, 5 SECONDARY, 4 CONTRADICTED, 16 SILENT. It is the only symptom that no framework models and that several actively predict against. The four contradictions are not incidental to the theories concerned but follow from their central commitments. Monoamine-deficiency accounts define depression as a catecholamine deficiency state and assign catecholamine *excess* to mania, so a deficiency predicts reduced motor activation [29-30]. Inflammatory accounts derive the depressive phenotype from sickness behavior, which is an energy-conserving hypoactivity program [27]. Learned-helplessness accounts have behavioral passivity as their signature output [19-20]. Hopelessness theory predicts retardation of voluntary responding and excludes agitation from its specified symptom set [31]. The five secondary claims are partial and in two cases equivocal: CRH-mediated hyperarousal in melancholia [17,18] low serotonergic function with behavioral disinhibition actual/ought discrepancy, whose "agitation-related emotions" are an affective rather than a psychomotor construct and so do not transfer without equivocation; clinical description of agitated self-reproach in the psychodynamic literature; and network theory's observation that agitation is among the least central DSM symptoms (Fried et al., 2016) a claim about its position in a symptom network, not about its production [12,32-35].

Against this, agitation is neither rare nor prognostically inert. In a consecutive series of 361 patients with major depressive disorder, 94 met Research Diagnostic Criteria for agitated depression, and the agitated group differed in phenomenology and outcome from matched non-agitated patients [36]. In a further MDD sample of 189, 74 were classified as agitated, and agitation predicted subsequent mood switching [37]. Agitated presentations have been argued to constitute a depressive mixed state with distinct treatment implications, including the antidepressant-suicidality question, and psychomotor disturbance more broadly is a well-documented transdiagnostic feature (Song et al., 2025) [38-39]. A symptom present in roughly a quarter of depressed patients, associated with switching risk and with a live treatment controversy, is the one symptom for which the etiological corpus offers no mechanism and four counter-predictions.

Hyperphagia / Weight Gain has 0 DIRECT codings, 7 SECONDARY, 1 CONTRADICTED and 17 SILENT. The seven secondary claims all route through association rather than mechanism: serotonergic satiety regulation with a carbohydrate-craving extension, immunometabolic association, the seasonal-affective phenotype, microbiome effects on host metabolism, interoceptive/allostatic budgeting that predicts dysregulated appetite without specifying *which pole appears in whom*, early-adversity associations with atypical features, and network theory's treatment of it as a node [26,33,40-46]. The single contradiction is the glucocorticoid-cascade account, whose anorexigenic CRH mechanism predicts the opposite [18].

The empirical record here is if anything stronger than for agitation. Atypical depression defined by hyperphagia, weight gain, hypersomnia and leaden paralysis is a recognized specifier with substantial prevalence, and the increased-appetite/weight dimension specifically indexes immunometabolic dysregulation: it shows greater genetic overlap with body mass index and C-reactive protein than other depressive presentations, tracks metabolic-syndrome components and

Galley Proof

inflammatory markers at the individual-symptom level, shows symptom-specific and potentially causal associations with CRP and IL-6 across UK Biobank and NESDA, shares genetic predisposition with immunometabolism traits, emerges from data-driven symptom-profile clustering as an immunometabolic subtype, and predicts poorer antidepressant outcomes across four pooled trials [3,4,15,16,41,47-49]. Leptin dysregulation is specifically associated with the atypical presentation [50]. The field has, in other words, a well-replicated biological signature for a symptom pole that no etiological framework in this set models directly.

Leaden Paralysis has 1 DIRECT coding against 11 SECONDARY, 12 SILENT and 1 CONTRADICTED the clearest instance in the matrix of a symptom that every energy-related theory can gesture at and almost none models. The single direct account is the immunometabolic variant of the inflammatory hypothesis [41]. The eleven secondary claims span monoamine deficiency, reward-circuit effort accounts, circadian models, mitochondrial bioenergetics, learned helplessness, Beck's, "paralysis of the will," hopelessness theory, behavioral activation, allostatic self-efficacy, the social risk hypothesis and network theory [22,25,44,51,52,53,54]. The contradiction is again the HPA account [18]. Eleven theories find leaden paralysis compatible with their mechanism; one predicts it.

Diurnal Mood Variation has 2 DIRECT codings, 1 SECONDARY and 22 SILENT the highest silence rate in the matrix (88%). The two direct accounts are the HPA axis, via cortisol diurnal-rhythm alteration, and circadian/social-zeitgeber theory, via phase relations between circadian and sleep–wake systems [17,44,55]. Network theory supplies the one secondary claim. This is a defining melancholic feature in the operational criteria, studied with morning–evening designs since at least the early 1990s, characterized in large treatment cohorts, measured with experience sampling in daily life, and reviewed as a core component of depressive phenomenology [56-60]. Twenty-three of 25 etiological frameworks have nothing to say about it.

Read together, the four orphans describe a coherent gap rather than four scattered ones. Three of them hyperphagia, leaden paralysis, hypersomnia's companion features constitute the atypical/immunometabolic presentation; agitation and diurnal variation sit at the melancholic end. The corpus is silent or counter-predictive precisely where the subtyping literature is most active.

Polarity Coverage Is Asymmetric, Not Bidirectional

The bidirectional criteria are the natural test of whether biological and psychological models between them span the phenotype: if each family covers one pole, their union might still be complete. This was tested directly by counting, within each theory group, how many frameworks cover each pole and how many cover both (Table 3).

Criterion	Pole A	Pole B	Theory group	n	A direct	B direct	A any	B any	Contradicts either	Covers both poles
Sleep	Insomnia	Hy-persomnia	Biolog-ical	11	2	2	5	2	1	2
Sleep	Insomnia	Hy-persomnia	Psy-chological/ social	4	1	8	4	0	3	
Appetite/ weight	Appetite / weight loss	Hy-perphagia / weight gain	Biolog-ical	11	2	0	2	4	1	1
Appetite/ weight	Appetite / weight loss	Hy-perphagia / weight gain	Psy-chological/ social	14	1	0	5	3	0	2
Psy-chomotor	Psy-chomotor retarda-tion	Psy-chomotor agitation	Biological	11	2	0	6	2	2	1
Psy-chomotor	Psy-chomotor retarda-tion	Psy-chomotor agitation	Psy-chological/ social	14	5	0	10	3	2	3

Table 3: Coverage of the three bidirectional DSM-5 criteria, by theory group. "Any" = DIRECT + SECONDARY. "Both poles" = frameworks with a non-silent claim on each pole.

The union is not complete. Among the 11 biological frameworks, 2 cover both sleep poles, 1 covers both appetite poles and 1 covers both psychomotor poles; among the 14 psychological/social frameworks the corresponding counts are 3, 2 and 3. Aggregating across groups, a total of 5, 3 and 4 frameworks respectively address both poles of a given criterion between 12% and 20% of the corpus. The retardation pole attracts 7 direct codings across both groups; the agitation pole attracts none. The appetite-loss pole attracts 3; the hyperphagia pole none.

Exactly one framework of 25 addresses both poles of all three bidirectional criteria: network theory. It does so by admitting each pole as a separate node whose explanation is relational, not by deriving either pole from a mechanism which is why its 15-symptom breadth coexists with 3 DIRECT codings (Table 1) [10,12]. The only framework that spans the bidirectional phenotype is the one that declines to explain symptoms individually.

Galley Proof

The expected division of labor between HPA-axis and inflammatory/metabolic models is real but much narrower than assumed. The glucocorticoid-cascade account delivers the melancholic pole directly insomnia with early-morning awakening, anorexigenic CRH-driven weight loss, altered cortisol diurnal rhythm and explicitly *contradicts* the atypical pole, with hypersomnia, weight gain and leaden paralysis all coded CONTRADICTED [18]. Its three contradictions are the largest number for any single framework (Table 1), and they are a mark of specificity rather than weakness: this is a theory that can be wrong about the phenotype.

The inflammatory account is the only framework that genuinely straddles a polarity, and it does so by splitting on timescale. Acute cytokine-induced sickness behavior yields anorexia and somnolence; chronic low-grade immunometabolic dysregulation yields the weight-gain and leaden-paralysis profile [27,41,60-61]. This is the one place in the matrix where a single mechanism is used to derive opposite poles rather than one of them, and it is also the mechanism with the strongest symptom-specific empirical support for the reverse-neurovegetative pattern [3,48]. But the inflammatory account too predicts hypoactivity and therefore contradicts agitation. Across all 25 frameworks, the union of biological and psychological coverage still leaves the agitated and hyperphagic presentations without a mechanism from either side.

Discipline Does Not Predict Explanatory Profile

Finally, the author asked whether the field's own taxonomy of theories carries information about what those theories explain. Ward clustering of the 25 coverage vectors recovers disciplinary families barely above chance: ARI = 0.04 for the five-cluster solution. Mean within-family profile distance (3.65) is only slightly below mean between-family distance (4.02), a separation of 0.37 units on a scale where the observed nearest-neighbor distances range from 0.00 to 3.87. And 13 of 25 theories have a nearest neighbor outside their own family a majority of the corpus is more similar, in what it explains, to a theory from a different discipline than to any theory from its own.

The five clusters are functional rather than disciplinary (Table 4, Figure 1).

Cluster	Members	Shared explanatory commitment	Selected nearest neighbors (Euclidean d)
1. Self-referential / meaning	Beck cognitive, Hopelessness, Psychodynamic, Evolutionary	DIRECT coverage of worthlessness/guilt and suicidal ideation	Psychodynamic ↔ Evolutionary, d = 2.65 †; Beck → Hopelessness, d = 2.83
2. Stress-reactivity / threat-sensitivity	HPA axis, Rumination, Self-discrepancy, Interpersonal, Attachment/adversity	Mood + somatic anxiety + rejection sensitivity	HPA axis → Rumination, d = 3.87 †; Hopelessness → Rumination, d = 2.65
3. Vegetative / somatic	Inflammation, Circadian/zeitgeber, Predictive processing, Network	Fatigue, sleep, appetite, psychomotor slowing	Inflammation ↔ Predictive processing, d = 2.83 †; Network ↔ Circadian, d = 3.00 †
4. Motivation–activation	Monoamine deficiency, Reward circuit, Mitochondrial, Learned helplessness, Behavioural activation	Reduced activity, anhedonia, low energy; three contradict or exclude agitation	Learned helplessness → Monoamine, d = 2.45 †; Monoamine → Mitochondrial, d = 2.24 †; Reward circuit ↔ Mitochondrial, d = 1.00; Behavioral activation → Monoamine, d = 3.00 †
5. Low-resolution	Serotonin, Receptor adaptation, Glutamate/plasticity, Neurotrophic, Microbiome, Diathesis–stress/G×E, Kindling	Claims pitched at the syndrome or at treatment response rather than at symptoms	Diathesis–stress ↔ Kindling, d = 0.00; Receptor adaptation → Diathesis–stress, d = 1.00 †

Table 4: Five-cluster solution from Ward linkage on coverage vectors, with each cluster's characteristic symptom commitments and selected nearest-neighbor distances. † = nearest neighbor lies outside the theory's own disciplinary family.

Cluster 1 (self-referential/meaning) groups Beck's cognitive theory, hopelessness theory, psychodynamic object-loss accounts and evolutionary theories, unified by direct coverage of worthlessness/guilt and suicidal ideation the two symptoms hardest to inherit (Section One hub, not three and the secondary-inheritance illusion). Its most informative feature is that psychodynamic and evolutionary accounts are each other's mutual nearest neighbors (d = 2.65), an alliance invisible to any disciplinary taxonomy and separated by eighty years and two incompatible methodologies. Both derive self-devaluation and suicidality from a functional interpretation of loss and social standing: Freud's (1917/1957) turning of reproach against an introjected object and Allen and Badcock's, social risk hypothesis reach the same symptoms from the same explanatory move.

Cluster 2 (stress-reactivity/threat-sensitivity) groups the HPA axis with rumination, self-discrepancy, interpersonal and

Galley Proof

attachment/adversity accounts. The HPA axis's nearest neighbor is response-styles theory (d = 3.87), which is the largest nearest-neighbor distance in the matrix and so the weakest of the cluster alliances; the shared profile is mood plus somatic anxiety plus rejection sensitivity. That a neuroendocrine cascade and a cognitive-style theory are closest to each other reflects a real commonality both are theories of sustained reactivity to social threat that the biological/psychological dichotomy cannot represent.

Cluster 3 (vegetative/somatic) groups inflammation, circadian/social-zeitgeber theory, predictive processing and network theory. Inflammation and allostatic predictive processing are mutual nearest neighbors (d = 2.83), converging on fatigue, sleep, appetite and psychomotor slowing; both treat the depressive phenotype as the output of a regulatory system managing bodily state under constraint, one immunologically and one computationally [25-27]. Network theory's nearest neighbor is circadian theory (d = 3.00), which is an artefact of breadth rather than substance: network theory is close to whatever covers the vegetative symptoms because it covers everything weakly.

Cluster 4 (motivation–activation) is the clearest cross-disciplinary alliance in the matrix. It groups monoamine deficiency, reward-circuit accounts, mitochondrial bioenergetics, learned helplessness and behavioral activation, and four of its five nearest-neighbor links cross family boundaries. Learned helplessness is nearer to monoamine deficiency (d = 2.45) than to any other cognitive theory, and behavioral activation likewise (d = 3.00); monoamine deficiency's own nearest neighbor is the mitochondrial hypothesis (d = 2.24), and reward-circuit and mitochondrial accounts are separated by d = 1.00. All five predict reduced activity, anhedonia and low energy, and three of them contradict or exclude agitation. A 1965 neurochemical hypothesis, a 1967 animal-learning paradigm and a 2001 behavioral treatment manual make nearly the same symptom-level commitments; the taxonomy that separates them into "neurochemical" and "cognitive–behavioral" records their methods, not their content.

Cluster 5 (low-resolution) collects the frameworks whose claims are pitched above the symptom level: serotonin, receptor adaptation, glutamate/plasticity, neurotrophic, microbiome, diathesis–stress/G×E and kindling. Diathesis–stress and kindling are separated by d = 0.00 identical coverage vectors, because both make exactly one non-silent claim (depressed mood) and are silent on the other 15. Receptor adaptation's nearest neighbor is diathesis–stress at d = 1.00. This cluster is not a theoretical alliance but an absence: these frameworks are close together because none of them say much.

The practical upshot is that the familiar biological-versus-psychological axis is nearly uninformative about explanatory content, while cluster membership predicts which symptoms a theory will and will not address including, in cluster 4's case, which symptom it will rule out.

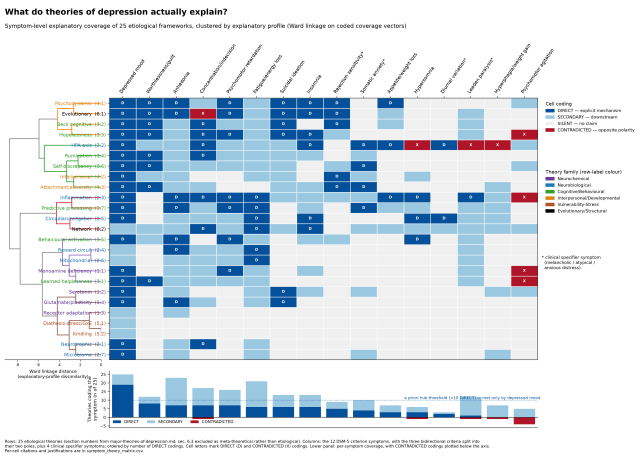


Figure 1: Clustered heatmap of the 25 × 16 weighted coverage matrix (DIRECT = 2, SECONDARY = 1, SILENT = 0, CONTRADICTED = −1), with theories ordered by Ward linkage on Euclidean distance and symptoms grouped by DSM criterion and specifier. The depressed-mood column is the only near-solid band; the psychomotor-agitation and hyperphagia columns are the only columns containing no DIRECT cells. [Source file: symptom_theory_coverage_heatmap.png]

DISCUSSION

Asked what they explain, 25 etiological theories of depression collectively answer: depressed mood, and for a minority with dedicated mechanisms anhedonia, fatigue, self-devaluation and suicidality. Nearly half of the theory-by-symptom space carries no claim. The apparent consensus about anhedonia and fatigue is produced by inheritance rather than by prediction. Four symptoms have no dedicated account, and one of them is actively contradicted by four frameworks. Coverage of the bidirectional criteria is asymmetric rather than joint, and the single framework that spans all three does so by not explaining symptoms. Disciplinary provenance is close to uninformative about any of this.

The author draws three sets of implications: for how theories are appraised, for how the diagnostic criteria are revised, and for how RDoC's construct set might be extended. Each is a proposal about practice, and each has a clinical or translational payoff stated explicitly.

For Theory Appraisal: Report Coverage, and Stop Crediting Inheritance

The first proposal is procedural. Symptom-level coverage should be reported alongside evidential support, as a separate and non-substitutable dimension of theory evaluation. A theory with strong evidence for a narrow claim reward-circuit accounts of anhedonia are the clearest example and a theory with weak evidence for a broad claim are not on the same scale, and the current practice of evaluating both against "depression" conflates them. Two theories can have identical category-level records and completely non-overlapping explananda. Nothing in standard appraisal makes that visible.

Concretely, the author suggests that new etiological proposals be accompanied by a coverage statement naming, for each

Galley Proof

symptom of the target phenotype, whether the theory predicts it, inherits it, is silent about it, or predicts against it. This costs an author very little and changes what reviewers can ask. The question moves from "is there evidence for this mechanism" to "which symptoms does this commit you to, and which do you rule out" the second question being the one that generates falsifiable predictions and the one this audit found the corpus has largely avoided answering.

The audit also identifies a specific appraisal failure that such a statement would make visible: secondary inheritance. A theory acquires apparent coverage of anhedonia or fatigue by describing it as a downstream consequence of the modelled process, without predicting its severity, its time course, its distribution across patients, or what would remove it. The author's position is that inheritance should not count as coverage, for a reason internal to the logic of appraisal rather than a matter of taste: because inheritance generates no symptom-specific prediction, no symptom-level observation can disconfirm it, so crediting it rewards unfalsifiability. The 16 secondary claims on anhedonia and 15 on fatigue (Table 2) are not sixteen and fifteen independent explanatory successes. They are sixteen and fifteen instances of a theory declining to be tested at that grain, and their aggregate effect is to make a thinly modelled symptom look like settled ground.

There is a corollary about contradiction that runs the other way. The HPA-axis account carries three CONTRADICTED codings and the monoamine, inflammatory, learned-helplessness, hopelessness and evolutionary accounts one each (Table 1). Under current appraisal habits a counter-prediction is a liability. Under symptom-level appraisal it is the mark of a theory specific enough to be wrong about the phenotype, and the eight contradicting cells in this matrix are among the most informative in it. Gold and Chrousos's (2002) explicit prediction that melancholic and atypical depression have opposite neuroendocrine signatures is a substantive, testable, subtype-specific claim of exactly the kind the corpus needs more of. A field that penalizes contradiction and rewards inheritance will select for theories that cannot be refuted by any fact about how depression presents.

Translational payoff

Coverage statements bear directly on mechanism-to-treatment reasoning. If a treatment is developed on a theory that models only depressed mood, expecting it to resolve hyperphagia or agitation is an inference the theory does not license and a trial powered on a sum score will not detect the difference between a treatment that lifts mood and leaves the vegetative profile intact and one that resolves the episode. Immunometabolic features predict poorer antidepressant response across pooled trials, which is what one expects if the treatments were developed against a mechanism that models a different part of the phenotype [49]. Symptom-level coverage statements would make such mismatches predictable in advance rather than discoverable in secondary analysis.

For Diagnostic Classification: Bidirectional Criteria Without Etiological Warrant

The orphan set is not a random sample of the criteria. Psychomotor agitation, hyperphagia/weight gain, leaden paralysis and diurnal mood variation are precisely the features that distinguish melancholic from atypical presentations the subtyping with the longest history of claimed treatment-moderating value. That the theoretical corpus is silent on them, and in the case of agitation, actively predicts their opposite, means the DSM's bidirectional criteria are doing classificatory work that no etiological model underwrites.

The historical explanation is straightforward and worth stating because it is not a criticism of the criteria's authors. The DSM criteria were constructed to maximize reliable clinical recognition, and bidirectional items serve that goal well: they capture the range of presentations a clinician encounters and avoid excluding patients whose vegetative symptoms run in the atypical direction. Theory construction then proceeded largely against the *modal* presentation, or against whichever pole the theorist's preferred mechanism could generate. The field has theorized the recognizable half. The result is a classification whose bidirectional structure encodes a genuine clinical observation that depressed patients sleep too little or too much, eat too little or too much, slow down or speed up and a theoretical literature that has an account of one branch of each disjunction and, for two of the three, nothing at all for the other.

The author makes one concrete recommendation for criteria revision, and it is deliberately modest. Any revision of the mood-disorder criteria that retains bidirectional items should state what etiological account, if any, licenses each pole. This is not a proposal to delete poles for which no theory exists clinical reality is not obliged to wait for theory, and the immunometabolic literature shows that a pole can have a well-replicated biological signature before it has a mechanism. It is a proposal that the warrant be made explicit, so that criterion sets carry an honest record of which of their elements are theoretically grounded and which are grounded only in clinical recognizability. Where the answer is "none," that is a research priority statement rather than an argument for deletion.

A second recommendation follows for how subtypes are validated. The melancholic and atypical specifiers are currently justified by their clinical and prognostic correlates, and for the atypical/immunometabolic specifier those correlates are now substantial: greater genetic overlap with BMI and CRP, symptom-specific inflammatory associations, shared genetic predisposition with immunometabolic traits, recovery from data-driven symptom-profile clustering, and differential treatment response [3,15,16,48,49]. This audit shows that the empirical program has outrun the theoretical one: the biology of the reverse-neurovegetative pattern is better characterized than any theory's account of why those symptoms occur together. The immunometabolic depression construct is the field's most developed attempt to close that gap, and it is also the only framework in this matrix that derives opposite poles of a criterion from one mechanism by splitting on timescale (Section Polarity coverage is asymmetric, not bidirectional). It

Galley Proof

is, on the evidence of this audit, the most promising template for what a pole-specific etiological account looks like.

Clinical Payoff

Diurnal mood variation illustrates what the theoretical silence costs at the bedside. It is assessed in routine practice, appears in standard rating instruments, is measurable with experience sampling, and has been studied as a predictor of treatment outcome in large cohorts yet 23 of 25 frameworks offer no account of it, so a clinician who observes it has no mechanism to reason from about what it implies for treatment selection [57-58]. The same holds for agitation, where the absence of a mechanism coexists with a live controversy about antidepressant-associated worsening and with evidence that agitation predicts mood switching [37-38]. Symptoms that are routinely observed, prognostically meaningful, and theoretically unmodelled are where clinical decisions are least supportable.

For Rdoc: The Construct Set Inherits the Asymmetry

RDoC replaced categories with constructs precisely to escape the problems the heterogeneity literature identified [14]. The present audit supports the core methodological bet and identifies a specific incompleteness in its execution.

The support first. RDoC's central wager is that units of analysis should be organized around function rather than around the discipline that studies it. This cluster analysis is, as far as the author knows, the first direct test of that wager at the level of *theoretical content* rather than method, and it comes out in RDoC's favour: explanatory profiles cross disciplinary lines (ARI = 0.04), and 13 of 25 theories are closest to a theory from another family. The alliances that emerge psychodynamic with evolutionary, inflammation with predictive processing, learned helplessness with monoamine deficiency are functional groupings that the discipline-based taxonomy actively obscures. If theories that were built with incompatible methods converge on the same explananda, then organizing research around explananda rather than methods is the right move, and it is vindicated by the content of the corpus and not merely by the inadequacy of DSM categories.

The incompleteness next. The construct set inherits the same asymmetry found in the theoretical literature, which is unsurprising given that it was assembled from that literature. Negative Valence Systems, Positive Valence Systems (reward learning, reward valuation, effort valuation) and Arousal and Regulatory Systems map cleanly onto the hub and onto the vegetative cluster: depressed mood, anhedonia, fatigue and sleep all have construct homes. Nothing in the framework indexes agitated hyperactivity in the context of low mood Arousal and Regulatory Systems cover arousal level but not the conjunction of elevated motor activation with depressed mood and retarded ideation that agitated depression presents. Nor does anything index the switch between appetite poles: the framework has constructs for reward valuation and for homeostatic regulation, but none whose parameter settings determine which patient becomes anorexic and which hyperphagic. Diurnal mood variation has no construct at all; it is a temporal property of mood rather than a level of any construct, and RDoC's matrix has no row for the phase relations that circadian theory specifies.

The audit therefore supplies a constructive target rather than a critique. The orphaned symptoms mark where new constructs, not merely new measures of existing constructs, are required. This distinction matters insofar as to how the gap gets closed. Improved measurements of Arousal and Regulatory Systems will not discover a mechanism for agitated depression, because the construct does not represent the conjunction that defines it. Improved measurement of homeostatic appetite regulation will not explain polarity, because the construct is specified without a polarity parameter. Three candidate additions follow from the current matrix: a construct for motor activation in the context of negative valence, dissociable from general arousal; a construct for allostatic set-point direction, specifying which way homeostatic regulation fails rather than that it fails; and a construct for circadian phase relations to affect, for which circadian/social-zeitgeber theory already supplies the mechanism and the empirical program supplies the measures [44,55,62].

The author also notes a tension worth pursuing. The functional groupings that emerge empirically from the current clustering self-referential, threat-reactivity, vegetative, motivation–activation do not coincide with RDoC's current domain structure. Motivation–activation cuts across Positive Valence Systems and Arousal/Regulatory; self-referential meaning-making is distributed across Social Processes and Negative Valence; threat-reactivity spans Negative Valence and Social Processes. If the theoretical corpus organizes itself into four functional groupings that the domain structure does not represent, that is at least a reason to ask whether the domains were drawn at the right joints.

Translational Payoff

The practical consequence is about where discovery effort is allocated. Construct-level measurement refinement is the dominant activity in the RDoC program, and for the hub and the vegetative cluster it is the right activity. For the orphans it is not: no amount of measurement refinement will discover mechanisms for symptoms no theory currently addresses, because there is nothing to measure more precisely. The orphans require theory construction, and the current matrix identifies the specific explananda a candidate theory would need to derive agitation from a mechanism that does not also predict hypoactivity, hyperphagia from a mechanism with a polarity parameter, leaden paralysis from something more specific than general energy depletion, and diurnal variation from something other than the two accounts that already have it.

What Symptom-Level Coverage Is Not

Three clarifications forestall over-reading, and they are stated because each corresponds to a misinterpretation the results invite.

Coverage is orthogonal to truth: Breadth of symptom-level claim is a property of how a theory was formulated, not evidence that its claims are correct. That psychodynamic object-loss accounts tie with the inflammatory hypothesis for widest

Galley Proof

DIRECT coverage (8 each) says that both were built to explain a described clinical picture; it says nothing about their evidential standing, which differs enormously. A narrow theory that is right is worth more than a broad theory that is wrong. Coverage and warrant are two axes, and the argument of this paper is that the field reports the second and not the first, not that the first should replace it.

Zero DIRECT coverage is not a defect: Diathesis–stress/G×E and kindling take the diagnostic category as their outcome by design: they predict who becomes depressed and when recurrence occurs. Receptor-adaptation accounts predict the latency of treatment response. These are legitimate explanatory projects and the current matrix scores them at zero because it measures something they do not attempt. What the zeros do establish is that these frameworks cannot be confirmed or embarrassed by any observation about the form an episode takes, which is worth knowing when they are cited in discussions of depressive phenotypes.

A low silence rate is not coverage: Network theory has the lowest silence rate in the matrix (1 of 16 cells) and 3 DIRECT codings. This is not a defect either, because network theory is not in the business of deriving symptoms from mechanisms its explanatory claim is about relations among symptoms, and the current rubric, which asks about mechanisms for individual symptoms, is orthogonal to it. The author included it because it belongs in the theory set and because its profile is informative: the only framework that spans all six poles of the bidirectional criteria is the one that treats each pole as a node rather than as an output. That is a genuine result about the corpus's structure, and it is also a reminder that a coverage matrix scores one kind of explanatory ambition and not all of them.

LIMITATIONS

The coding is interpretive, single-rater and unblinded. This is the main limitation, and it bounds every quantity reported here. The appropriate next step is independent multi-rater coding against a published manual with reported inter-rater reliability, which would also permit formal inference on the cluster structure; the ARI, the within/between distance comparison and the nearest-neighbor counts are currently descriptive statistics on a single coding, with no sampling distribution attached. The author has made the full coded matrix with per-cell justifications available specifically so that recoding is possible without repeating the frame construction.

The DIRECT/SECONDARY boundary is the most consequential judgement in the scheme and the one on which independent raters will most often disagree. The current operationalization does the source license a symptom-specific prediction that data could contradict is defensible but not the only defensible one, and a stricter or looser threshold would move the headline percentages. It would not, however, move the qualitative results: the hub/orphan structure depends on gaps of 12 or more codings, and the four orphans have 0, 0, 1 and 2 direct codings against thresholds that no plausible re-reading closes.

Locators are argument- or section-level rather than verified page images. Where a theory has been restated over decades, it was coded against the primary statement and canonical contemporary developments, which means a claim first made in a later paper by a different author may be credited to the framework. This is the right choice for auditing frameworks rather than papers, but it blurs the boundary of what "the theory" claims.

Theory boundaries and the aggregation of variants within rows affect coverage counts. The aggregation choices made here were coverage-inflating by design (Section Theory set and taxonomy construction): rows that aggregate variants receive credit for any variant's claim. A finer-grained row set would lower the direct rates for the aggregated frameworks and would not raise them anywhere.

The symptom space is the DSM's, plus four specifiers. This inherits the criteria's own selection: symptoms that clinicians recognize and that operational criteria encode. Features outside the criteria anger and irritability, anxiety beyond the anxious-distress specifier, somatic pain, cognitive complaints beyond concentration are outside the current frame, and Fried et al.'s (2016) work on non-DSM symptoms suggests some of them behave at least as informatively as criterion symptoms. An audit against a wider phenotype would likely find more silence, not less.

Finally, the theory set is a curated map of principal accounts of unipolar depression, not an exhaustive census. Sociological and structural accounts, cultural-psychiatric frameworks and tripartite affect models are not represented, and their inclusion would change family-level aggregates. The author does not anticipate it would change the orphan set, since none of these traditions makes pole-specific vegetative predictions, but that is a prediction rather than a result.

CONCLUSION

The heterogeneity critique established that depression is not a consistent syndrome. This audit shows that the theoretical literature is not a consistent explanation either, and that its gaps are systematic rather than incidental. Twenty-five etiological frameworks converge on one symptom, inherit two more without modelling them, and leave four the four that carry the melancholic/atypical distinction with no dedicated account at all. Coverage of the bidirectional criteria is asymmetric on every criterion, and the only framework that spans them declines to explain symptoms individually. The taxonomy the field uses to organize these theories records their methods and not their content.

The author recommends three changes in practice. Symptom-level coverage matrices should be published alongside new etiological proposals, with secondary inheritance reported as such rather than counted as explanation. Any revision of the mood-disorder criteria that retains bidirectional items should state what etiological account licenses each pole. And RDoC's construct set should be extended where the orphaned symptoms show it to be silent motor activation under negative valence, allostatic set-point direction, circadian phase relations to affect since no refinement of existing measures will discover mechanisms for symptoms no theory addresses.

Galley Proof

The question to ask of a theory of depression is not whether it explains depression. It is which symptoms it commits to, which it inherits, and which it rules out. On present evidence the corpus has answered the first of those honestly, the second by default, and the third almost not at all.

DECLARATIONS

Author Note. This audit was produced as a structured appraisal of an existing theory map rather than as a systematic review of primary evidence: it inventories what theories *claim*, not what is *true*. Coverage should be read as a property of theoretical formulation and not as a proxy for evidential support.

FUNDING

No funding was received for this work.

DECLARATION OF COMPETING INTEREST

The author declares no competing interests.

ETHICS

No human participants or identifiable patient data were involved; institutional review was not required.

DATA AND MATERIALS AVAILABILITY

The complete coded matrix with per-cell citations and justifications (symptom_theory_matrix.csv), the wide code grid (symptom_theory_matrix_wide.csv), theory-, symptom- and polarity-level summaries (theory_coverage_summary.csv, symptom_coverage_summary.csv, polarity_coverage_summary.csv), the derived statistics (audit_statistics.json), the clustered heatmap (symptom_theory_coverage_heatmap.png) and the fully reproducible coding and analysis script (build_symptom_theory_matrix.py) accompany this manuscript. Every reported quantity can be recomputed from these files; see the accompanying figure and table manifest for the mapping from each exhibit to its source.

REFERENCES

1. Østergaard SD, Jensen SOW, Bech P. The heterogeneity of the depressive syndrome: When numbers get serious. Acta Psychiatr Scand. 2011;124(6):495-496.
2. Fried EI, Nesse RM, Zivin K, Guille C, Sen S. Depression is more than the sum score of its parts: Individual DSM symptoms have different risk factors. Psychol Med. 2014;44(10):2067-2076.
3. Lamers F, Milaneschi Y, de Jonge P, Giltay EJ, Penninx BWJH. Metabolic and inflammatory markers: Associations with individual depressive symptoms. Psychol Med. 2018;48(7):1102-1110.
4. Milaneschi Y, Kappelmann N, Ye Z, Lamers F, Moser S, Jones PB, et al.. Association of inflammation with depression and anxiety: Evidence for symptom-specificity and potential causality from UK Biobank and NESDA cohorts. Mol Psychiatry. 2021;26(12):7393-7402.
5. Fried EI, Nesse RM. Depression is not a consistent syndrome: An investigation of unique symptom patterns in the STARD study. J Affect Disord. 2015;172:96-102.
6. Fried EI, Nesse RM. Depression sum-scores don't add up: Why analyzing specific depression symptoms is essential. BMC Med. 2015;13:72.
7. van Loo HM, Aggen SH, Kendler KS. The structure of the symptoms of major depression: Factor analysis of a lifetime worst episode of depressive symptoms in a large general population sample. J Affect Disord. 2022;307:115-124.
8. Kendler KS. The dappled nature of causes of psychiatric illness: Replacing the organic–functional/hardware–software dichotomy with empirically based pluralism. Mol Psychiatry. 2012;17(4):377-388.
9. Borsboom D, Cramer AOJ. Network analysis: An integrative approach to the structure of psychopathology. Annu Rev Clin Psychol. 2013;9:91-121.
10. Borsboom D. A network theory of mental disorders. World Psychiatry. 2017;16(1):5-13.
11. Cramer AOJ, van Borkulo CD, Giltay EJ, van der Maas HLJ, Kendler KS, Scheffer M, et al.. Major depression as a complex dynamic system. PLoS One. 2016;11(12):e0167490.
12. Fried EI, Epskamp S, Nesse RM, Tuerlinckx F, Borsboom D. What are "good" depression symptoms? Comparing the centrality of DSM and non-DSM symptoms of depression in a network analysis. J Affect Disord. 2016;189:314-320.
13. Robinaugh DJ, Hoekstra RHA, Toner ER, Borsboom D. The network approach to psychopathology: A review of the literature 2008–2018 and an agenda for future research. Psychol Med. 2020;50(3):353-366.
14. Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K, et al.. Research Domain Criteria (RDoC): Toward a new classification framework for research on mental disorders. Am J Psychiatry. 2010;167(7):748-751.
15. Milaneschi Y, Kan C, Pistis G, Strohmaier J, Peyrot WJ, Mullins N, et al.. Genetic association of major depression with atypical features and obesity-related immunometabolic dysregulations. JAMA Psychiatry. 2017;74(12):1214-1225.
16. Badini I, Coleman JRI, Hagenaars SP, Hotopf M, Breen G, Lewis CM, et al.. Depression with atypical neurovegetative symptoms shares genetic predisposition with immuno-metabolic traits and alcohol consumption. Psychol Med. 2022;52(4):726-736.
17. Carroll BJ, Feinberg M, Greden JF, Tarika J, Albala AA, Haskett RF, et al.. A specific laboratory test for the diagnosis of melancholia: Standardization, validation, and clinical utility. Arch Gen Psychiatry. 1981;38(1):15-22.
18. Gold PW, Chrousos GP. Organization of the stress system and its dysregulation in melancholic and atypical depression: High vs low CRH/NE states. Mol Psychiatry. 2002;7(3):254-275.
19. Seligman MEP, Maier SF. Failure to escape traumatic shock. J Exp Psychol. 1967;74(1):1-9.
20. Maier SF, Seligman MEP. Learned helplessness at fifty: Insights from neuroscience. Psychol Rev. 2016;123(4):349-367.
21. Nestler EJ, Carlezon WA Jr. The mesolimbic dopamine reward circuit in depression. Biol Psychiatry. 2006;59(12):1151-1159.
22. Treadway MT, Zald DH. Reconsidering anhedonia in depression: Lessons from translational neuroscience. Neurosci Biobehav Rev. 2011;35(3):537-555.
23. Pizzagalli DA. Depression, stress, and anhedonia: Toward a synthesis and integrated model. Annu Rev Clin Psychol. 2014;10:393-423.
24. Huys QJM, Pizzagalli DA, Bogdan R, Dayan P. Mapping anhedonia onto reinforcement learning: A behavioural meta-analysis. Biol Mood Anxiety Disord. 2013;3:12.

Galley Proof

25. Stephan KE, Manjaly ZM, Mathys CD, Weber LAE, Paliwal S, Gard T, et al.. Allostatic self-efficacy: A metacognitive theory of dyshomeostasis-induced fatigue and depression. Front Hum Neurosci. 2016;10:550.
26. Barrett LF, Quigley KS, Hamilton P. An active inference theory of allostasis and interoception in depression. Philos Trans R Soc Lond B Biol Sci. 2016;371(1708):20160011.
27. Dantzer R, O'Connor JC, Freund GG, Johnson RW, Kelley KW. From inflammation to sickness and depression: When the immune system subjugates the brain. Nat Rev Neurosci. 2008;9(1):46-56.
28. Felger JC, Treadway MT. Inflammation effects on motivation and motor activity: Role of dopamine. Neuropsychopharmacology. 2017;42(1):216-241.
29. Schildkraut JJ. The catecholamine hypothesis of affective disorders: A review of supporting evidence. Am J Psychiatry. 1965;122(5):509-522.
30. Bunney WE Jr, Davis JM. Norepinephrine in depressive reactions: A review. Arch Gen Psychiatry. 1965;13(6):483-494.
31. Abramson LY, Metalsky GI, Alloy LB. Hopelessness depression: A theory-based subtype of depression. Psychol Rev. 1989;96(2):358-372.
32. Åsberg M, Träskman L, Thorén P. 5-HIAA in the cerebrospinal fluid: A biochemical suicide predictor? Archives of General Psychiatry, 33(10), 1193–1197..
33. Coppen A. The biochemistry of affective disorders. British Journal of Psychiatry. 1967;113(504):1237-1264.
34. Higgins ET. Self-discrepancy: A theory relating self and affect. Psychol Rev. 1987;94(3):319-340.
35. Abraham K. Notes on the psycho-analytical investigation and treatment of manic-depressive insanity and allied conditions. In: Selected papers of Karl Abraham. Hogarth Press; 1927. p. 137-156. Original work published 1911.
36. Maj M, Pirozzi R, Magliano L, Fiorillo A, Bartoli L. Agitated "unipolar" major depression: Prevalence, phenomenology, and outcome. J Clin Psychiatry. 2006;67(5):712-719.
37. Iwanami T, Maeshima H, Baba H, Satomura E, Namekawa Y, Shimano T, et al.. Psychomotor agitation in major depressive disorder is a predictive factor of mood-switching. J Affect Disord. 2015;170:185-189.
38. Akiskal HS, Perugi G, Benazzi F, Rihmer Z. Agitated "unipolar" depression re-conceptualized as a depressive mixed state: Implications for the antidepressant-suicide controversy. J Affect Disord. 2005;85(3):245-258.
39. Song L, Lyu Z, Xie J, Zhang Y, Zhao D, Chen A, et al.. Assessing psychomotor disturbances among mental disorders: A systematic review and meta-analysis. Psychother Psychosom. 2026;95(1):28-38.
40. Wurtman RJ, Wurtman JJ. Carbohydrates and depression. Sci Am. 1989;260(1):68-75.
41. Milaneschi Y, Lamers F, Berk M, Penninx BWJH. Depression heterogeneity and its biological underpinnings: Toward immunometabolic depression. Biol Psychiatry. 2020;88(5):369-380.
42. Howren MB, Lamkin DM, Suls J. Associations of depression with C-reactive protein, IL-1, and IL-6: A meta-analysis. Psychosom Med. 2009;71(2):171-186.
43. Rosenthal NE, Sack DA, Gillin JC, Lewy AJ, Goodwin FK, Davenport Y, et al.. Seasonal affective disorder: A description of the syndrome and preliminary findings with light therapy. Arch Gen Psychiatry. 1984;41(1):72-80.
44. Wirz-Justice A. Biological rhythm disturbances in mood disorders. Int Clin Psychopharmacol. 2006;21(Suppl. 1):S11-S15.
45. Zheng P, Zeng B, Zhou C, Liu M, Fang Z, Xu X, et al.. Gut microbiome remodeling induces depressive-like behaviors through a pathway mediated by the host's metabolism. Mol Psychiatry. 2016;21(6):786-796.
46. Nelson J, Klumparendt A, Doebler P, Ehring T. Childhood maltreatment and characteristics of adult depression: Meta-analysis. British Journal of Psychiatry. 2017;210(2):96-104.
47. Łojko D, Rybakowski JK. Atypical depression: Current perspectives. Neuropsychiatric Disease and Treatment. 2017;13:2447-2456.
48. de Kluiver H, Milaneschi Y, Jansen R, van Sprang ED, Giltay EJ, Hartman CA, et al.. Associations between depressive symptom profiles and immunometabolic characteristics in individuals with depression and their siblings. World J Biol Psychiatry. 2021;22(2):128-138.
49. Vreijling SR, Chin Fatt CR, Williams LM, Schatzberg AF, Usherwood T, Nemeroff CB, et al.. Features of immunometabolic depression as predictors of antidepressant treatment outcomes: Pooled analysis of four clinical trials. Br J Psychiatry. 2024;224(3):89-97.
50. Milaneschi Y, Lamers F, Bot M, Drent ML, Penninx BWJH. Leptin dysregulation is specifically associated with major depression with atypical features: Evidence for a mechanism connecting obesity and depression. Biol Psychiatry. 2017;81(9):807-814.
51. Gardner A, Boles RG. Beyond the serotonin hypothesis: Mitochondria, inflammation and neurodegeneration in major depression and affective spectrum disorders. Prog Neuropsychopharmacol Biol Psychiatry. 2011;35(3):730-743.
52. Beck AT. Depression: Clinical, experimental, and theoretical aspects. Harper & Row; 1967.
53. Martell CR, Addis ME, Jacobson NS. Depression in context: Strategies for guided action. W. W. Norton; 2001.
54. Allen NB, Badcock PBT. The social risk hypothesis of depressed mood: Evolutionary, psychosocial, and neurobiological perspectives. Psychol Bull. 2003;129(6):887-913.
55. Lyall LM, Wyse CA, Graham N, Ferguson A, Lyall DM, Cullen B, et al.. Association of disrupted circadian rhythmicity with mood disorders, subjective wellbeing, and cognitive function: A cross-sectional study of 91,105 participants from the UK Biobank. Lancet Psychiatry. 2018;5(6):507-514.
56. Moffoot APR, Carroll S, Bennie J, Dick H, Ebmeier KP, O'Carroll RE, et al.. Diurnal variation of mood and neuropsychological function in major depression with melancholia. J Affect Disord. 1994;32(4):257-269.
57. Morris DW, Trivedi MH, Fava M, Wisniewski SR, Balasubramani GK, Khan AY, et al.. Diurnal mood variation in outpatients with major depressive disorder. Depress Anxiety. 2009;26(9):851-863.
58. Peeters F, Rottenberg J, Delespaul P, Berkhof J, Nicolson NA. Diurnal mood variation in major depressive disorder. Emotion. 2006;6(3):383-391.
59. Wirz-Justice A. Diurnal variation of depressive symptoms. Dialogues Clin Neurosci. 2008;10(3):337-343.
60. Sapolsky RM, Krey LC, McEwen BS. The neuroendocrinology of stress and aging: The glucocorticoid cascade hypothesis. Endocr Rev. 1986;7(3):284-301.
61. Miller AH, Raison CL. The role of inflammation in depression: From evolutionary imperative to modern treatment target. Nat Rev Immunol. 2016;16(1):22-34.
62. Ehlers CL, Frank E, Kupfer DJ. Social zeitgebers and biological rhythms: A unified approach to understanding the etiology of depression. Arch Gen Psychiatry. 1988;45(10):948-952.

Galley Proof