What Mental Health Is: Integration, Teleology, and the Normative Foundations of Psychiatry

By Luke Benton

Contents

Abstract

The classification of mental disorders rests on an unresolved normative foundation. Despite decades of conceptual refinement, psychiatry lacks a principled criterion for distinguishing mental disorders from other forms of psychological variation—a gap exposed by the 1973 removal of homosexuality from the DSM, which turned on shifting cultural consensus rather than new empirical findings. The dominant philosophical solution, Wakefield's harmful dysfunction analysis, attempts to ground disorder in evolutionary dysfunction constrained by culturally mediated harm judgments, but this hybrid approach fails because the normative component remains tethered to the very cultural values the framework was designed to discipline.

This paper argues that the normativity required for a scientifically robust definition of mental health can be objectively grounded in the teleological structure of living systems. Drawing on neo-Aristotelian naturalism, it contends that human flourishing is constitutively relational, such that individual well-being depends on the health of the larger wholes—families, communities, ecosystems—in which individual humans are embedded. Mental health is defined as systemic alignment: the reliable tendency of the decision-making system to produce thoughts, emotions, and behaviours that sustain and generate integrative wholes. Mental illness, correspondingly, is repulsive processing: a pattern that systematically degrades the cooperative connections upon which flourishing depends, often maintained through emotional capture loops in which distressing affect and distorted cognition reinforce one another.

The framework generates specific, falsifiable predictions operationalized through the Integrative Processing Orientation (IPO) construct, a proposed disposition measurable via a four-factor scale. It provides a unifying theoretical rationale for the mechanisms underlying cognitive-behavioural, acceptance-based, psychodynamic, and mindfulness interventions, and it yields determinate, non-arbitrary guidance in contested cases, as demonstrated through a detailed analysis of gender dysphoria in a non-affirming environment. The paper thus offers a resolution to the normative crisis in mental health that is simultaneously empirically responsible, philosophically rigorous, and clinically orienting.

Chapter 1: The Normative Crisis in Mental Health

1.1. The Real-World Problem

In 1973, the American Psychiatric Association voted to remove homosexuality from the Diagnostic and Statistical Manual of Mental Disorders. The decision was not prompted by new scientific discoveries about same-sex attraction but by shifting cultural norms and political pressure. Whatever one thinks of the outcome, the episode exposes a conceptual fracture at the foundation of mental health practice: the field lacks a principled, scientifically grounded way to distinguish mental disorders from other forms of psychological variation. The definition of mental disorder is the load-bearing wall of the entire clinical edifice—it determines who receives treatment, what counts as a diagnosis, how insurance reimbursements are structured, and where the boundary lies between criminal responsibility and exculpatory pathology. Yet that wall rests on sand.

The problem is not merely academic. Consider the successive editions of the DSM. With each revision, diagnostic categories expand, contract, appear, and vanish. Premenstrual Dysphoric Disorder, Binge Eating Disorder, and Internet Gaming Disorder entered the official nosology not because underlying pathological mechanisms were discovered but because committees of clinicians reached consensus that certain patterns of distress and impairment warranted medical attention. Hoarding was once a symptom of Obsessive-Compulsive Personality Disorder; it is now a standalone diagnosis. Asperger's Syndrome was introduced in DSM-IV and absorbed into Autism Spectrum Disorder in DSM-5. Narcissistic Personality Disorder nearly disappeared from DSM-5 only to be retained after professional outcry. These are not merely taxonomic refinements; they are substantive decisions about which forms of human suffering and strangeness count as medical conditions, and they are made through processes that blend empirical evidence, clinical experience, and cultural negotiation in proportions no one can quite specify.

The stakes are immense. Overdiagnosis exposes individuals to unnecessary stigma and medication. Underdiagnosis denies treatment to those who suffer. Misdiagnosis sends people down clinical pathways that compound their difficulties. And the broader culture absorbs psychiatric categories into its self-understanding, such that the language of diagnosis—trauma, depression, anxiety, ADHD—becomes the vernacular through which ordinary people interpret their inner lives. When the definitions governing this system lack conceptual coherence, the entire apparatus operates without a clear mandate.

The root of the problem is straightforward to state and extraordinarily difficult to resolve: What is a mental disorder? What makes Major Depressive Disorder a disorder and grief a normal human experience? What makes Generalized Anxiety Disorder a condition requiring treatment and the anxious temperament of a successful professional merely a personality trait? What makes a psychotic delusion pathological but a religious vision culturally valorised? The field has attempted to answer these questions for decades, and while progress has been made, no consensus definition exists that withstands philosophical scrutiny while remaining clinically useful.

1.2. The DSM Definition and Its Discontents

The DSM-5 offers the field's most influential definition. A mental disorder is:

“A syndrome characterized by clinically significant disturbance in an individual's cognition, emotion regulation, or behaviour that reflects a dysfunction in the psychological, biological, or developmental processes underlying mental functioning. Mental disorders are usually associated with significant distress or disability in social, occupational, or other important activities. An expectable or culturally approved response to a common stressor or loss, such as the death of a loved one, is not a mental disorder. Socially deviant behaviour (e.g., political, religious, or sexual) and conflicts that are primarily between the individual and society are not mental disorders unless the deviance or conflict results from a dysfunction in the individual, as described above.”

This definition attempts to exclude several categories: culturally approved responses, socially deviant behaviour that is not caused by individual dysfunction, and expectable responses to stressors. Yet each exclusionary clause generates its own difficulties.

The cultural exclusion—"an expectable or culturally approved response"—presupposes that we can reliably distinguish what a culture "approves" from what it merely tolerates, and that cultural approval is itself a valid criterion for demarcating health from pathology. But cultures approve of many things that seem injurious—foot binding, honour killings, ritual scarification—and disapprove of many things that seem benign. More troublingly, the clause suggests that the same psychological phenomenon could be a disorder in one culture and not in another, which raises the question of whether mental disorders are real entities or cultural constructions. If depression is a disorder in New York but a spiritual trial in rural Nepal, in what sense is it a ‘medical’ condition at all?

The social deviance exclusion attempts to prevent psychiatry from becoming an instrument of social control, but it contains a fatal loophole: deviance is not a disorder unless it results from a dysfunction in the individual. This merely relocates the problem. What counts as a dysfunction? If we define dysfunction in terms of statistical deviance from normal functioning, we risk pathologizing eccentricity and genius. If we define it in terms of distress or impairment, we risk excluding conditions marked by anosognosia—the inability to recognize one's own illness—which characterizes many severe mental disorders. If we define it biologically, we confront the fact that few mental disorders have established biological markers.

The distress-or-disability criterion faces its own difficulties. Many conditions universally recognized as mental disorders do not reliably cause distress. Manic episodes are often experienced as intensely pleasurable. Antisocial Personality Disorder is characterized by a lack of remorse. Anorexia Nervosa involves a denial of the seriousness of low body weight that can border on the delusional. Conversely, many conditions that cause intense distress—poverty, bereavement, romantic rejection—are not considered mental disorders. Distress is neither necessary nor sufficient.

Disability fares no better. Many people with diagnosable mental disorders function well occupationally and socially—the high-functioning depressive, the anxious professional whose hypervigilance drives career success. Conversely, many people without mental disorders experience significant disability—the physically disabled, the intellectually disadvantaged, the socially marginalized. If the criterion is "clinically

significant" impairment, we simply push the question back: significant by what standard? Compared to what baseline?

1.3. Wakefield's Harmful Dysfunction Analysis

The most philosophically sophisticated attempt to solve these problems is Jerome Wakefield's "harmful dysfunction analysis" (HDA), first proposed in 1992 and refined over subsequent decades. Wakefield argues that the concept of mental disorder is a hybrid of two components: a factual component (dysfunction) and a value component (harm). A condition counts as a disorder when (a) some internal mechanism fails to perform its naturally selected function, and (b) this failure causes harm to the individual as judged by cultural standards.

The dysfunction component is grounded in evolutionary biology. Wakefield draws on the etiological account of function, according to which a trait's function is the effect for which it was naturally selected. The heart's function is to pump blood because hearts were selected for blood-pumping, not for making thumping noises. A heart that fails to pump blood adequately is dysfunctional regardless of whether anyone judges this to be bad. This provides an objective, scientifically grounded criterion for the "dysfunction" part of the definition. It is not a matter of cultural opinion whether the heart is malfunctioning; it is a matter of biological fact.

The harm component is grounded in social values. A dysfunction does not constitute a disorder unless it causes harm—distress, disability, or increased risk of death or suffering—as evaluated by cultural standards. This explains why some biological dysfunctions are not disorders. A genetic mutation that prevents wisdom teeth from developing is a dysfunction in Wakefield's sense (the mechanism for producing third molars fails to perform its naturally selected function), but it is not a disorder because it causes no harm and arguably confers a benefit. Conversely, conditions that cause harm but involve no dysfunction—illiteracy, poverty, social oppression—are not disorders. The HDA thus promises to carve nature at its joints while accommodating the role of value judgments in medical classification.

The HDA has been enormously influential. It provides the theoretical framework underlying much of the DSM's conceptual architecture, and it has shaped how clinicians and researchers think about the boundary between disorder and non-disorder. Yet it faces serious objections that ultimately render it inadequate as a solution to the normative crisis.

1.4. The Limits of the Harmful Dysfunction Analysis

The HDA's difficulties fall into several categories: problems with the evolutionary account of function, problems with the harm criterion, problems applying the framework to specific mental conditions, and deeper conceptual problems that suggest the entire approach may be insufficient.

Evolutionary Problems - The dysfunction component requires identifying the naturally selected function of the relevant psychological mechanism. For many physical disorders, this is straightforward—we know what hearts and kidneys were selected for. But for most mental disorders, we do not know what mechanisms are involved, let alone their evolutionary functions. What is the naturally selected function of the mechanism that fails in Major Depressive Disorder? In Schizophrenia? In Generalized Anxiety Disorder? The brain is organized in ways we only dimly understand, and its psychological functions—mood regulation, reality testing, impulse control—are not localized in discrete anatomical structures with transparent evolutionary histories. Wakefield's framework requires a level of mechanistic knowledge that psychiatry simply does not possess.

Moreover, the evolutionary account of function is itself contested among philosophers of biology. Critics argue that many biological traits have effects that were not specifically selected for but that nevertheless contribute to the organism's current fitness. The etiological account excludes these as "functions" in the relevant sense, but this seems arbitrary and potentially distorting. If a psychological mechanism that was originally selected for one purpose has been co-opted for another, is dysfunction relative to the original function or the current one? Evolution does not furnish unambiguous answers to such questions.

Selection vs. Function Mismatch. Even when we can identify an evolutionary function, the relationship between that function and current adaptive value is complex. The human capacity for anxiety was presumably selected for threat detection in ancestral environments. Does this mean that anxiety disorders involve a dysfunction of the threat-detection mechanism, or might they involve a properly functioning mechanism responding to environments for which it was not designed? Wakefield's framework must distinguish genuine dysfunctions from "adaptive responses to maladaptive environments," but this distinction often collapses under scrutiny. Is the increased anxiety reported by residents of violent neighbourhoods a dysfunction of their threat-detection system, or an adaptive calibration to genuine danger? Different answers yield different judgments about who is mentally ill, and the evolutionary facts alone cannot settle the matter.

Harm Problems. The harm criterion is supposed to explain why some dysfunctions are not disorders. But harm is an irreducibly evaluative concept, and it introduces all the cultural variability the HDA was meant to constrain. What counts as harmful varies across cultures, historical periods, and even individuals. The person with a manic episode may not experience their state as harmful; society may judge otherwise. The person with a personality disorder may harm others more than themselves; does the harm criterion count third-party harm? If so, by what standard? The HDA relies on value judgments to distinguish disorders from non-disordered dysfunctions, but it provides no account of how those judgments should be made or by whom.

Furthermore, harm and dysfunction are not independent criteria in practice. Our judgments about whether a dysfunction is harmful are shaped by our prior judgments about whether the condition is a disorder. This introduces a circularity that undermines the HDA's claim to provide a non-arbitrary foundation. The reason wisdom tooth agenesis is not considered harmful is partly because we don't consider it a medical disorder; if we did, we would likely find reasons to consider it harmful

(malocclusion, reduced chewing efficiency). The harm criterion does not independently constrain the dysfunction criterion; it is shaped by the same cultural forces that the HDA was meant to discipline.

Application Problems. The HDA struggles with specific conditions that psychiatry routinely treats. What is the dysfunctional mechanism in Antisocial Personality Disorder? One might point to deficits in empathy or conscience, but what is the evolutionary function of these capacities? If they function to facilitate cooperation and group living, then a person who successfully exploits others without remorse may have a 'different' functioning mechanism, not a dysfunction—a predatory strategy that works perfectly well for the individual, however harmful to others. To classify this as a disorder, we must privilege certain functions (group-oriented ones) over others (individual-exploitative ones), and this is a value choice, not a biological fact.

Similarly, what is the dysfunction in Paraphilic Disorders? Wakefield has argued that the relevant mechanism is the sexual arousal system, whose function is to direct sexual interest toward reproductively viable targets. But this requires an account of "natural" sexual function that is difficult to sustain given what we know about the diversity of sexual behaviour in nature and human history. The classification of paraphilias as disorders depends heavily on cultural judgments about acceptable sexuality, and the evolutionary veneer appears thin.

The Deeper Problem. The most fundamental difficulty with the HDA is not any specific technical flaw but its failure to address the underlying normative crisis. The HDA is an attempt to marry objective biological facts with subjective cultural values and produce a coherent hybrid. But the marriage is unstable because the partners speak different languages. The dysfunction component speaks the language of biological description—what mechanisms 'do', what they were selected 'for'. The harm component speaks the language of normative evaluation—what we 'care about', what we judge to be 'bad'. The HDA places these two languages side by side without explaining how they are related or why the combination should carry normative force beyond what either component provides alone.

This reflects a deeper limitation of the entire approach that the HDA exemplifies. The dominant strategy in the philosophy of psychiatry has been to treat mental disorder as a boundary problem: how do we draw the line between healthy and pathological, normal and abnormal, problems in living and genuine medical conditions? The implicit assumption is that "health" is the unproblematic default state—what remains when dysfunction and harm are absent—and that the task of definition is simply to specify the criteria that mark the departure from this default. But this assumption is itself question-begging. It presumes that we already understand what mental health is, that it requires no positive characterization, that it is merely the absence of pathology rather than a distinct state with its own identifiable features. This is analogous to defining physical health as the absence of disease and then defining disease as whatever deviates from normal functioning, without ever specifying what "normal functioning" consists in or why it constitutes the standard of evaluation. The entire enterprise rests on an unexamined normative foundation.

The HDA acknowledges this foundation implicitly through its harm criterion—harm is what we care about avoiding—but it provides no account of what we ought to care about or why. The values that

determine harm are simply taken as given, as though the fact that a culture judges something harmful settles the matter. This is precisely the kind of normative abdication that renders the DSM vulnerable to the cultural winds that produced the 1973 homosexuality vote. If our definition of disorder depends on culturally variable judgments of harm, and if those judgments can shift with political pressure, then the definition of mental disorder is ultimately political, not scientific. The HDA dresses this political reality in evolutionary language but does not fundamentally alter it.

1.5. Beyond the Harmful Dysfunction Model: Why a New Approach Is Needed

The problems with the HDA suggest that the field needs not merely a better set of diagnostic criteria but a different kind of answer to the question "What is a mental disorder?" The HDA attempts to solve the normative crisis by combining an objective component (dysfunction) with a subjective component (harm) and hoping that the combination yields a principled boundary. But this strategy fails because the normative dimension of mental disorder is not adequately captured by cultural judgments of harm, and the objective dimension is not adequately captured by evolutionary dysfunction.

What is needed is an account that grounds the normativity of mental health in something more stable than cultural consensus but more substantive than evolutionary function—an account that specifies what mental health positively consists in, not merely what it is the absence of. Such an account would need to explain why certain psychological states count as healthy and others as pathological in a way that is responsive to empirical evidence, conceptually coherent, and capable of guiding clinical practice without collapsing into cultural relativism or scientific reductionism.

This is a tall order. It requires engaging with questions that modern psychiatry has largely bracketed: What is the nature of human psychological functioning? What constitutes flourishing for creatures like us? Is there an objective standard by which mental health can be evaluated, or are we stuck with the shifting sands of cultural consensus? These are philosophical questions, but they are also practical ones, because the answers we give will determine who gets diagnosed, who gets treated, and who gets stigmatized.

The chapters that follow attempt to provide such an account. I will argue that the normativity required for a scientifically robust definition of mental health can be objectively grounded in the teleological structure of the cosmos, where "healthy" mental function is defined as the cognitive and behavioural alignment with the universal, systemic principle of generating and maintaining integrative wholes. This is a bold claim, and it requires careful defence. Chapter 2 establishes the descriptive foundation: the observation that reality, from quantum physics to human civilization, exhibits a recurrent pattern in which individuals enter into cooperative relationships to form more encompassing systems. This pattern is offered as a descriptive "is"—a factual generalization about how organized complexity arises and persists—not as a normative claim. Chapter 3 builds the teleological bridge from this descriptive pattern to evaluative conclusions, arguing that living systems have objective functional ends that ground judgments of health and dysfunction, and that the human good consists in promoting the health of the larger systems—family, community, biosphere—in which human life is embedded. Chapter 4 applies this

framework to mental health, defining it as systemic alignment: the condition in which an individual's decision-making system reliably produces actions, thoughts, and emotions that sustain and generate integrative wholes. Chapter 5 derives testable predictions from this theory and demonstrates its capacity to unify existing therapeutic approaches. Chapter 6 addresses objections spanning the entire framework.

The ambition is to provide what the field currently lacks: a conception of mental health that is simultaneously scientifically grounded, philosophically rigorous, and practically orienting. Whether this ambition can be realized is for the argument to demonstrate. But the need for such a conception is clear. The normative crisis in mental health is not a parlour game for philosophers; it is a practical problem that affects millions of lives. The DSM's definition of mental disorder governs a global apparatus of diagnosis, treatment, and reimbursement. If that definition rests on an unexamined normative foundation, then the entire apparatus is operating without a clear mandate. The chapters that follow aim to provide that foundation—or, at minimum, to show what kind of foundation is required.

Chapter 2: The Architecture of Complexity: Cooperative Integration as a Recurrent Structural Pattern

A striking observation emerges from the sciences when one surveys the architecture of organized complexity across scales: reality exhibits a recurrent pattern in which individuals enter into cooperative relationships to form more encompassing systems, which themselves become individuals at higher levels of organization. This pattern is not posited as a single causal law operating uniformly across all domains. The mechanisms that produce it are diverse—strong force binding at the subatomic scale, electromagnetic attraction at the chemical scale, natural selection and self-organization at the biological scale, shared intentionality and cultural evolution at the human scale. Yet the structural outcome recurs with sufficient regularity to constitute a robust inductive generalization about how complexity arises and persists. The purpose of this chapter is to articulate this pattern clearly, to demonstrate its empirical grounding across multiple scientific domains, and to mark the boundary that separates the description of how organized systems are structured from any normative claim about how humans ought to behave. In doing so, it sets a foundation of understanding that later chapters can employ to explore prescriptive frameworks—but the pattern itself is offered here solely as a feature of the observable universe, not as a commandment derived from it.

The chapter proceeds as follows. Section 2.1 briefly acknowledges the pattern's appearance across the physical and chemical domains, establishing that cooperative integration is a possibility condition for organized complexity at any scale—a fact that will prove relevant when we later consider why repulsive processing is self-undermining. Section 2.2 examines the biological extension, where the pattern acquires the functional organization essential for teleological evaluation. Section 2.3 traces the pattern into social systems and the human scale, where it becomes directly relevant to mental health. Section 2.4 characterizes the pattern explicitly as a recurrent structural isomorphism rather than a monolithic law, and Section 2.5 distinguishes description from prescription with the necessary clarity.

2.1. The Pattern Across Scales: A Brief Survey

The pattern of cooperative integration is visible even at the most fundamental levels of physical organization, though it operates through mechanisms entirely indifferent to purpose or intention. Quarks bind via the strong nuclear force into protons and neutrons—entities whose existence consists in nothing other than the ongoing, mutually constraining relationships of their constituents. Protons and neutrons enter into further cooperative arrangements to form atomic nuclei, which subsequently attract electrons to create atoms with emergent properties irreducible to any single component. Chemistry extends this logic: atoms form cooperative bonds to produce molecules whose collective behaviours—polarity, solvent capacity, reactivity—are not present in the isolated elements. Lipids self-assemble into membranes, amino acids polymerize into proteins that fold into functional shapes, and nucleotides organize into information-storing polymers. At each stage, the whole is constituted by stable relationships among individual components, and the persistence of the whole depends on those relationships remaining within certain bounds.

What these physical instances reveal is not a normative template for human conduct—electrons do not obligate humans—but a structural condition: wherever organized complexity exists and persists across time, it does so because the relationships among its constituent parts are sufficiently cooperative to maintain the system's integrity. If the cooperative dynamics fall below a critical threshold, the system ceases to exist as that particular organized entity. This condition holds from atomic nuclei to human institutions: cooperative integration is a possibility condition for the emergence and persistence of organized complexity at any scale. A system whose internal dynamics become predominantly repulsive—in which parts degrade rather than sustain the relationships that constitute the whole—is a system in the process of dissolution. This is not a moral claim but a factual observation about the conditions under which organization is possible. Its relevance to the argument of later chapters is this: if human mental health consists in the proper functioning of a processing system embedded within larger relational wholes, then repulsive processing—the systematic degradation of cooperative connections—is not merely unfortunate but structurally self-undermining. It violates the very conditions that make organized psychological and social life possible. The physical survey thus provides not normative premises but a reminder that the framework's later claims about mental illness as self-defeating are instances of a more general structural truth.

2.2. The Biological Extension: Life as Relational Process

The transition from non-living to living systems does not break the pattern of cooperative integration; it extends it into a new regime of complexity. What distinguishes living entities from simpler organized structures is the degree of adaptive coordination among their components, the presence of internal constraints that must be satisfied for the system to persist, and the capacity to process information in ways that guide activity toward constraint satisfaction. Yet the underlying structural principle remains one of individuals forming relationships that give rise to greater functional wholes. It is here, in the biological domain, that the pattern begins to acquire the normative significance that later chapters will develop.

The single-celled organism is already a cooperative enterprise. The modern eukaryotic cell, with its nucleus, mitochondria, and intricate internal membranes, bears the signature of ancient symbiotic mergers—relationships between once-independent prokaryotes that became obligate, each partner eventually specializing in metabolic or reproductive functions while the combined system acquired novel capacities. This is organizational cooperation, not sentiment, and its outcome is a new individual at the cellular scale. In multicellular organisms, individual cells cede a degree of autonomy to specialize and coordinate. Neurons cannot survive in isolation; they exist to relate. Their electrical and chemical conversations, mediated by neurotransmitters and gap junctions, produce networks whose emergent property is rapid information processing, enabling the organism to sense and navigate the world. Muscle cells contract in synchrony, epithelial cells form protective barriers, immune cells patrol and communicate—the body is a society of cells whose cooperative interactions constitute the life of the whole.

Reproduction and development recapitulate the same theme. A fertilized zygote is a single cell, but through division, differentiation, and morphogenetic signalling, it self-assembles into a structured embryo. The process is not directed by an external blueprint but by local interactions: cells respond to chemical gradients, mechanical stresses, and contact with neighbours, and their cumulative decisions produce tissues, organs, and the body plan. The organism emerges from a cascade of relationship-based construction, with each developmental step depending on the coordinated activity of the cells that preceded it.

Crucially, the pattern at this level is not merely structural but functional. The relationships among parts are not only stable; they serve ends. The heart's relationship to the circulatory system maintains blood flow. The immune system's relationship to pathogens maintains bodily integrity. The nervous system's relationship to the environment maintains adaptive responsiveness. These functional relationships provide the basis for teleological evaluation—for judgments about proper and defective functioning—that will be developed in Chapter 3. A heart that fails to pump blood adequately is not merely differently organized; it is dysfunctional relative to the cooperative system it serves. This is not a cultural projection but an objective fact about the role the heart plays in maintaining the multicellular whole. And what makes that dysfunction consequential is precisely the structural truth articulated in Section 2.1: when a component's operations become destructive to the system that sustains it, the system's persistence is threatened, and the component's own conditions of existence erode.

Thus, life, in its very ontology, is a manifestation of the pattern of cooperative integration. Wherever organisms exist, they exist as integrated systems of cooperative parts. The deeper one looks, the more evident it becomes that the "individual" organism is itself a collective achievement—a nested hierarchy of cooperative relationships extending from organelles to organ systems. This biological instance of the pattern will bear genuine normative weight, for it is here that talk of function, health, and flourishing first becomes intelligible.

2.3. Social Systems and the Human Scale

The same relational logic scales further when organisms interact with one another. Social behaviour, from bacterial quorum sensing to the intricate colonies of ants and bees, exhibits individuals coordinating actions in ways that create supra-organismic systems. An ant colony operates with a division of labour, a collective memory in the form of pheromonal trails, and a resilience that no single ant could achieve. The colony is a system whose identity is defined by the relationships among its members, not merely by their physical proximity.

In human beings, the pattern assumes its most flexible and abstract expressions. Language itself is a cooperative system—a shared set of symbols and rules that allows individuals to coordinate thought and action across time and space. A family is a relational system that sustains and socializes its members, with bonds maintained through shared practices, emotional attunement, and mutual care. A community coordinates resource distribution, defence, and cultural preservation through norms, laws, and shared narratives. A scientific enterprise links investigators across centuries in a cooperative search for reliable

knowledge, mediated by publications, peer review, and shared methodological standards. At every scale, human individuals enter into relationships—economic, legal, emotional, communicative—that generate larger wholes with properties irreducible to the individuals who constitute them.

Even the internal self, the seat of personal identity, develops through relational experience with caregivers, peers, and the broader culture. The mind, considered as an ongoing process of self-organization, is itself a product of cooperative interplay between neural systems within the brain and the linguistic and social environments in which the brain is embedded. The boundary between self and world proves porous, for the self is constituted through the very relationships that the pattern describes.

It is essential to note that the presence of this pattern does not deny conflict, competition, or disintegration. Organisms prey upon one another; cells become cancerous; societies fracture into civil war. The fact of destructive and entropic processes is as much a part of the descriptive landscape as the constructive ones. However, the persistence of any complex system over time indicates that, on balance, the relationships among its parts are more cooperative than destructive. If the cooperative component fell below a critical threshold, the system would cease to exist as that particular organized entity. The pattern described is not a law that prohibits dissolution but a condition that must be satisfied wherever enduring complexity is found—a condition whose violation produces, with structural necessity, the degradation of the system in question. Destructive dynamics are real, but they are not what build; they are what systems must resist or compensate for to persist.

2.4. The Pattern as a Recurrent Structural Isomorphism

From this survey of physical, biological, and social domains, a concise characterization becomes defensible: reality exhibits a nested hierarchy of systems, wherein individuals at one level enter into stable, relationship-based interactions to constitute new individuals at a higher level, with novel properties emerging that were not present in the parts alone. This is not a single causal law but a recurrent structural isomorphism—a pattern that recurs across domains for different domain-specific reasons. I offer it as a unifying heuristic principle: a lens through which diverse phenomena can be integrated and their common architecture recognized, without claiming that the integration is driven by a single underlying force.

Several features of this characterization warrant emphasis.

First, the pattern is grounded in empirical observation. Physicists and chemists directly detect the bonding relationships that produce atoms, molecules, and materials. Biologists detail the cooperative interactions among cells that underpin physiology. Social scientists map the networks of communication, exchange, and authority that define organizations. The evidence for each instance is domain-specific and reproducible; it does not depend on a particular cultural or philosophical predisposition. One can observe that cells coordinate to form organisms regardless of one's metaphysical commitments.

Second, the mechanisms that produce the pattern vary by level. The strong force and electromagnetic attraction produce binding at the subatomic and chemical scales. Natural selection, self-organization, and energetic efficiency produce functional integration at the biological scale. Shared intentionality, cultural evolution, and institutional design produce cooperation at the human social scale. The pattern is a regularity of outcomes, a common architecture that recurs because certain kinds of relational arrangements happen to be stable under the relevant dynamical rules at each level, not because a single mechanism operates uniformly. This distinguishes the present framework from reductionist programs that would explain all organization in terms of physics, as well as from vitalist or idealist programs that would posit a special organizing force.

Third, the pattern is robust but not universal in the sense of inevitably dominating every local region. Coherent systems form where conditions permit; they also dissolve. The pattern identifies what exists when organization succeeds, leaving open the question of when and why it fails. It is therefore a factual description of success cases—the structures that populate the cosmos—rather than a deterministic prediction that all parts must cooperate or that complexity will always increase. Entropy, competition, and dissolution are equally real aspects of the world. The pattern simply records that wherever organized complexity endures, cooperative integration is present.

Fourth, the normative significance of the pattern varies by level. The physical instances—quarks to atoms to molecules—establish that cooperative integration is a possibility condition for organized complexity, but they are normatively inert in themselves. The biological instances introduce functional organization, and with it the possibility of teleological evaluation: a heart can function well or poorly relative to the system it serves. The social instances introduce reflective choice, and with it the possibility of moral evaluation: a community can be structured in ways that promote or undermine the flourishing of its members. The normative argument of Chapter 3 will build primarily on the biological and social instances, where functional and evaluative judgments are independently defensible. The structural truth that repulsive dynamics undermine the persistence of organized systems—visible from physics to social organization—will provide context for understanding why repulsive processing in the mental health domain is self-defeating, but it will not itself serve as a normative premise.

2.5. The Boundary Between Description and Prescription

A careful treatment of this pattern requires distinguishing the descriptive "is" from any prescriptive "ought" that might be derived from it. The fact that atoms form molecules does not logically entail that humans should form families or that societies should prioritize collective welfare. To move from "organized complexity arises through cooperative integration" to "one ought to cooperate" is to commit a well-recognized philosophical mistake—the naturalistic fallacy—unless one inserts additional evaluative premises that do not inhere in the observation itself.

Yet the separation is not an argument that the pattern is irrelevant to human concerns. A map of how stable systems actually subsist can inform choices, just as understanding the physiology of the body can inform practices of health. One might, for instance, choose to align one's decisions with the conditions

that have demonstrably enabled complexity and flourishing elsewhere in nature, but that choice is a separate, values-driven act. The present chapter is concerned solely with establishing the map, not with specifying how a person ought to walk upon it.

The normative work of this larger project will depend not on a direct inference from cosmic pattern to human obligation but on a teleological argument of the following form: (1) living systems possess objective functional ends that ground judgments of health and dysfunction; (2) the human life-form is constitutively relational, such that individual flourishing depends on the health of the larger systems in which individual humans are embedded; and (3) therefore, mental health consists in the alignment of individual processing with the conditions that sustain those larger systems. This argument, developed in Chapter 3, draws on the biological and social instances of the pattern described here but does not derive its normative force from the pattern itself. The pattern provides the empirical context; the normative force comes from the nature of living systems and the human life-form.

What the pattern does provide—and what the brief physical survey of Section 2.1 makes visible in its most abstract form—is a structural truth of considerable importance for understanding why certain kinds of dysfunction are self-defeating. If cooperative integration is a possibility condition for organized complexity at any scale, then repulsive processing—the systematic degradation of cooperative connections that Chapter 4 will define as mental illness—is not merely normatively undesirable but structurally unsustainable. It is the psychological equivalent of a system whose internal dynamics have fallen below the cooperative threshold required for its own persistence. This is not a moral judgment dressed in scientific language but a factual observation about the conditions under which organized psychological and social life is possible. The moral weight, when it comes, will arrive through the teleological argument of Chapter 3, not through the structural observation of Chapter 2.

2.6. Conclusion

From the binding of quarks to the architecture of human institutions, existence displays a remarkable structural pattern: individuals form cooperative relationships that create more complex systems, which in turn become individuals in higher-level systems. This is not a poetic metaphor imposed on nature by hopeful imagination but a recurrent structural isomorphism emerging from the empirical findings of physics, chemistry, biology, and the social sciences. It describes the architecture of a universe that builds complexity relationally—not through a single mechanism operating across all scales, but through diverse mechanisms that converge on a common structural outcome.

The value of this characterization lies in its capacity to unify disparate areas of knowledge under a single, empirically grounded perspective without reducing them to a single explanatory framework. It provides a lens through which one can view any organized phenomenon—a galaxy, a cell, a legal system—and ask the same clarifying questions: What are the individuals? What relationships bind them? What greater whole do they constitute? By answering these questions, one maps the architecture of organized complexity without yet pronouncing on its moral meaning.

That mapping is the aim of this chapter, and it stands as an open invitation: if the observable universe displays a recurrent way of building—if cooperative integration is the condition under which organized complexity arises and persists—then understanding that architecture is a prerequisite for any thoughtful discussion of how human beings might build well. The subsequent chapter will take up that discussion, arguing that the normativity required for a scientifically robust definition of mental health can be grounded not in a direct inference from cosmic pattern to human obligation but in the teleological structure of living systems, for whom the pattern of cooperative integration is not merely an observable regularity but a condition of flourishing that their own nature requires them to serve.

Chapter 3: From "Is" to "Good": The Teleological Bridge

The preceding chapter articulated a descriptive pattern observable across the cosmos: individuals enter into cooperative relationships to form more encompassing systems, a pattern visible from quantum binding to human institutions. That pattern was offered strictly as an "is"—a factual generalization about how organized complexity arises and persists. Yet if the larger work is to address how humans might live and what mental health consists in, we must confront a formidable philosophical obstacle. David Hume famously observed that moralists routinely slide from statements about how things "are" to conclusions about how we "ought" to behave, without acknowledging the logical chasm between these two kinds of claim. This chapter argues that the chasm can be bridged—not by deriving an "ought" deductively from a bare "is," but by recognizing that living systems possess objective functional ends, or teloi, which ground evaluative judgments internal to their nature. Drawing on neo-Aristotelian naturalism, particularly the work of Philippa Foot and Michael Thompson, I contend that "health" for any living system is functioning in a way that fulfils its natural end, and that a "good" human life is one that promotes the health of the larger systems—family, community, biosphere—in which human life is embedded.

Three further refinements strengthen this bridge. First, I introduce a graduated account of normative standing, distinguishing systems by their processing capacities, to explain why human mental health possesses the specific normative structure it does without attributing full moral status to every aggregate. Second, I employ the concept of "carriers"—the physical mediators through which all relationships are realized—to ground the abstract language of "cooperation" and "integration" in empirically respectable terms. Third, I develop an account of "emotional capture" to explain why the repulsive processing that constitutes mental illness is often self-reinforcing even when the individual recognizes its destructiveness, thereby addressing a gap in purely cognitive accounts of psychopathology.

3.1. Hume's Challenge and Its Limits

Hume's famous remark in “A Treatise of Human Nature” sets the terms of the problem:

“In every system of morality, which I have hitherto met with, I have always remark'd, that the author proceeds for some time in the ordinary way of reasoning, and establishes the being of a God, or makes observations concerning human affairs; when of a sudden I am surpriz'd to find, that instead of the usual copulations of propositions, “is”, and “is not”, I meet with no proposition that is not connected with an “ought”, or an “ought not”. This change is imperceptible; but is, however, of the last consequence. For as this “ought”, or “ought not”, expresses some new relation or affirmation, 'tis necessary that it shou'd be observ'd and explain'd; and at the same time that a reason should be given, for what seems altogether inconceivable, how this new relation can be a deduction from others, which are entirely different from it.”

The standard interpretation—often called Hume's Law or the is-ought gap—holds that no set of purely descriptive premises can logically entail a prescriptive conclusion. If this were the final word, any attempt to ground normative claims in the descriptive pattern of Chapter 2 would be doomed from the outset.

But Hume's challenge, properly understood, is a logical point about deduction, not an absolute prohibition on grounding norms in facts. He demands that a reason be given for the transition, and that the new relation be explained. The teleological tradition, revived in modern neo-Aristotelian naturalism, provides precisely such a reason: evaluative conclusions follow not from bare factual statements but from facts about the nature of a thing—specifically, facts about what it is to be a well-functioning instance of its kind.

3.2. The Return of Teleology: Objective Functions in Living Systems

To see how this is possible, we must recover a concept that modern philosophy, under the influence of mechanistic science, largely abandoned: telos—end, purpose, or function. The expulsion of final causes from physics did not eliminate teleology from biology. A heart has the function of pumping blood; an eye has the function of enabling sight; a root system has the function of absorbing water and nutrients. These are not projections of human preference. They are objective facts about what these structures do within the organized system of a living organism, facts that explain why the structures exist in the form they do and what counts as their proper or defective operation. A heart that fails to pump blood adequately is dysfunctional, regardless of anyone's opinion about it. The concept of function is ineliminable from biology, and with it comes an ineliminable normativity internal to the description of living things.

This normativity is not an externally imposed value; it is constitutive of what it means to be a living organism. Living things are processing systems—entities that maintain themselves far from thermodynamic equilibrium by receiving, processing, and reacting to matter and energy. To continue existing as the kind of thing they are, they must perform certain activities successfully. This self-maintaining, self-organizing character is their natural end: to live, to flourish, to reproduce. The term "flourishing" here is not a sentimental add-on; it denotes the objectively describable state of functioning well according to the life-form that defines the organism's kind.

It is worth pausing to clarify the physical basis of the functional organization under discussion, for doing so pre-empts the charge that "cooperation" and "integration" are merely metaphorical. All relationships between individuals—from quarks binding into protons to humans forming communities—are physically realized through the exchange of mediators. In physical systems, these are gauge bosons: photons mediate electromagnetic interactions, gluons mediate strong nuclear binding. In chemical systems, shared electrons mediate covalent bonds. In biological systems, signalling molecules—hormones, neurotransmitters, cytokines—mediate the coordination of cellular and organismic activity. In social systems, symbols—words, gestures, written texts—mediate the coordination of human action. I will call these mediators carriers: the physical entities whose exchange constitutes the relationships that bind individuals into systems. The concept is not a metaphysical addition but a recognition that relationships are always physically realized. When I speak of "cooperative relationships" in what follows, I refer to

patterns of carrier exchange that sustain the integrity of the systems they constitute. This provides an empirical anchor for what might otherwise seem abstract.

3.3. Philippa Foot and the Naturalistic Criterion of Goodness

Philippa Foot, in “Natural Goodness” (2001), developed precisely this line of argument. She observed that we routinely make evaluative judgments about plants and animals without any sense of projecting subjective values onto them. We say that a tree has "good" roots or that a wolf has "poor" vision, and we understand these as factual assessments. The standard by which we judge is the species life-form—the characteristic pattern of development, sustenance, and reproduction that defines what it is to be a healthy instance of that species. A "good" oak is one whose roots anchor it firmly, whose leaves photosynthesize efficiently, and whose acorns are viable. A "defective" oak is one that fails in these respects. The judgment is objective, grounded in the natural history of the organism, and it yields a kind of "natural normativity." There are things an oak "needs"—water, sunlight, soil nutrients—and things that are "good for it," quite independently of human desires.

Foot's insight resolves a long-standing confusion about the naturalistic fallacy. G.E. Moore charged that any attempt to define "good" in natural terms commits a fallacy because one can always ask, "But is that natural property itself good?" However, Foot's account does not identify goodness with any simple natural property like pleasure or evolutionary success. Rather, goodness is a relational status: it consists in the relation between an individual organism's actual functioning and the life-form that specifies what functioning well consists in for its kind. This is analogous to the way "health" operates in medicine—it is not a single measurable quantity but a complex, norm-governed condition defined by the proper functioning of interdependent systems. The life-form provides the standard; the individual's state is measured against it. The question "But is good functioning itself good?" is otiose, because the life-form constitutes what goodness consists in for members of that species. There is no further fact to appeal to.

Michael Thompson's “Life and Action” (2008) strengthens this account by arguing that life-form judgments possess a distinctive logical form. When we say "the bobcat has four young in a litter," we are not making a statistical generalization about what most bobcats do. We are making a normative claim about the bobcat life-form—what it is to be a flourishing bobcat—that holds even if most actual bobcats, due to injury or environmental stress, produce fewer. The life-form concept provides the standard against which individual performance is measured, and this standard is not reducible to statistical norms. This is why a trait can be statistically common yet pathological (e.g., dental caries in populations with high-sugar diets) or statistically rare yet healthy (e.g., exceptional longevity). The normativity is grounded in the life-form, not in the distribution.

3.4. Alternative Metaethical Frameworks and Why Neo-Aristotelian Naturalism Is Preferable

The neo-Aristotelian naturalist position developed above is contested. Philosophers committed to different metaethical frameworks reject the claim that evaluative judgments about living systems can be objective discoveries rather than expressions of attitude or constructions of practical reason. While a comprehensive defence would require a separate treatise, acknowledging the two strongest alternatives—and explaining why the present framework is preferable to each—demonstrates that the position adopted here is not naively assumed but chosen in light of the available options.

3.4.1. Expressivism

Expressivism holds that moral and evaluative judgments do not function primarily as descriptions of fact but as expressions of non-cognitive attitudes—desires, emotions, plans, or commitments. In its most sophisticated contemporary forms (Gibbard, 1990; Blackburn, 1998; Schroeder, 2008), expressivism explains how normative language can behave logically like descriptive language while remaining fundamentally expressive in function. To call a pattern of processing "mentally healthy" is not, on this view, to report a fact about that pattern but to express one's endorsement of it, to prescribe its cultivation, or to commit oneself to norms that license certain responses to it.

Expressivism directly challenges the framework's central claim: that evaluative judgments about mental health are "internal to the description of living systems." For the expressivist, when the veterinarian judges a wolf's heart defective, the factual component (the heart fails to pump blood adequately) can be objectively true, but the normative component (this failure is bad, a defect warranting concern) expresses an attitude that is not itself answerable to facts in the way the factual component is. Applied to mental health: the expressivist can grant all the descriptive claims in Chapters 2 and 4—that humans are constitutively relational, that certain processing patterns degrade cooperative connections—while maintaining that the judgment that such patterns constitute disorder expresses collective attitudes rather than tracking objective normative facts. The 1973 homosexuality vote was, on this view, simply a shift in collective attitudes, and no deeper resolution of the normative crisis is available.

Why neo-Aristotelian naturalism is preferable. Three considerations favour the present framework over expressivism.

First, expressivism cannot adequately account for the phenomenology of normative discovery in the life sciences. When a biologist identifies a cellular mechanism as functioning to repair DNA, or when a clinician recognizes that a particular emotional pattern systematically undermines relationships, the experience is one of finding something—of detecting a normative structure that was already there—not of projecting an attitude onto neutral facts. The expressivist must claim that this phenomenology is misleading, that the sense of discovery attaches only to the factual substrate while the normative valence is supplied by the attitude. But this bifurcation is artificial. The judgment "this mechanism is dysfunctional" presents itself as a unified assessment that can be correct or incorrect depending on how the mechanism actually relates to the system's persistence. The neo-Aristotelian framework respects this phenomenology; expressivism must explain it away.

Second, expressivism struggles to explain why certain evaluative judgments are non-optional in ways that mere expressions of attitude are not. It is not simply a cultural preference to say that a heart incapable of pumping blood is defective. Someone who insisted that such a heart was "functioning perfectly well according to its own unique way of being" would be making a mistake, not merely expressing a different attitude. The expressivist can appeal to shared linguistic norms to explain the appearance of objectivity, but the deeper question is whether those norms are answerable to anything beyond further attitudes. The neo-Aristotelian answers that they are: the norms governing functional evaluation are answerable to facts about what living systems of a given kind need to flourish. A wolf needs a functioning heart to live; this need is not a projection but an objective constraint on the wolf's persistence. Normative judgments that respect such constraints are correct; those that ignore them are mistaken. Expressivism cannot vindicate this distinction between norms that track genuine needs and norms that merely express parochial attitudes.

Third, and most decisively for the purposes of this paper, expressivism leaves psychiatric classification hostage to the very cultural contingency the framework was designed to escape. If the judgment that a pattern of processing is a "disorder" ultimately expresses collective attitudes, then the 1973 homosexuality vote was a collective attitude change—no more or less justified, in objective terms, than the previous attitude that pathologized homosexuality. The expressivist can say that the new attitude is better by standards we now endorse, but cannot say that it is more accurate—that it better tracks the normative structure of human flourishing. This is precisely the situation that the DSM's critics have identified as the normative crisis. The present framework offers a genuine alternative: it claims that some evaluative judgments about mental health are more accurate than others because human beings have an objective nature that specifies what flourishing consists in, and that this nature is discoverable through empirical investigation and philosophical reflection. Expressivism cannot offer this, and the normative crisis persists unresolved under its auspices.

3.4.2. Kantian Constructivism

Constructivism, in its most influential contemporary form (Korsgaard, 1996, 2009), holds that normative truths are not discovered but constructed through procedures of practical reason. For Korsgaard, the source of normativity is the structure of reflective consciousness itself: as self-conscious agents, we must act on reasons, and this necessitates valuing our own humanity—our capacity for rational choice. From this starting point, universal moral obligations are generated through the exercise of practical reason. Constructivism is not expressivism; it holds that there are correct answers to normative questions. But these answers are constituted by what rational agents would agree to under idealized conditions, not by correspondence to mind-independent normative facts. Normative truths are real but constructed by practical reason rather than discovered in nature.

Constructivism challenges the framework's claim that normative standards for mental health can be grounded directly in the teleological structure of living systems. For the constructivist, the fact that organisms have functional ends does not, by itself, generate normative reasons for reflective agents. The

wolf's heart has the function of pumping blood—this is a biological fact. But the claim that I ought to respect the wolf's flourishing, or that the wolf's flourishing is good, requires a further step: the exercise of practical reason that constructs value from the standpoint of a reflective agent. The framework, on this view, conflates biological normativity (what a system needs to function) with practical normativity (what agents have reason to do or value). These are, the constructivist argues, different kinds of normativity, and the move from one to the other requires rational construction that the framework does not adequately theorize.

Constructivism also offers an alternative diagnosis of the normative crisis in psychiatry. The crisis, on this view, arises not from the absence of objective normative facts but from the failure to conduct practical reasoning properly—to include all affected parties, to respect the autonomy of patients, to reason in ways that could be universally endorsed. The 1973 homosexuality vote was a failure of this kind: the reasoning that pathologized homosexuality could not be endorsed by homosexual persons themselves under conditions of reflective autonomy. The solution is not to ground psychiatry in natural teleology but to ensure that psychiatric classification is governed by procedures of practical reason that respect the autonomy of all persons.

Why neo-Aristotelian naturalism is preferable. Four considerations favour the present framework over constructivism.

First, constructivism's sharp distinction between biological and practical normativity is itself questionable. The neo-Aristotelian argues that practical normativity is not a separate domain added to biological facts by an act of rational construction; it is an extension and refinement of the normativity already present in living systems. Consider the judgment: "This wound needs to be cleaned." This is simultaneously a biological claim (uncleaned wounds become infected, compromising the organism's integrity) and a practical claim (one has reason to clean the wound). The biological and practical dimensions are not independent; the practical claim is motivated by and grounded in the biological reality of the organism's vulnerability. The neo-Aristotelian argues that this integration of biological and practical normativity is the norm in our actual evaluative practices, and that constructivism's attempt to bifurcate them creates an artificial problem that it then struggles to solve.

Second, constructivism faces a version of the "empty formalism" objection. Kantian constructivism grounds normativity in the structure of rational agency—in what any reflective agent must value to be an agent at all. But this starting point is extraordinarily thin. It generates formal constraints on action (universalizability, respect for autonomy) but struggles to yield substantive conclusions about what constitutes human flourishing. Korsgaard can explain why we must respect others' autonomy, but she cannot easily explain why connection, contribution, or nature-connectedness are constitutive of well-being. These look like substantive claims about human nature that constructivism's procedural starting point cannot generate. For a theory of mental health—which requires distinguishing healthy from pathological processing in rich, domain-specific ways—this thickness is a significant advantage of the neo-Aristotelian approach, which starts from a thick description of the human life-form and generates substantive conclusions about flourishing from that description.

Third, constructivism's response to the normative crisis in psychiatry is less determinate than it appears. The constructivist proposes that psychiatric classification should be governed by procedures that respect the autonomy of all affected parties. But specifying these procedures non-question-beggingly is notoriously difficult. Consider the gender dysphoria case analysed in Chapter 6. In a non-affirming community, the dysphoric individual's reflective endorsement of transition conflicts with family members' reflective rejection of it. The constructivist appeals to idealized conditions of reasoning, but how these conditions adjudicate the conflict is unclear. The neo-Aristotelian framework, by contrast, provides substantive criteria—mutual benefit, emergent capacity, dynamic stability—for distinguishing genuine integrative wholes from pseudo-wholes, and for determining when an individual's processing is healthy. These criteria are contestable, but they provide more guidance than a procedural standard alone.

Fourth, and most fundamentally, constructivism cannot explain why the functional ends of living systems matter in the way the framework claims they do without making those ends dependent on an act of rational endorsement. For the constructivist, a wolf's needs become normative for me only through my reflective endorsement of valuing the wolf's flourishing. But this seems to reverse the actual phenomenology of value. I do not first decide to value the wolf's flourishing and then recognize its needs as reason-giving. I encounter the wolf's needs as already reason-giving, and this recognition motivates my reflective endorsement, not vice versa. The neo-Aristotelian argues that this is because value is not constructed by practical reason but encountered in the world—specifically, in the flourishing and suffering of living beings whose teleological structure makes them the kind of things that can flourish or suffer. The framework's graduated account of normative standing (3.5 below) captures this insight: different levels of processing capacity generate different kinds of normative claims, and these claims are features of the beings themselves, not products of our reflective procedures.

3.4.3. The Defensible Core

Neither expressivism nor Kantian constructivism is incoherent. Both are sophisticated research programs with significant achievements. But each purchases its philosophical advantages at a cost: expressivism cannot vindicate the distinction between norms that track genuine human needs and norms that merely express parochial attitudes, and constructivism cannot generate the thick, substantive conclusions about human flourishing that a theory of mental health requires. Neo-Aristotelian naturalism, for all its own challenges, stays closer to the normative phenomena it aims to explain—the way evaluative judgments actually operate in the life sciences and clinical practice, as discoveries about what living systems need to flourish, discoveries that are simultaneously factual and normative and that carry genuine practical force without requiring a separate act of rational construction or attitudinal expression. It is this framework that the remainder of this chapter develops.

3.5. A Graduated Account of Normative Standing

The argument thus far establishes a teleological basis for evaluating individual organisms. But the living world exhibits enormous variation in organizational complexity, and not all living systems merit the same kind of normative consideration. A bacterium, an oak, a wolf, and a human being are all living systems with objective functional ends, yet we rightly recognize that their interests carry different weight. The neo-Aristotelian framework can accommodate this intuition without abandoning its naturalistic foundation by distinguishing levels of processing capacity.

Consider three broad categories along a continuous spectrum:

Minimal Experience. The most basic living systems—bacteria, single cells, simple multicellular organisms—receive information from their environment, process it relative to internal constraints (maintaining membrane integrity, acquiring energy, avoiding toxins), and react adaptively. They have interests in the minimal sense that certain states of affairs satisfy their constitutive constraints and others violate them. This is normativity in its most primitive form: a factual relation between an organism's state and its conditions of persistence. We need not attribute consciousness or sentience to recognize that a bacterium deprived of nutrients is failing to flourish relative to its life-form.

Sentience. More complex organisms—vertebrates, mammals, birds, and likely some cephalopods and arthropods—possess centralized nervous systems capable of integrating sensory information with emotional valence. They do not merely register stimuli; they feel them as pleasant or aversive. This capacity, which we may call core consciousness or sentience, introduces a qualitatively new dimension of normative standing. A wolf does not merely have interests; it has a subjective point of view from which those interests matter. Its suffering is not merely a functional failure but an experienced state that calls for alleviation. The capacity to suffer generates stronger normative claims than the mere violation of constitutive constraints in systems that cannot feel.

Reflective Consciousness. Human beings—and possibly some cetaceans, great apes, and future artificial systems—possess an additional capacity: we can reflect on our own processing, form second-order desires about our first-order desires, evaluate our lives against normative standards, and deliberately modify our decision-making in light of those evaluations. This capacity for reflective self-awareness does not make human interests more real than those of sentient animals, but it does make them more complex. A human being can suffer not only from pain but from meaninglessness; can fail to flourish not only from starvation but from alienation; can experience impairment not only in bodily function but in the narrative coherence of a life. Mental health, as a distinct domain of normative concern, emerges precisely because human beings are the kind of processing systems whose flourishing depends on the quality of their reflective, meaning-generating activity.

This graduated account has three implications for the argument to come. First, it explains why mental health possesses the specific normative structure it does without overgeneralizing. Mental health is not merely the absence of biological dysfunction; it is the proper functioning of a self-reflective processing system whose constitutive constraints include meaning, connection, and contribution beyond the self. Second, it locates human normativity on a continuum with the rest of the living world, demonstrating

that the teleological bridge does not require an unmotivated leap from physical description to human ethics but proceeds through stages of increasing complexity, each grounded in the one before. Third, it provides resources for addressing concerns about anthropocentrism: the framework recognizes genuine normative standing in non-human systems without collapsing all distinctions.

3.6. From Individual Health to Systemic Health

The argument thus far establishes a teleological basis for evaluating individual organisms. But the cosmic pattern described in Chapter 2 reveals something further: individuals are themselves constituted by cooperative relationships among parts, and they, in turn, enter into cooperative relationships that constitute larger systems. An organism is a society of cells; a family is a society of organisms; an ecosystem is a society of species. If teleological evaluation applies to individual organisms, it should in principle apply to these larger systems as well—provided we can identify what it is for such a system to function well, to flourish, to be "healthy" in a structurally analogous sense.

A system, as defined in the previous chapter, is a set of individuals whose interrelationships, mediated by carriers, give rise to a new whole with emergent properties. The system persists only as long as the cooperative relationships holding it together remain within certain bounds. If those relationships degrade beyond a threshold, the system disintegrates—the cell dies, the family dissolves, the ecosystem collapses. The system's telos, its natural end, is nothing other than its own continued existence as the integrated whole it is. For a system to be "healthy" is for the cooperative dynamics among its parts to be robust, adaptive, and self-stabilizing—functioning in a way that maintains the system's integrity over time. For it to be "unhealthy" is for those dynamics to degrade, for cooperation to give way to destructive conflict or parasitic exploitation that threatens the whole.

This is not a metaphor. In ecology, the concept of "ecosystem health" is operationalized through measures of resilience, nutrient cycling, biodiversity, and productivity—all of which track the system's capacity to maintain its organized complexity. In sociology, a "healthy" family or community is one whose relational bonds enable members to flourish while sustaining the collective unit. In each case, the standard of evaluation is objective, grounded in the functional requirements of the system's persistence as a system of that type.

Crucially, this standard is not invented by the evaluator; it is discovered by observing what conditions enable such systems to persist versus what conditions lead to their breakdown. The pattern of Chapter 2 is precisely a map of those conditions: wherever complex, organized wholes exist, they exist because cooperative relationships—realized through specific patterns of carrier exchange—have been sufficiently stable to constitute them. That is the "is." The "good" for any such system is simply the continuation and enhancement of that stability—functioning well according to the relational architecture that defines it. The carriers that mediate cooperative exchange at each level—photons for atoms, neurotransmitters for neural circuits, symbols for human communities—are not metaphors but physical realities whose adequate functioning constitutes systemic health.

3.7. The Human Situation: Nested Teleology

Human beings are peculiar in two respects relevant here. First, we are self-aware: we can reflect on our own functioning and deliberately choose actions that affect it. This reflective capacity, as noted above, is what gives mental health its distinctive normative structure. Second, we are embedded in multiple overlapping systems simultaneously—families, communities, nations, the global economy, the biosphere. Our individual flourishing is not independent of these larger systems. A human being cannot flourish in isolation; our characteristic life-form includes language, culture, cooperation, and social bonds. What is "good for" an individual human, therefore, is not definable merely by reference to that individual's biological survival. It includes the health of the relational systems without which the individual cannot fully realize the human life-form.

This yields a nested teleology. The telos of a cell is served in part by contributing to the health of the organ it belongs to; the telos of the organ is served by contributing to the organism; the telos of the organism is served by contributing to the family and community; and so on. There is no sharp boundary where "my" good ends and "the whole's" good begins, because my identity as the kind of being I am is partly constituted by my participation in these larger wholes. I am a relational being through and through—biologically, psychologically, and socially. My capacity for language, thought, and selfhood itself emerged through relationships with caregivers and culture, mediated by the exchange of symbols and emotional signals that serve as carriers at the social level. The cosmic pattern runs straight through the centre of my being.

It follows that a "good" human decision is not simply one that maximizes my individual pleasure or even my individual biological health. It is one that promotes the health of the larger systems of which I am a part, because my own flourishing depends on their flourishing. Acting in a way that undermines family bonds, community cohesion, or ecological integrity is, on this view, akin to a cell turning cancerous—pursuing its own replication at the expense of the organism, thereby destroying the very conditions of its own survival. The judgment that cancer is "bad" is not a subjective cultural opinion; it is an assessment grounded in the functional requirements of the multicellular system. Likewise, the judgment that certain kinds of human action are "bad" follows from the functional requirements of the social and ecological systems that sustain human life.

3.8. Emotional Capture: Why Repulsive Processing Is Self-Reinforcing

Before completing the bridge from "is" to "ought," we must address a psychological phenomenon that threatens the practical applicability of the framework. Chapter 4 will define mental illness as repulsive processing—a pattern of decision-making that degrades the cooperative connections upon which flourishing depends. But why, once such patterns take hold, are they so difficult to escape? Why do individuals persist in processing that demonstrably undermines their own well-being and the well-being of those they care about?

The answer lies in what I will call emotional capture. The human decision-making system is not a unified, rational executor. It comprises multiple interacting subsystems, including phylogenetically ancient emotional circuits that evolved for rapid evaluation of threats and opportunities in ancestral environments. These circuits generate powerful motivational valences—fear, desire, anger, attachment—that shape which information is attended to, how it is interpreted, and which responses feel available.

Under normal conditions, these emotional systems serve integrative functions: fear motivates avoidance of genuine danger, attachment motivates care for dependent others, anger motivates defence against exploitation. But these systems can become decoupled from their proper objects and instead capture the entire decision-making apparatus. When this occurs, conceptual processing—the capacity to reason, evaluate evidence, and consider alternatives—ceases to function as an independent check on emotional reactivity and instead becomes its servant. The individual does not lose the capacity to reason; they deploy reason selectively in the service of emotional imperatives that are no longer being critically examined.

The result is what I will call a meaning-capture loop: emotionally generated meanings shape conceptual processing, which generates interpretations that reinforce the emotional meanings, which further entrench the concepts, and so on. A depressed individual's anhedonia and despair generate the conviction "I am worthless and nothing will help." This conviction then screens out evidence of worth and possibility, confirming the despair. An anxious individual's hypervigilance generates the interpretation "That ambiguous expression means they're angry with me." This interpretation then justifies further vigilance and avoidance, preventing the disconfirming experience of discovering the expression was neutral. A narcissistically organized individual's need for admiration generates the conviction "I am exceptional and entitled to special treatment." This conviction then filters out feedback that contradicts it, entrenching the very processing patterns that alienate others and thwart genuine connection.

Emotional capture explains why repulsive processing is not simply corrected by new information. The depressed person is not merely ignorant of their worth; their processing system is organized to reject evidence of it. The anxious person is not merely unaware that their fears are disproportionate; their processing system is organized to interpret any ambiguity as confirmation. The narcissist is not merely overconfident; their processing system is organized to dismiss challenges to their self-conception. In each case, the carrier exchanges that normally mediate integrative connection—the words, gestures, and emotional signals that bind individuals into cooperative wholes—are systematically distorted. The depressed person's flat affect and withdrawal signal to others that connection is unwelcome, even when connection is desperately needed. The anxious person's reassurance-seeking exhausts the very relationships it seeks to preserve. The narcissist's grandiosity pre-empts the mutual recognition upon which genuine intimacy depends.

The clinical implication is significant: therapeutic intervention must do more than provide corrective information. It must interrupt the meaning-capture loop itself—creating conditions in which emotional processing can be recalibrated through new relational experiences that the old processing patterns would otherwise foreclose. This insight will prove central to the therapeutic analysis in Chapter 5.

3.9. Bridging the Gap: How "Is" Yields "Ought"

We are now in a position to explain precisely how the "ought" emerges from the "is," satisfying Hume's demand for a reason.

The critical premise, which Hume's empiricist framework tended to obscure, is that living beings have a nature. They are not mere collections of particles in arbitrary arrangements; they are organized systems with objective functional requirements that must be met for them to persist as the kind of thing they are. This is not a metaphysical postulate added to the empirical facts; it is a fact discernible empirically—one can observe what a wolf needs, what damages it, what enables it to flourish. The "is" of a wolf includes its teleological structure.

From this factual premise, together with the fact that a given individual is a wolf, we can derive evaluative conclusions: this wolf's heart is defective; that wolf's diet is inadequate; this behaviour is good for the pack. These are not expressions of sentiment, though they may evoke sentiment; they are factual judgments with truth conditions that can be investigated. A veterinarian who says "This wolf is in poor health" is making a claim about the world, not merely expressing a personal attitude.

The extension from biological evaluation to moral evaluation requires an additional step, but not a discontinuous leap. If the human life-form includes participation in larger relational systems as an essential component—if we are the kind of beings who realize our nature through cooperative bonds, mediated by the exchange of symbols, emotional signals, and shared practices—then actions that damage those bonds are, in the most literal sense, bad for us. They undermine the conditions of our flourishing. "Moral" goodness is not a separate domain from "prudential" goodness; it is prudential goodness extended across the nested systems that constitute the human self. To harm my community is to harm myself, because my self is partly constituted by my relational place in that community. This is not mysticism; it is a recognition of the interdependence that Chapter 2 established as a general feature of reality.

The graduated account of normative standing clarifies the scope of this claim. The framework does not entail that every living system has the same moral status as a human being. Minimal experience, sentience, and reflective consciousness generate different kinds and degrees of normative consideration. But the structure of normativity—the relation between a system's state and its constitutive constraints—is the same across levels. Human mental health is a particular instance of a general phenomenon: the flourishing of a processing system whose natural end includes the health of the larger wholes it participates in.

Thus, the "ought" is not deduced from a pure "is" in the way that a geometric theorem is deduced from axioms. Rather, it is unpacked from a more complete understanding of what the "is" contains. Once we recognize that a human being is a teleologically structured system embedded within larger teleologically structured systems, the question "How ought I to live?" becomes answerable by reference to the functional requirements of those systems. The ought is internal to the is: if one is a human being, then

one needs certain goods to flourish, and those goods include the health of the relational wholes one participates in. The only way to reject the ought is to reject the is—to deny that human beings have a nature, or that their nature is constitutively relational. But such a denial flies in the face of everything biology, psychology, and the cosmic pattern tell us.

Crucially, the emotional capture analysis explains why recognizing the "ought" is not always sufficient to produce the corresponding motivation. The is-ought gap, often treated as a purely logical problem, also has a psychological dimension: it is the experience of a processing system whose conceptual recognition of what flourishing requires is not yet integrated with its emotional and motivational architecture. This is not a refutation of the teleological bridge but a diagnosis of why it can be difficult to cross in practice. The gap is real, but it is a gap within the processor, not a gap in the structure of reality. Just as a depressed individual may recognize intellectually that they are not worthless while remaining emotionally captured by the conviction that they are, a person may recognize intellectually that their flourishing depends on cooperative integration while remaining motivationally captured by narrower self-interest. The task of moral education and therapeutic intervention alike is to bring emotional processing into alignment with accurate conceptual recognition—to close the gap not by argument alone but by practices that recalibrate the whole person.

3.10. Conclusion

The teleological bridge rests on four pillars. First, living systems have objective functional ends that are discernible features of their nature, not subjective projections. This claim has been defended against expressivist and constructivist alternatives, with each shown to purchase its philosophical advantages at the cost of losing touch with the normative phenomena it aims to explain. Second, health is the state of functioning well relative to those ends, and "good" is what promotes or constitutes that state. Third, the human life-form is constitutively relational, such that the health of the individual depends on the health of the larger cooperative systems—families, communities, biosphere—that the cosmic pattern generates at every scale. Fourth, the gap between recognizing what flourishing requires and being motivated to pursue it is not a logical refutation of natural normativity but a psychological phenomenon—emotional capture—that explains why repulsive processing can persist even when it is cognitively understood to be destructive.

The graduated account of normative standing situates this framework within the broader living world, showing that human mental health is a particular expression of principles observable across all processing systems. The carrier concept anchors the language of "cooperation" and "integration" in the physical exchanges—from photon to neurotransmitter to symbol—that constitute all relationships. And the emotional capture analysis illuminates the self-reinforcing character of the repulsive processing that the next chapter will define as mental illness.

The descriptive "is" of Chapter 2 thus turns out to contain within itself the seed of an "ought," not by logical trickery but by revealing that the ordered, relational structure of reality is not normatively inert. To be a part of that structure is to have a stake in its flourishing. The question "How ought I to live?"

finds its answer not in a leap from fact to value, but in a deeper apprehension of what the facts actually are: that we are cooperative systems within cooperative systems, that our good is inseparable from the good of the wholes we participate in, and that health—from the cell to the biosphere—is the objective measure of our success in aligning our decisions with the order that sustains us. Chapter 2 mapped the territory; this chapter has shown that the map is also a guide. The way forward is not imposed from outside but drawn from within, from the nature of the beings we are and the living world we inhabit. The next chapter applies this teleological framework to the specific domain of mental health, defining it as systemic alignment and mental illness as repulsive processing—a pattern of emotional capture that degrades the cooperative connections upon which flourishing depends, and which requires more than information to be undone.

Chapter 4: Mental Health as Systemic Alignment

The preceding chapters have established two foundational claims. Chapter 2 demonstrated that the cosmos exhibits a recurrent descriptive pattern: individuals enter into cooperative relationships to form more encompassing systems, a nested hierarchy observable from quantum binding to human institutions—and that cooperative integration functions as a possibility condition for organized complexity at any scale. Chapter 3 argued that this descriptive pattern grounds a teleological conception of goodness, wherein health consists in the proper functioning of a system relative to its natural end, and that human flourishing is constitutively relational—dependent upon the integrity of the larger wholes we inhabit. With these foundations in place, the present chapter advances a definition of mental health that follows directly from them. Mental health is systemic alignment: the condition in which an individual's decision-making system reliably produces reactions (sensory, emotional, language contents of mind, and behaviour) that sustain and generate integrative wholes. Mental illness, correspondingly, is a pattern of processing that degrades these connections—a repulsive dynamic that unravels the relational fabric upon which both individual and collective flourishing depend.

4.1. The Decision-Making System as the Locus of Mental Function

Every living organism is a processing system that receives, processes, and reacts to matter, energy, and information. In human beings, this processing architecture has attained a distinctive level of complexity through the evolution of a dedicated neural information system—the brain—capable of generating not merely reflexive responses but a continuous stream of internally represented contents: sensory imagery, emotional valences, and linguistic thought. These contents, together with overt behaviour, constitute the reactions of the human processing system. They are the observable outputs of decision-making, whether that decision-making occurs through intentional, conscious deliberation or through the automatic operation of deeply ingrained patterns.

The term "decision-making system" designates the totality of structures and processes—neural, emotional, conceptual, and behavioural—that determine how an individual moves through the world. It includes both the subconscious architecture of learned associations and the conscious capacity for reflection and choice. This system is not static; it is shaped continuously by experience, by the information it receives from the senses, from the body's internal states, and crucially, from its own prior outputs via the feedback loop of conscious awareness. The contents of mind are simultaneously reactions and inputs, and this recursive structure means that the decision-making system is self-modifying. How one processes today shapes how one will process tomorrow.

It is important to clarify that describing the mind as part of a "processing system" does not entail an internalist or computationally reductionist picture of human psychology. The framework advanced here converges in significant respects with 4E approaches to cognition—which emphasize that cognition is embodied, embedded, enacted, and extended—while differing on certain points of theoretical emphasis. Like 4E theorists, the present framework insists that human mental life cannot be understood

in isolation from the body, the social environment, and the broader systems in which persons are embedded. The decision-making system described here is constitutively relational: its "inputs" are not raw data from a detached sensorium but the felt presence of other persons, the shared symbols of culture, the affordances of the natural world, and the accumulated history of one's own prior interactions. Its "outputs" are not merely internal representations but actions that feed back into and reshape the relational fields from which future experience arises. The self that processes is itself constituted through these ongoing exchanges—a point the framework shares with enactive and embedded approaches that treat the boundary between organism and environment as permeable and dynamically negotiated.

Where the framework differs from some 4E approaches is in its emphasis on the normative structure of this relational embeddedness. The 4E tradition has been primarily descriptive, showing how cognition actually occurs across brain, body, and world. The present framework adds a teleological dimension: the constitutive relationality of human selfhood is not merely a fact about how we happen to think but a feature of our nature that grounds normative judgments about health and dysfunction. Mental health consists not simply in being appropriately embedded—a description that could apply to anyone enmeshed in a toxic system—but in processing that sustains genuinely integrative wholes, as defined by the criteria of mutual benefit, emergent capacity, and dynamic stability introduced in Chapter 3 and elaborated below. The "system" in "processing system" is thus not a bounded internal computer but a node in a network, whose processing is healthy when it strengthens the network and unhealthy when it degrades it. This is not internalism; it is relational through and through. But it is a relational account that retains the capacity to distinguish healthy from pathological forms of embeddedness—a capacity that purely descriptive 4E frameworks do not, on their own, provide.

Because mental health pertains to this decision-making system, its assessment must focus on the quality of the system's outputs—the reactions it produces—and on the principles that govern its ongoing self-organization. A decision-making system is not healthy merely because it produces outputs that feel pleasant to the individual or that satisfy immediate desires. Pleasure and pain are evolutionarily ancient signals that orient organisms toward or away from stimuli, but they are not, in themselves, reliable indicators of systemic health. A cell turning cancerous may experience a kind of local "success" in its unconstrained replication, yet this very success destroys the multicellular system upon which its own existence depends. Analogously, a human decision-making system may produce thoughts, emotions, and behaviours that yield short-term gratification while simultaneously degrading the relational wholes—family, community, ecosystem—that sustain the individual's broader flourishing. The measure of health must therefore be relational and systemic, not merely subjective or hedonic.

4.2. Defining Health as Integrative Processing

Health, in its most general teleological sense, is the condition in which a system functions well relative to its natural end. For a living system, that natural end is the continuation and enhancement of its organized complexity—the maintenance of the cooperative dynamics among its parts that constitute it

as the kind of thing it is. A healthy organism is one whose subsystems operate in an integrated manner, each contributing to the persistence and adaptive capacity of the whole. Disease, conversely, is a breakdown of integration: a condition in which some component or process begins to operate in a way that undermines the cooperative relationships upon which systemic integrity depends.

Applied to the human decision-making system, this teleological framework yields a precise criterion. The natural end of the human processing system is not arbitrary; it is specified by the kind of being a human is. Humans are constitutively relational creatures. Our evolutionary history, our developmental trajectory, and our lived experience all testify that we become fully human only through participation in larger wholes. Language acquisition, the very medium of distinctively human thought, occurs through relational immersion. Selfhood emerges through attunement with caregivers. The capacities we value most—reason, empathy, creativity, love—are realized not in isolation but through engagement with other humans, with cultural traditions, with the living systems that sustain our bodies. To be a well-functioning human decision-making system is to process information in ways that strengthen these constitutive connections.

Mental health, therefore, is integrative processing. It is the reliable tendency of the decision-making system to produce reactions—thoughts, emotions, and behaviours—that sustain and create integrative wholes. "Integrative wholes" are systems in which the relationships among parts are cooperative rather than destructive, enabling the emergence of properties and capacities that exceed what any part could achieve alone. A family is an integrative whole when its members support one another's flourishing while maintaining the bonds that define the family as a unit. A community is an integrative whole when its institutions and norms enable cooperation at scale. An ecosystem is an integrative whole when its species interact in ways that maintain nutrient cycles, energy flows, and biodiversity. A mentally healthy individual is one whose decision-making contributes to, rather than extracts from, these larger patterns of integration.

Crucially, this is not a demand for self-sacrifice in the sense of self-negation. The teleological framework reveals that the distinction between "my good" and "the good of the whole" is a false dichotomy when the whole is constitutive of my very identity. A neuron does not sacrifice itself by participating in the brain's information processing; it fulfils its nature precisely through that participation. A parent does not lose herself by caring for her child; she actualizes a dimension of her being that would otherwise remain dormant. Integrative processing is simultaneously self-enhancing and whole-enhancing because the self is not a bounded atom but a node in a network whose flourishing depends on the network's health.

4.3. Mental Illness as Repulsive Processing

If mental health is integrative processing, mental illness is its opposite: repulsive processing. The term "repulsive" is used here not primarily in its emotional sense—though emotional repulsion is often involved—but in its systemic sense. Repulsive processing is a pattern of decision-making that degrades, severs, or prevents the formation of the cooperative connections that constitute integrative wholes. It is

a dynamic that pushes apart what would otherwise cohere, that extracts value from systems without replenishing them, that treats relationships as instrumental rather than constitutive.

The phenomenology of mental illness aligns closely with this characterization. Consider the cognitive patterns characteristic of depression: the withdrawal from social connection, the rumination that isolates the individual in a loop of self-referential thought, the conviction of worthlessness that severs the felt link between self and community. These are not merely painful states; they are processes that systematically dismantle the individual's integration into larger wholes. The depressed person's decision-making system produces thoughts ("I am a burden"), emotions (anhedonia, despair), and behaviours (withdrawal, inactivity) that weaken the relational ties upon which flourishing depends. The illness is self-reinforcing precisely because it degrades the very connections that might otherwise restore health.

Anxiety disorders exhibit a similar structure. The anxious decision-making system over-weights threat, producing avoidance behaviours that contract the individual's sphere of engagement. Opportunities for cooperative participation are foreclosed. The world narrows. Relationships that might provide corrective experience are pre-emptively severed. The processing pattern is repulsive in the literal sense of pushing the individual away from integrative contact with reality and others.

At the more severe end of the spectrum, psychotic disorders involve a breakdown in the shared conceptual frameworks that enable cooperative coordination. When an individual's decision-making system generates beliefs and perceptions radically disconnected from those of others, the possibility of mutual understanding—and thus of the cooperative relationships that mutual understanding enables—is compromised. The individual becomes isolated not by choice but by the very structure of their information processing.

Even conditions often framed in terms of behavioural dysregulation, such as substance use disorders or impulse control disorders, can be understood through this lens. The addicted decision-making system prioritizes the consumption of a substance over the maintenance of relationships, responsibilities, and roles that constitute the individual's integration into family, work, and community. The pursuit of the substance becomes a kind of repulsive attractor, pulling the individual out of orbit around the various integrative wholes that previously structured their life.

In every case, the common thread is not a particular symptom profile but a directional tendency: the decision-making system produces reactions that degrade connection and cooperation, thereby undermining the conditions of its own systemic health. Mental illness, on this view, is a pattern of processing that is self-defeating at the systemic level. Like a cancer cell whose unchecked proliferation destroys the organism it depends on, the mentally ill decision-making system undermines the relational wholes that are the condition of its own flourishing.

4.4. The Spectrum and Its Implications

Positioning integrative and repulsive processing as poles of a spectrum yields several conceptual advantages. First, it normalizes mental health variation without trivializing suffering. Every human decision-making system operates somewhere on this continuum, and movement along it is a function of experience, circumstance, and deliberate effort. One is not simply "healthy" or "ill" in a binary sense; one's processing patterns are more or less integrative, more or less conducive to the creation and maintenance of wholes. This destigmatizes mental illness by framing it as a directional deviation rather than a categorical otherness, while simultaneously preserving the objective reality of dysfunction—a system genuinely can be more or less aligned with the conditions of its own flourishing.

Second, the spectrum model accounts for the cultural and contextual variability of mental health presentations without collapsing into relativism. The specific forms that integrative and repulsive processing take will vary with cultural context, because the wholes that individuals participate in—families, communities, meaning systems—vary in their particular architectures. What counts as relationship-sustaining behaviour in one cultural setting may differ from what counts in another. Yet the underlying principle—that health consists in sustaining the cooperative connections constitutive of the wholes one inhabits—provides a universal metric. One can assess, across any cultural context, whether a given pattern of decision-making tends to build or erode the cooperative dynamics upon which the relevant systems depend. The standard is objective even though its application requires local knowledge.

Third, this framework integrates the biological, psychological, and social dimensions of mental health into a unified account. The biological substrate of the decision-making system—neural circuits, neurotransmitter systems, genetic predispositions—sets parameters on what kinds of processing are possible and how easily they can be modified. Psychological processes—learned associations, cognitive schemas, narrative self-understandings—constitute the proximate mechanisms through which integrative or repulsive patterns are enacted. Social structures—family dynamics, economic conditions, cultural narratives—shape the field of experience that feeds the decision-making system and determine which integrative wholes are available for participation. No single level of analysis is privileged; mental health is an emergent property of the total configuration.

4.5. Situating the Definition Within Established Frameworks

The definition of mental health as systemic alignment does not emerge ex nihilo; it converges with and provides a unifying rationale for several influential approaches in clinical psychology and psychiatry.

The biopsychosocial model, which holds that mental health and illness arise from the interaction of biological, psychological, and social factors, finds in the systemic alignment framework a specification of what "health" across these domains consists in. Biological factors are healthy when they support the organism's capacity for cooperative participation; psychological factors are healthy when they dispose the individual toward integrative processing; social factors are healthy when they provide opportunities for and reinforce connection to larger wholes. The model's strength—its recognition of multiple interacting levels—is complemented by the systemic framework's provision of a common metric across levels.

Cognitive-behavioural approaches identify dysfunctional patterns of thought and behaviour as central to mental illness and aim to replace them with more adaptive alternatives. The systemic alignment framework enriches this approach by specifying what makes a pattern adaptive: it promotes integration rather than repulsion. A thought is not dysfunctional merely because it is negative, but because it degrades the thinker's capacity to participate in and contribute to the cooperative wholes that constitute their flourishing. Behavioural activation, a well-validated treatment for depression, works precisely because it interrupts the repulsive spiral of withdrawal and re-establishes the individual's engagement with rewarding, connection-sustaining activities.

Humanistic and existential traditions emphasize self-actualization, meaning, and authentic connection as central to psychological well-being. The systemic alignment framework grounds these values not in subjective preference but in the objective structure of human existence. Self-actualization is not arbitrary self-expression; it is the realization of one's nature as a constitutively relational being. Meaning is not a subjective overlay on a meaningless world; it is the experience of participating in integrative wholes that transcend the individual. Authenticity is not doing whatever one feels; it is aligning one's decision-making with the relational realities that constitute one's being.

Even psychodynamic perspectives, with their attention to unconscious conflict and early relational patterns, can be understood through this lens. The "repulsive" patterns that characterize mental illness often originate in early experiences where cooperative connection was thwarted or betrayed. The child who learns that attachment figures are dangerous develops processing patterns organized around avoidance—a repulsive adaptation to a repulsive environment. Therapy, on this view, succeeds by providing a relational context in which integrative processing can be experienced and internalized, gradually replacing the older repulsive patterns.

4.6. Conclusion

Mental health, properly understood, is not the absence of symptoms, the attainment of subjective happiness, or conformity to culturally specific norms. It is systemic alignment: the condition in which an individual's decision-making system reliably produces actions, thoughts, and emotions that sustain and generate integrative wholes. This definition follows with necessity from the descriptive pattern of cosmic organization and the teleological account of goodness elaborated in prior chapters. If the universe builds complexity through cooperative relationships that form ever-more-encompassing systems, and if human beings are constitutively relational creatures whose flourishing depends on the health of the wholes they participate in, then the well-functioning human mind is precisely one that processes experience in ways that strengthen rather than weaken those vital connections.

The decision-making system at the centre of this account is not an isolated internal processor but a constitutively relational node whose boundaries are permeable to the social, cultural, and ecological systems that shape and are shaped by its activity. The framework's language of "processing" and "systems" is thus not a retreat into computational internalism but an attempt to describe with precision how a mind embedded in relational fields can function well or poorly relative to the wholes that sustain

it—a project that shares 4E cognition's commitment to situatedness while adding the normative resources necessary to distinguish healthy embeddedness from pathological enmeshment.

Mental illness, correspondingly, is repulsive processing: a pattern of decision-making that degrades, severs, or prevents the cooperative connections upon which individual and collective flourishing depend. This is not a metaphor or a moral condemnation but a structural description. Just as a cancerous cell's processing (its pattern of replication and resource consumption) undermines the multicellular system that sustains it, repulsive human processing undermines the relational systems that sustain human life. And just as the physical survey of Chapter 2 established that cooperative integration is a possibility condition for organized complexity at any scale, repulsive processing is self-undermining because it violates the very conditions under which organized psychological and social life can persist.

This framework does not promise easy answers to every clinical or ethical question. It does, however, provide an intellectually robust, empirically grounded, and practically orienting conception of what mental health is—one that respects the complexity of human psychology while situating it within the broader architecture of a relational universe. The subsequent chapters will explore the practical implications of this definition for the cultivation and maintenance of mental health in individual lives and in the social systems that shape them.

Chapter 5: Predictions and Practical Application

The Testable Implications of Systemic Alignment Theory

A scientific theory earns its place not merely through conceptual elegance but through its capacity to generate empirically testable predictions and to organize existing findings within a unifying framework. The preceding chapters have constructed a theoretical edifice: Chapter 2 established that the observable universe exhibits a recurrent structural pattern in which individuals enter into cooperative relationships to form more encompassing systems—a pattern observable across physical, biological, and social domains, though the normative argument draws primarily on the biological and social instances. Chapter 3 erected a teleological bridge from this pattern to a conception of the good grounded in the objective functional ends of living systems, arguing that human flourishing is constitutively relational and that health consists in functioning well relative to the nested wholes that sustain us. Chapter 4 then applied this framework to mental health, defining it as systemic alignment—the condition in which an individual's decision-making system reliably produces integrative rather than repulsive processing.

While the philosophical architecture supporting these claims has been carefully articulated, the framework would remain speculative if it could not be brought into contact with empirical data. The present chapter demonstrates that the theory generates clear, falsifiable predictions about psychological well-being, and it illustrates how these predictions can be operationalized and tested. Its central contribution is the conceptual specification of the Integrative Processing Orientation (IPO) construct—the proposed disposition to process experience in ways that sustain and generate cooperative wholes—and the detailed articulation of a scale for its measurement. I present the construct's theoretical foundations, its proposed factor structure, its predicted relationships with established measures, and the means by which its discriminant validity can be established. I then show how the systemic alignment framework reframes existing therapeutic modalities, not by replacing their techniques but by providing a unifying theoretical rationale that explains why effective interventions work. The chapter concludes by acknowledging limitations, specifying the conditions under which the framework would be falsified, and identifying the empirical program required to validate or challenge its claims.

5.1. Core Predictions Derivable from the Framework

The systemic alignment theory of mental health yields a central empirical hypothesis: psychological well-being will be positively associated with the degree to which an individual's processing patterns are integrative—that is, oriented toward the creation, maintenance, and enhancement of cooperative relationships that constitute larger wholes. This central hypothesis can be decomposed into several specific, testable predictions.

Prediction 1: Prosocial Integration. The framework holds that human flourishing is constitutively relational, such that the health of the individual depends on the health of the larger cooperative systems in which they participate. From this, it follows that individuals who report higher levels of meaningful social connection, contribution to collective well-being, and a felt sense of belonging to communities

should score higher on standard measures of psychological well-being. This prediction is not reducible to the trivial claim that social contact is pleasant; it specifically predicts that integrative social participation—participation that strengthens the cooperative fabric of families, groups, and communities—will be more strongly associated with well-being than mere social interaction frequency or network size. The quality of relational integration, characterized by mutual benefit, emergent capacity, and dynamic stability, matters more than the quantity of social contact.

Prediction 2: Sense of Purpose Oriented Toward Contribution. The teleological framework elaborated in Chapter 3 grounds human flourishing in the fulfilment of functional ends that extend beyond the individual self. A well-functioning decision-making system, on this account, orients its processing toward goals that contribute to wholes larger than the immediate ego. Consequently, individuals who report a strong sense of purpose—particularly purpose defined in terms of contribution to family, community, or broader systems—should exhibit higher levels of psychological well-being. Moreover, the framework predicts that purpose-derived well-being will be distinguishable from mere hedonic pleasure; it will be associated with eudaimonic measures emphasizing meaning and self-realization rather than simply positive affect. Specifically, contribution-oriented purpose should predict well-being over and above self-referential purpose (career success, wealth accumulation).

Prediction 3: Nature-Connectedness. Chapter 2 demonstrated that the human organism is embedded within nested systems extending from the cellular to the biospheric, and Chapter 3 argued that the health of these larger systems is constitutive of individual flourishing. If mental health consists in alignment with these systems, then a felt sense of connection to the natural world—the most encompassing living system of which humanity is a part—should correlate with psychological well-being. Nature-connectedness, operationalized as an individual's affective, cognitive, and experiential sense of belonging to the natural world, should predict well-being independently of other forms of social connection, reflecting the unique contribution of alignment with the broader ecological system that sustains human life. This prediction distinguishes the framework from theories that treat nature's benefits solely in terms of stress reduction or aesthetic preference.

Prediction 4: Integrative Processing Orientation as a Unifying Latent Factor. Beyond these specific domain predictions, the framework posits a general mechanism: mental health is a function of the degree to which an individual's decision-making system generates integrative rather than repulsive outputs across domains. This suggests that the relationship between the aforementioned predictors (prosocial integration, contribution-oriented purpose, nature-connectedness) and well-being should be mediated by a general factor of Integrative Processing Orientation—a disposition to frame experience and choose actions in ways that build cooperative wholes. If adequately measured, IPO should account for significant variance in well-being across multiple domains and should incrementally predict well-being beyond established personality and values measures.

Prediction 5: Repulsive Processing and Psychopathology. The obverse of the framework's health predictions concerns dysfunction. The theory characterizes mental illness as repulsive processing—patterns of decision-making that degrade, sever, or prevent cooperative connections, often maintained through the emotional capture dynamics described in Chapter 3. This yields the prediction that clinical

populations, particularly those characterized by social withdrawal (depression), threat-biased avoidance (anxiety), reality-disconnection (psychosis), or relational destruction (certain personality disorders), will exhibit processing patterns measurably lower in integrative quality and higher in repulsive dynamics compared to non-clinical populations. Furthermore, therapeutic improvement should be accompanied by measurable shifts toward integrative processing, not merely symptom reduction. Specifically, the emotional capture loop—in which emotionally generated meanings shape conceptual processing in ways that reinforce the original emotional patterns—should be measurably weaker following effective intervention.

5.2. The Integrative Processing Orientation (IPO) Construct

The central empirical contribution of this framework is the proposal of a novel psychological construct: Integrative Processing Orientation. This section provides its theoretical specification.

5.2.1. Theoretical Definition

Integrative Processing Orientation (IPO) is defined as the stable disposition of an individual's decision-making system to produce outputs—thoughts, emotional responses, interpretations, and behaviours—that sustain and generate cooperative relationships constituting larger wholes. IPO is not merely a preference for social harmony or a tendency toward agreeableness. It is a structural property of the processing system itself: the degree to which the system's default operations treat the health of nested systems (family, community, humanity, biosphere) as functionally relevant to its own flourishing, as the teleological framework of Chapter 3 argues they in fact are.

IPO is thus conceived as the psychological manifestation of what the normative framework identifies as health: the alignment of individual processing with the objective relational structure of human existence. Individuals high in IPO do not merely value connection; they process information in ways that reliably generate it. Individuals low in IPO do not merely devalue connection; their processing patterns systematically undermine it, often (as the emotional capture analysis would predict) without conscious recognition that this is occurring.

Phenomenologically, this dispositional gradient is experienced as a fundamental expansion in the boundaries of the self. At low levels of IPO, the self is coterminous with the biological organism—a bounded entity housed within skin and bone, for whom the external world, including other people, appears as a set of instruments or obstacles to be managed. As IPO increases, this felt sense of self progressively enlarges: others cease to be merely ‘not-me’ and become part of the domain whose wellbeing registers as intrinsically self-relevant, first with intimate relations, then extending to broader communities, other sentient creatures, and ultimately encompassing the entire living planetary system. The high-IPO individual does not experience the biosphere as an environment in which they happen to live, but as a larger body of which their personal organism is one expression—a shift in which the

distinction between self-interest and systemic flourishing dissolves, not as a moral achievement, but as a direct feature of how reality is perceived.

5.2.2. Conceptual Distinctions from Related Constructs

To establish discriminant validity, IPO must be distinguished from several established constructs with which it might be conflated.

IPO vs. Agreeableness. Agreeableness (Costa & McCrae, 1992) is a broad personality dimension encompassing trust, straightforwardness, altruism, compliance, modesty, and tender-mindedness. While IPO may correlate positively with agreeableness—both involve prosocial orientation—the constructs are theoretically distinct in several respects. First, IPO is specifically about the capacity to generate and sustain cooperative wholes, not about interpersonal pleasantness per se. A person high in agreeableness but low in IPO might maintain pleasant surface interactions while avoiding the difficult conversations required for genuine relational integration. Conversely, a person moderate in agreeableness but high in IPO might engage in necessary confrontation that temporarily disrupts harmony but ultimately strengthens cooperative bonds. Second, agreeableness is defined primarily at the interpersonal level, whereas IPO spans multiple system levels, including community, institutional, and ecological integration. Third, the theoretical grounding of IPO in the teleological framework provides it with normative content—what counts as "integrative" is not merely socially conventional but functionally related to the health of nested systems—that agreeableness lacks.

IPO vs. Social Desirability. Social desirability (Crowne & Marlowe, 1960) reflects a tendency to present oneself in culturally approved ways. IPO is not a self-presentation strategy but a processing disposition. While socially desirable responding might inflate IPO self-report scores, several features of the construct and its measurement can mitigate this. First, IPO items can be constructed to focus on cognitive and emotional processing patterns rather than overtly evaluative self-descriptions (e.g., "When making an important decision, I naturally consider how it will affect the groups I belong to" rather than "I am a caring person"). Second, the inclusion of reverse-scored items capturing repulsive processing tendencies provides a check on consistency. Third, the construct's predictive validity should survive statistical control for social desirability; if IPO contributes incrementally to well-being beyond a social desirability measure, the concern is substantially mitigated.

IPO vs. Interdependent Self-Construal. Interdependent self-construal (Markus & Kitayama, 1991) captures the degree to which individuals define themselves through relationships and group memberships. IPO shares this relational emphasis but differs in focusing on the processing consequences of such self-construal rather than the self-concept itself. An individual could hold an interdependent self-construal—defining themselves in terms of family roles, for instance—while processing information in ways that are repulsive (e.g., through enmeshment, emotional fusion, or controlling caregiving that undermines others' autonomy). IPO predicts not merely that one sees oneself as connected but that one's cognitive and emotional processing actually serves to build and maintain healthy connection, as defined by the mutual benefit, emergent capacity, and dynamic stability criteria articulated in Chapter 4.

IPO vs. Empathy. Empathy—both cognitive (perspective-taking) and affective (emotional contagion, empathic concern)—is a component of IPO but not equivalent to it. Empathy provides the informational input for integrative processing: understanding others' internal states is necessary for coordinating action with them. However, IPO includes additional capacities: the ability to translate empathic understanding into actions that build cooperative wholes, the capacity to maintain integration even when empathic connection is temporarily strained, and the extension of integrative concern beyond individual others to the larger systems (communities, ecosystems, future generations) that empathy alone may not reach. High empathy without high IPO might produce acute sensitivity to others' distress without the capacity to contribute to systemic solutions; high IPO without exceptionally high empathy might produce steady, reliable integrative contribution without intense emotional resonance.

IPO vs. Light Triad. The Light Triad (Kaufman, Yaden, Hyde, & Tsukayama, 2019) measures Kantianism (treating people as ends), Humanism (valuing dignity and worth), and Faith in Humanity (believing in human goodness). IPO shares conceptual space with these constructs, particularly Kantianism's emphasis on treating persons as ends rather than means. However, IPO is broader in scope: it extends beyond interpersonal ethics to include ecological integration and contribution to impersonal systems (knowledge, institutions, future generations). Moreover, IPO is grounded in a specific theoretical claim about the structure of reality and human flourishing, whereas the Light Triad is primarily descriptive. The theoretical grounding of IPO generates predictions (e.g., about nature-connectedness, about the mediating role of systemic thinking) that the Light Triad alone does not.

5.2.3. Proposed Factor Structure

Based on the theoretical framework, I propose that IPO is a multi-dimensional construct comprising four correlated but distinguishable factors, corresponding to the domains in which integrative processing manifests.

Factor 1: Relational Integration (RI). This factor captures the disposition to process interpersonal experiences in ways that build and maintain cooperative dyadic and small-group relationships. High RI is characterized by: (a) default attention to others' perspectives and needs; (b) interpretive habits that assume good faith until evidence indicates otherwise; (c) emotional responses that motivate repair when relationships are strained rather than withdrawal or retaliation; and (d) behavioural tendencies toward mutual benefit rather than zero-sum competition. Sample items might include: "When someone upsets me, my first instinct is to understand why rather than to defend myself"; "I find that considering others' perspectives usually leads to better decisions for everyone involved"; "When a relationship is struggling, I feel a strong pull to address the issue rather than let it drift."

Factor 2: Systemic Contribution (SC). This factor captures the disposition to orient one's actions and decisions toward the health of larger collectives—organizations, communities, institutions, and society. High SC is characterized by: (a) attention to system-level outcomes beyond immediate personal consequences; (b) interpretive habits that recognize systemic interdependence; (c) emotional responses of responsibility and generativity toward collective well-being; and (d) behavioural tendencies to

contribute to shared infrastructure, norms, and institutions even when personal return is indirect. Sample items: "I feel a sense of responsibility for the health of the communities I belong to"; "When I think about my career, I consider not just what I'll earn but what I'll contribute"; "Maintaining shared resources matters to me even when I don't use them personally."

Factor 3: Ecological Connection (EC). This factor captures the disposition to process the natural world as a system of which one is a constitutive part, and to experience its flourishing as relevant to one's own. High EC is characterized by: (a) attention to natural systems and one's embeddedness within them; (b) interpretive habits that recognize ecological interdependence; (c) emotional responses of connection, awe, and care toward the more-than-human world; and (d) behavioural tendencies toward ecological stewardship. Sample items: "I feel a genuine sense of belonging when I spend time in nature"; "The health of ecosystems feels personally relevant to me, not just an abstract concern"; "My sense of what a good life includes is shaped by my relationship with the natural world."

Factor 4: Temporal Depth (TD). This factor captures the disposition to process present decisions in light of their consequences for future systems—both future generations of humans and the ongoing integrity of the wholes one participates in. High TD is characterized by: (a) attention to long-term consequences beyond one's own lifespan; (b) interpretive habits that consider intergenerational effects; (c) emotional investment in leaving things better than one found them; and (d) behavioural tendencies toward patience, investment, and stewardship that extend beyond personal returns. Sample items: "When I make important decisions, I consider how they will affect people not yet born"; "I feel a sense of responsibility to preserve what I've inherited for those who come after"; "Short-term gains don't feel worth it if they come at the cost of long-term damage."

The four-factor structure is hypothesized to be the best fit to the data, with a higher-order IPO factor accounting for correlations among the first-order factors. The higher-order factor represents the general disposition toward integrative processing, while the first-order factors capture domain-specific expressions. An alternative model—a single-factor structure with no domain differentiation—should be tested and compared for fit.

5.2.4. Item Development Principles

Item development for the IPO Scale should follow established psychometric principles (DeVellis, 2017; Boateng et al., 2018) while reflecting the theoretical specificity of the construct.

- Item Content. Items should capture processing dispositions—characteristic ways of attending, interpreting, feeling, and responding—rather than abstract values or self-evaluations. The stem "I..." should be followed by statements about what the respondent tends to do, feel, or think, not about what they believe they should do. This grounds the scale in psychological process rather than normative commitment, aiding discriminant validity.
- Directionality. Each factor should include both positively and negatively keyed items. Negatively keyed items capture repulsive processing tendencies: the disposition to process experience in ways

that degrade or neglect cooperative wholes. Sample reverse-keyed items: "When I'm focused on my own problems, I lose sight of how others are affected" (RI, reversed); "I don't really think about how my daily choices affect larger systems" (SC, reversed); "Nature is nice to visit, but I don't feel any deep connection to it" (EC, reversed); "The long-term future is too uncertain to worry about" (TD, reversed).

- Response Format. A 7-point Likert scale from "Strongly Disagree" to "Strongly Agree" provides adequate granularity for capturing individual differences while remaining accessible. Items should avoid extreme or absolutist wording that would compress variance.
- Readability. Items should be comprehensible to adults with a range of educational backgrounds. Technical terms like "systemic" or "ecological" should be defined implicitly by the item content rather than presumed as prior knowledge.

5.2.5. Initial Item Pool

An initial item pool of approximately 60 items (15 per factor) should be generated, drawing on the theoretical framework, existing related measures (adapted with theoretical rationale), and expert review. Table 5.1 presents sample items for each factor, illustrating the intended scope and style.

Table 5.1: Sample IPO Scale Items by Factor

Factor	Sample Integrative Items	Sample Repulsive Items (Reverse-Scored
Relational Integration	"When I'm in conflict with someone I care about, I look for what we both need, not just who's right."	"I often find myself pulling away from people when things get difficult between us."
Relational Integration	"I naturally notice when someone I'm close to seems off, even if they haven't said anything."	"Other people's emotions are their business, not mine."
Systemic Contribution	"I think about what my community needs, not just what I want	"I don't see why I should contribute to things that don't directly benefit me."
Systemic Contribution	"I feel a sense of ownership for the groups I belong to—their problems are my problems."	"Institutions and organizations will take care of themselves; I focus on my own life
Ecological Connection	"I feel most myself when I'm connected to the natural world." |	"I don't understand why people get emotional about environmental issues."
Ecological Connection	"Being in nature reminds me of what really matters."	"My well-being is mostly separate from what happens to the environment."
Temporal Depth	"I try to act in ways that will make sense to my grandchildren."	"The future will sort itself out; I focus on the present."
Temporal Depth	"I feel accountable to the generations that will come after me."	"I don't owe anything to people who don't exist yet."

The full item pool would then undergo content validity assessment by expert raters familiar with the theoretical framework and with scale development methodology. Items rated as ambiguous, poorly

representing the intended construct, or excessively overlapping with established measures (particularly agreeableness and social desirability) would be revised or eliminated.

5.3. Validation Strategy

5.3.1. Phase 1: Exploratory Factor Analysis and Item Reduction

An initial administration of the 60-item pool to a large, demographically diverse sample (N ≥ 500) would permit exploratory factor analysis (EFA) to identify the underlying factor structure. Parallel analysis and Velicer's MAP test would guide factor retention. Items with low primary loadings (< .40), high cross-loadings (> .30 on multiple factors), or low communalities (< .30) would be candidates for elimination. The goal would be a reduced scale of approximately 24–32 items (6–8 per factor) with clean simple structure.

The hypothesized four-factor structure should emerge from EFA, with factors corresponding to Relational Integration, Systemic Contribution, Ecological Connection, and Temporal Depth. If a different factor structure emerges—for instance, if Relational Integration and Systemic Contribution collapse into a single interpersonal factor, or if Ecological Connection and Temporal Depth form a single transpersonal factor—the theoretical framework would require revision. Such an outcome would not necessarily falsify the broader theory but would require reconceptualizing how integrative processing differentiates across domains.

5.3.2. Phase 2: Confirmatory Factor Analysis

A second, independent sample (N ≥ 500) would permit confirmatory factor analysis (CFA) testing several competing models:

- Model 1: Single-factor (all items loading on a general IPO factor)

- Model 2: Four correlated factors (as hypothesized)

- Model 3: Four first-order factors with a higher-order IPO factor

- Model 4: Bifactor model (general IPO factor plus four specific factors)

Fit indices (CFI ≥ .90, RMSEA ≤ .08, SRMR ≤ .08) would be evaluated for each model. The theoretical framework predicts that Model 3 (higher-order) or Model 4 (bifactor) will provide the best fit, consistent with both a general disposition and domain-specific expressions. The bifactor model, if well-fitting, would permit the calculation of both a general IPO score and specific factor scores, maximizing flexibility for research applications.

5.3.3. Phase 3: Convergent and Discriminant Validity

Convergent Validity. The IPO Scale should demonstrate significant positive correlations with theoretically related constructs:

- Prosociality measures: Social Connectedness Scale (Lee & Robbins, 1995), Inclusion of Other in Self (Aron, Aron, & Smollan, 1992), Generativity Scale (McAdams & de St. Aubin, 1992)

- Purpose measures: Purpose in Life subscale of Ryff's Psychological Well-Being Scales (Ryff, 1989), Meaning in Life Questionnaire (Steger et al., 2006), with stronger correlations expected for contribution-oriented purpose than self-oriented purpose

- Nature-connectedness measures: Nature Relatedness Scale (Nisbet, Zelenski, & Murphy, 2009), Connectedness to Nature Scale (Mayer & Frantz, 2004)

- Empathy measures: Interpersonal Reactivity Index (Davis, 1983), particularly Perspective-Taking and Empathic Concern subscales

- Light Triad Scale (Kaufman et al., 2019)

Convergent correlations should be substantial (r = .40–.70) but not so high as to suggest redundancy (r > .80), which would indicate that IPO is not a distinct construct.

Discriminant Validity. The IPO Scale should demonstrate only modest correlations with constructs from which it is theoretically distinct:

- Agreeableness (IPIP-NEO or BFI-2): Expected r = .30–.50. IPO should explain well-being variance beyond agreeableness.

- Social Desirability (Marlowe-Crowne or BIDR): Expected r = .20–.40. IPO should predict outcomes independently of socially desirable responding.

- Neuroticism: Expected r = -.30 to -.50. Some shared variance is expected (repulsive processing correlates with emotional instability), but IPO should not be reducible to low neuroticism.

- Political ideology (liberal-conservative continuum): Expected r = near zero or modest. The framework's emphasis on integration spans domains that different ideologies value differently, and repulsive processing can occur across the political spectrum. This is an important discriminant test: if IPO correlates strongly with liberalism, it may be capturing political values rather than a more fundamental processing disposition.

Discriminant validity should be demonstrated through hierarchical regression analyses showing that IPO predicts well-being and psychopathology outcomes incrementally beyond these established measures. If IPO adds no incremental variance, the construct is redundant with existing measures and the theoretical framework loses empirical support.

5.3.4. Phase 4: Criterion Validity

Prediction of Well-Being. The IPO Scale should predict multiple dimensions of psychological well-being, assessed through:

- Ryff's Psychological Well-Being Scales (Ryff, 1989): autonomy, environmental mastery, personal growth, positive relations with others, purpose in life, self-acceptance

- Satisfaction With Life Scale (Diener et al., 1985)

- Positive and Negative Affect Schedule (Watson, Clark, & Tellegen, 1988)

- Flourishing Scale (Diener et al., 2010)

The framework predicts that IPO will correlate more strongly with eudaimonic well-being dimensions (purpose, personal growth, positive relations) than with hedonic dimensions (positive affect), though both should show associations. This pattern would support the claim that IPO reflects alignment with objective conditions of flourishing rather than merely pleasant feeling.

Prediction of Psychopathology. The IPO Scale should correlate negatively with measures of psychopathology, particularly those characterized by relational disruption:

- Depression: Beck Depression Inventory (BDI-II) or PHQ-9. The emotional capture analysis (Chapter 3, §3.8) predicts that low IPO (high repulsive processing) will be associated with depression, and that this association will be mediated by social withdrawal and rumination.

- Anxiety: Beck Anxiety Inventory (BAI) or GAD-7. The framework predicts that anxiety's association with low IPO operates through threat-biased avoidance that degrades integrative participation.

- Personality disorder features: PID-5 (Krueger et al., 2012), particularly detachment and antagonism domains.

- Substance use disorders: measures of addictive behaviour severity. The framework predicts that addictive processing represents a form of repulsive processing in which the substance becomes a "repulsive attractor" pulling the individual out of orbit around integrative wholes.

Clinical vs. Non-Clinical Comparisons. The framework predicts significantly lower IPO scores in clinical populations compared to matched controls. Effect sizes should be substantial ($d \geq 0.50$) for conditions characterized by relational disruption. Pre-post therapy designs should show increases in IPO accompanying effective treatment, and these increases should mediate improvements in symptom measures. If symptom improvement occurs without corresponding IPO change, the framework's claim that IPO is a mechanism of health rather than merely a correlate would be challenged.

5.3.5. Phase 5: Incremental Validity and Mediation

Incremental Validity. Hierarchical regression analyses should demonstrate that IPO predicts well-being and psychopathology outcomes incrementally beyond:

- Big Five personality traits

- Social desirability

- Existing prosociality measures (empathy, agreeableness, Light Triad)

- Demographics (age, gender, SES, education)

If IPO fails to contribute incremental variance, the construct does not earn its place in the nomological network.

Mediation Analyses. The framework proposes that IPO is the mechanism through which various positive psychological factors exert their effects on well-being. Specifically:

- The relationship between social connectedness and well-being should be partially mediated by IPO. Mere social contact is insufficient; contact that reflects and reinforces integrative processing is what matters.

- The relationship between nature-connectedness and well-being should be partially mediated by IPO. Nature connection's benefits arise not merely from stress reduction but from aligning the self with a larger integrative whole.

- The relationship between purpose and well-being should be partially mediated by IPO, with contribution-oriented purpose showing stronger mediation than self-oriented purpose.

Conversely, the framework proposes that IPO protects against psychopathology in part by interrupting emotional capture loops. Low IPO should moderate the relationship between stress and negative outcomes: individuals low in IPO should show stronger stress-reactivity because their repulsive processing amplifies rather than buffers the effects of adversity. Longitudinal designs testing this moderation hypothesis would provide strong evidence for the framework's causal claims.

5.4. Falsification Conditions: What Would Count Against the Framework

A theory that cannot be falsified is not a scientific theory. The systemic alignment framework makes substantive claims about the structure of human flourishing and the nature of mental health, and these claims must be vulnerable to empirical disconfirmation. This section specifies, in advance of the empirical program outlined above, the patterns of results that would constitute falsification of the framework's core claims, as distinct from results that would require modification or refinement. Drawing

this distinction explicitly addresses the concern, anticipated in Chapter 6, that the framework might be insulated from refutation by ad hoc auxiliary hypotheses.

5.4.1. The Falsification Principle

The framework's core empirical claim is that Integrative Processing Orientation—the disposition to process experience in ways that sustain and generate cooperative wholes—is the psychological manifestation of mental health, such that variations in IPO systematically covary with variations in psychological well-being and psychopathology. This claim generates the predictions articulated in section5.1 and the validation expectations specified in section 5.3. The falsification question is: what pattern of results, if obtained across multiple rigorous, well-powered, independent studies, would require abandoning this core claim?

A responsible answer must distinguish between results that would falsify the framework and results that would require its modification. Frameworks typically evolve in response to anomalous findings; the question is whether the anomaly can be accommodated by refining auxiliary hypotheses or whether it strikes at the framework's central commitments. The following specifies both.

5.4.2. Results That Would Falsify the Framework

The framework would be abandoned if, after a decade of programmatic research including scale validation in multiple samples, longitudinal designs, clinical comparisons, and independent replication by multiple laboratories, the following pattern of results obtained:

F1: Complete failure of the IPO factor structure. If, across multiple large, demographically diverse samples, exploratory and confirmatory factor analyses consistently fail to identify any coherent latent factor uniting the proposed domains of Relational Integration, Systemic Contribution, Ecological Connection, and Temporal Depth—if these domains prove to be entirely independent dimensions with no shared variance attributable to a general IPO factor—the framework's central construct would lack empirical support. This would not merely require modification (e.g., dropping one factor); it would undermine the claim that there exists a general disposition toward integrative processing that manifests across domains. The framework's theoretical architecture depends on IPO being a coherent psychological reality, not merely a conceptual abstraction. If the domains are empirically unrelated, the framework's claim that mental health consists in a unified orientation toward integration is false.

F2: Zero or negative association between IPO and well-being. If meta-analyses of the relationship between validated IPO measures and multiple indices of psychological well-being (both hedonic and eudaimonic) yield effect sizes indistinguishable from zero ($r < .05$) or significantly negative, the framework's central prediction is disconfirmed. The framework claims that integrative processing constitutes mental health; if IPO scores are unrelated or inversely related to well-being, this claim is false. This finding would be particularly damning if obtained while established predictors (social support,

purpose, nature-connectedness) continue to show positive associations with well-being, indicating that the null result is not an artifact of measurement or sample characteristics.

F3: IPO fully redundant with existing constructs. If hierarchical regression analyses consistently show that IPO adds no incremental variance in predicting well-being or psychopathology beyond established personality traits (particularly agreeableness, neuroticism, and conscientiousness) and existing prosociality measures (empathy, Light Triad, social connectedness)—if the ΔR^2 attributable to IPO is consistently non-significant across multiple studies and outcomes—the construct does not earn its place. The framework claims to identify a distinctive psychological disposition that existing measures do not capture; if IPO is empirically indistinguishable from being agreeable and socially connected, the theoretical framework is otiose.

F4: Clinical populations indistinguishable from controls on IPO. If validated IPO measures show no significant differences between clinical populations (particularly those characterized by the relational disruption the framework emphasizes: major depressive disorder, social anxiety disorder, borderline personality disorder, substance use disorders) and matched non-clinical controls ($d < 0.15$ across multiple comparisons), the framework's account of mental illness as repulsive processing is disconfirmed. The framework predicts substantial differences; their absence would indicate that whatever underlies psychopathology, it is not the degradation of integrative processing.

F5: Therapeutic improvement without IPO change. If longitudinal clinical studies consistently demonstrate that effective treatments (CBT, ACT, psychodynamic therapy, MBIs) produce significant symptom reduction and functional improvement without corresponding increases in IPO—if pre-post IPO changes are negligible while well-being and symptom measures show large effects—the framework's claim that IPO is a mechanism of therapeutic change is disconfirmed. The framework predicts that effective therapy works in part by shifting processing from repulsive to integrative patterns; if this shift does not occur when therapy is effective, the mechanism is not operative.

F6: IPO does not mediate established relationships. If formal mediation analyses consistently fail to show that IPO mediates the relationships between social connectedness, purpose, nature-connectedness, and well-being—if the indirect effects through IPO are non-significant while direct effects remain robust—the framework's claim to provide a unifying mechanism is disconfirmed. The framework asserts that integrative processing is why these factors promote well-being; if they promote well-being through other pathways entirely, the framework's explanatory claim fails.

F7: Nature-connectedness adds no unique variance. This is a distinctive prediction of the framework that differentiates it from standard social-psychological accounts. If nature-connectedness measures show no unique association with well-being after controlling for social connectedness and other interpersonal variables—if the partial correlation is zero—a specific and risky prediction of the framework is disconfirmed. While this alone might not falsify the entire framework (the other predictions could hold), it would remove one of the framework's most distinctive empirical signatures and substantially reduce its comparative advantage over theories that treat well-being as exclusively social.

5.4.3. Results That Would Require Modification

Not all disconfirmatory findings would falsify the framework. Some patterns of results would require modification of auxiliary hypotheses while leaving the core theoretical commitments intact. Distinguishing these from falsifying results is essential for honest theory development.

M1: Factor structure diverges from the hypothesized four-factor model. If EFA and CFA consistently identify a different factor structure—for example, two factors (Interpersonal Integration and Transpersonal Integration) rather than four, or five factors with Temporal Depth splitting into separate present-future and intergenerational dimensions—this would require reconceptualizing the structure of integrative processing but would not falsify the claim that such processing exists and matters. The four-factor model is a theoretical proposal; the empirical structure may be different while still reflecting a general disposition toward integration.

M2: IPO correlates more strongly with some outcomes than predicted. The framework predicts stronger associations with eudaimonic than hedonic well-being. If IPO proves equally predictive of both, or more strongly predictive of hedonic well-being, this would require revising the claim that IPO specifically reflects alignment with objective conditions of flourishing rather than general positivity. It would not falsify the framework but would suggest that integrative processing promotes well-being through both meaning-based and pleasure-based pathways.

M3: Some proposed items or factors perform poorly. If specific items from the initial pool fail to load on their intended factors, or if a particular factor (e.g., Ecological Connection) proves difficult to measure reliably, this would require psychometric refinement but not theoretical abandonment. Scale development is iterative; poor initial items reflect measurement challenges, not necessarily theoretical error.

M4: IPO shows cultural variability in expression but not in underlying structure. The framework predicts that the specific forms of integrative processing vary with cultural context, while the underlying principle is universal. If the four-factor structure replicates in some cultures but not others, or if item functioning varies across cultural groups, this would require cultural adaptation of the scale but would not falsify the universalist claim—provided that locally adapted measures of integrative processing continue to predict well-being across cultures.

M5: IPO change is a consequence rather than a cause of well-being change. The framework predicts bidirectional relationships but emphasizes IPO as a mechanism of health. If longitudinal studies consistently show that changes in well-being precede changes in IPO rather than the reverse—or if experimental manipulations of IPO (through interventions designed to increase integrative processing) fail to produce well-being improvements—the causal model would require revision. The framework might need to acknowledge that IPO is more a marker of health than a mechanism, or that the causal relationship is primarily from well-being to IPO. This would weaken the framework's practical implications but would not falsify its central claim that IPO and health are intimately linked.

5.4.4. The Burden of Proof

The framework's proponents bear the burden of demonstrating that its distinctive predictions survive rigorous testing. The falsification conditions specified above are not merely rhetorical hedges; they are commitments. If the empirical program outlined in this chapter is undertaken and the results consistently align with the falsification pattern (F1–F7) rather than the modification pattern (M1–M5), the framework should be abandoned. If the results align with the modification pattern, the framework should be refined. If the results align with the framework's positive predictions—a coherent IPO factor structure, substantial associations with well-being, incremental validity beyond established measures, clinical differences, mediation of established relationships, and unique variance from nature-connectedness—the framework earns empirical support.

What cannot be responsibly claimed is that the framework is immune to empirical challenge. The conditions specified here ensure that it is not.

5.5. How the Framework Reframes Existing Therapies

Beyond generating novel predictions and a new measurement instrument, the systemic alignment framework provides a unifying theoretical lens through which existing evidence-based therapies can be understood. This reframing does not claim that therapists must adopt the framework's terminology to practice effectively; rather, it shows that the mechanisms underlying diverse therapeutic approaches converge on the promotion of integrative processing—whether or not this convergence is explicitly recognized. The emotional capture analysis of Chapter 3 (3.8) provides the theoretical basis for this convergence: diverse therapies work, on this account, because they all interrupt the meaning-capture loops in which repulsive processing becomes self-reinforcing.

Cognitive-Behavioural Therapy (CBT). CBT identifies and modifies dysfunctional thoughts and behaviours. The systemic alignment framework provides a criterion for what makes a thought or behaviour dysfunctional: it degrades the individual's capacity to participate in and contribute to integrative wholes. A thought is not problematic merely because it is negative—negative thoughts can be accurate and adaptive when they guide corrective action within relationships. Rather, a thought is dysfunctional when it functions repulsively, pushing the individual away from the connections upon which flourishing depends. The cognitive distortion of "mind reading" (assuming one knows what others are thinking) is problematic not because it is a logical error but because it substitutes a repulsive fantasy about others' judgments for the integrative work of actual communication. Behavioural activation, one of CBT's most potent components for depression, works by interrupting the repulsive spiral of withdrawal—the emotional capture loop in which low mood produces inactivity, which produces further low mood—and re-engaging the individual in activities that restore connection to rewarding, often relational, contexts. The framework thus explains why behavioural activation works: it directly counteracts repulsive processing by engineering integrative engagement, breaking the self-reinforcing cycle that emotional capture theory predicts.

Acceptance and Commitment Therapy (ACT). ACT describes psychological flexibility as the capacity to contact the present moment and act on chosen values. The systemic alignment framework reframes this as the operational signature of a decision-making system shifting from repulsive to integrative processing. The six ACT processes are thereby reorganized around a single functional axis: diffusion and acceptance interrupt the repulsive loop in which avoidance of private experience severs relational connection; self-as-context establishes the perspective from which the individual is recognized as a node embedded within larger systems rather than a bounded self besieged by symptoms; values clarification and committed action are the explicit reorientation of the decision-making system toward the generation of integrative wholes—family, community, ecosystem—that the framework identifies as the objective ground of human flourishing. The emotional capture analysis illuminates why diffusion is therapeutically central: it creates the gap between emotional activation and conceptual elaboration that allows choice to enter where automatic repulsive processing previously operated. What ACT presents as a pragmatic clinical model, the framework reinterprets as the practical working-out of a teleological truth: that mental health consists in processing that builds connection because human beings are constitutively relational creatures whose good is inseparable from the health of the nested systems they inhabit.

Psychodynamic Therapies. Psychodynamic approaches trace current dysfunctional patterns to early relational experiences, particularly those involving attachment disruption or trauma. The systemic alignment framework provides a language for understanding how early repulsive environments—those characterized by neglect, abuse, or inconsistent caregiving—shape a decision-making system organized around defensive disconnection. The "repulsive adaptations" that once protected the child in an unsafe environment become self-defeating in adulthood, precisely because they systematically degrade the relational connections that adult flourishing requires. The therapeutic relationship, from this perspective, succeeds by providing a relational context in which integrative processing can be experienced and internalized. The therapist's consistent, empathic presence creates a new kind of relational experience that gradually rewrites the old repulsive programming. Transference interpretations work by making the patient aware of how old emotional capture patterns—"this person will betray me as my caregiver did"—are being imposed on current relationships, opening the possibility of choosing differently. The framework's account of how repulsive processing becomes self-reinforcing through emotional capture provides a mechanistic complement to the psychodynamic concept of repetition compulsion: what is repeated is not merely a behavioural pattern but a meaning-capture loop in which early emotional learning continues to shape interpretation in ways that prevent new learning.

Humanistic and Existential Therapies. Humanistic approaches emphasize self-actualization, authentic relating, and the inherent growth tendency of the organism. The systemic alignment framework provides a theoretical grounding for these values that goes beyond mere assertion. Self-actualization is not arbitrary self-expression but the realization of one's nature as a constitutively relational being whose flourishing depends on integrative participation in larger systems. Carl Rogers' "core conditions" of empathy, unconditional positive regard, and congruence are not merely facilitative techniques; they are integrative acts offered by the therapist that model and invite corresponding integration in the client. Existential therapy's confrontation with meaning and mortality can be understood as an invitation to

locate one's purpose within systems that transcend the individual lifespan—a conceptual expansion of the self that occurs as the individual grows to treat larger wholes as functionally as important as their individual organism. The existential emphasis on authentic choice aligns with the framework's recognition that autonomy is valuable precisely because it enables chosen integration rather than coerced compliance.

Family Systems and Couples Therapy. These modalities perhaps most transparently align with the systemic framework, as they already conceptualize problems in terms of relational patterns rather than individual pathology. The framework enriches this perspective by connecting it to the broader architecture described in Chapter 2: the family is one level in the nested hierarchy of integrative wholes, and family dysfunction represents a breakdown of cooperative dynamics at that level. Therapeutic interventions that restructure family communication patterns, clarify boundaries, and foster mutual support are directly promoting the conditions for family-level health as defined by the criteria of mutual benefit, emergent capacity, and dynamic stability. The framework also suggests that family health cannot be fully understood in isolation from the larger systems—community, culture, biosphere—in which families are embedded, providing a theoretical rationale for expanding systemic interventions to address these broader contexts.

Mindfulness-Based Interventions (MBIs). MBIs are conventionally understood as techniques for reducing rumination and emotional reactivity through non-judgmental present-moment awareness. The systemic alignment framework reframes mindfulness practice as a systematic training in suspending emotional capture—the automatic decision-making outputs that contract experience, amplify self-referential loops, and sever the individual from direct contact with the larger wholes they inhabit. The core MBI instruction to observe thoughts and sensations without fusion or avoidance is, in this light, the deliberate cultivation of a processing mode that does not push away, grasp at, or elaborate upon mental content; it thus creates the internal conditions for integrative perception, wherein the practitioner can register their actual embeddedness in the present moment, the body, and the social and ecological systems that sustain them. The framework therefore repositions the well-documented benefits of MBIs—reduced anxiety, improved affect regulation, enhanced well-being—not as primary outcomes but as downstream consequences of a fundamental shift: the quieting of repulsive dynamics and the corresponding restoration of the mind's natural tendency toward integrative functioning. This reframing generates a testable prediction: changes in IPO should mediate the relationship between mindfulness practice and well-being outcomes.

5.6. Practical Implications for Intervention Design

If the systemic alignment framework is correct, several implications follow for how mental health interventions might be designed or enhanced.

Common Factor Integration. The framework provides a theoretical account of the "common factors" that research consistently identifies as driving therapeutic change—the therapeutic alliance, empathy, and the provision of a credible rationale. These factors are all integrative in nature: they involve the

formation of a cooperative relationship (the alliance), the cognitive bridging between self and other (empathy), and the provision of a meaning framework that connects individual suffering to a larger understanding. The framework suggests that these common factors are not merely non-specific "placebo" elements but are direct expressions of the integrative processing that constitutes mental health. A therapeutic relationship characterized by mutual benefit, emergent capacity (the dyad can achieve what the individual cannot), and dynamic stability is itself an integrative whole, and participation in it may be curative regardless of specific technique.

Multi-Level Intervention Targeting. Because mental health is a function of processing at multiple systemic levels—biological, psychological, social, and ecological—effective interventions should ideally address multiple levels simultaneously. A treatment for depression, for instance, might combine biological interventions (medication or exercise to support neural function) with psychological interventions (cognitive restructuring to shift repulsive thought patterns) and social interventions (structured opportunities for prosocial contribution to rebuild integrative connections) and ecological interventions (nature contact to restore connection to the biospheric whole). The framework provides a rationale for such multi-level approaches that goes beyond mere eclecticism.

Emphasis on Contribution, Not Just Coping. Many existing therapies focus heavily on symptom reduction and coping skill development. While these are valuable, the systemic alignment framework suggests that the deepest well-being comes from active contribution to larger wholes—from being a functional, generative part of families, communities, and ecosystems. Interventions might therefore place greater emphasis on helping individuals identify and engage in activities that allow them to contribute to systems they value, whether through volunteering, caregiving, creative production, or environmental stewardship. This is not mere altruism training; it is enabling individuals to experience themselves as functional parts of larger wholes, which the framework argues is constitutive of flourishing.

Nature-Connectedness as a Therapeutic Resource. The prediction that nature-connectedness contributes independently to well-being suggests that interventions incorporating structured contact with nature—not merely as a pleasant backdrop but as an opportunity to cultivate felt relationship with the living world—may have benefits beyond stress reduction. Ecotherapy approaches that explicitly foster nature connection may be understood as working to align the individual with one of the most encompassing integrative wholes—the biosphere—and thereby to expand the self beyond its contracted, symptom-focused boundaries.

Prevention Through Systemic Design. The framework's implications extend beyond clinical treatment to prevention. If mental health depends on opportunities for integrative participation, then social structures that foster such participation—walkable communities that encourage informal interaction, workplaces structured around cooperative missions, educational systems that teach relational skills alongside academic content, urban planning that integrates natural spaces—are not merely amenities but are public health interventions. The framework thus connects individual mental health to broader questions of social organization, providing a theoretical rationale for policies that might otherwise seem merely aesthetic or communitarian in their appeal.

5.7. Limitations and Future Directions

A responsible presentation of the framework's empirical implications requires acknowledging its limitations and identifying areas where further theoretical and empirical work is needed.

Measurement Challenges. The IPO Scale proposed here requires full development and validation. The item pool, factor structure, and validation strategy articulated in this chapter are conceptual specifications, not yet empirical demonstrations. The scale's ultimate factor structure may diverge from the hypothesized four-factor model in ways that require theoretical revision. The relationship between self-reported IPO and behavioural measures of integrative processing—observation of actual relationship-sustaining behaviour, ecological momentary assessment of processing patterns, informant reports from close others—requires investigation. Self-report measures are vulnerable to biases that the construct itself suggests may be systematically present: individuals low in IPO may lack the self-awareness to recognize their repulsive processing patterns. Multi-method assessment, including performance-based measures, implicit association tests, and informant reports, would strengthen the empirical base.

Cultural Variability. The specific forms that integrative and repulsive processing take will vary with cultural context, as acknowledged in Chapter 4. What constitutes relationship-sustaining behaviour in one cultural setting may differ from another. The framework predicts that the underlying principle—health as alignment with the cooperative dynamics of the wholes one inhabits—is universal, but the operationalization of that principle in the IPO Scale will require cultural adaptation. Items that effectively capture Ecological Connection in a Western, educated, industrialized context may require revision in contexts where relationship to nature is structured differently. The scale should be developed and validated in multiple cultural contexts, with particular attention to whether the factor structure generalizes across cultural boundaries.

Causal Complexity. The framework posits bidirectional relationships: integrative processing promotes well-being, and well-being facilitates integrative processing. Disentangling these causal pathways requires longitudinal designs, experience sampling, and experimental manipulations that may be challenging to implement. The relationship between individual and systemic health is recursive—healthy individuals build healthy communities, which in turn support individual health—and demonstrating these feedback dynamics requires sophisticated modelling (e.g., multilevel structural equation modelling, dynamic systems approaches) that captures cross-level interactions over time.

Boundary Conditions. A central task for future work is specifying the framework's boundary conditions with greater precision. As noted in Chapter 4 and elaborated in Chapter 6, the theory must clarify when alignment becomes pathological—as in enmeshment, co-dependence, or destructive loyalty to toxic systems—and when repulsion is adaptive, as in the necessary withdrawal from abusive relationships or oppressive communities. The IPO Scale must be constructed to avoid scoring enmeshment as high integration and healthy differentiation as repulsion. Items should emphasize mutual benefit and dynamic

stability rather than mere closeness, and the validation strategy should demonstrate that IPO does not correlate with measures of unhealthy fusion or lack of self-differentiation.

Alternative Explanations. The predicted relationships between IPO, prosocial integration, purpose, nature-connectedness, and well-being are consistent with other theoretical frameworks, including self-determination theory (autonomy, competence, relatedness), attachment theory, and evolutionary psychological accounts. The systemic alignment framework's distinctive contribution is its claim that these relationships are expressions of a deeper pattern—the recurrent structural isomorphism described in Chapter 2—and that they are mediated by a general disposition toward integrative processing that spans interpersonal, systemic, ecological, and temporal domains. Demonstrating this requires showing (a) that IPO is a coherent latent factor across these domains, (b) that it predicts well-being incrementally beyond established related constructs, and (c) that it mediates the effects of domain-specific predictors (social connection, purpose, nature-connectedness) on well-being. These demonstrations await the empirical program outlined above.

5.8. Conclusion

The systemic alignment theory of mental health is not an untestable metaphysical speculation; it is a framework that generates specific, falsifiable predictions about the correlates and causes of psychological well-being. The central prediction—that individuals whose decision-making systems reliably produce integrative outputs will demonstrate higher well-being than those whose processing is predominantly repulsive—can be operationalized through the Integrative Processing Orientation Scale proposed in this chapter and tested through correlational, longitudinal, clinical, and experimental designs.

The IPO construct, with its four hypothesized factors (Relational Integration, Systemic Contribution, Ecological Connection, and Temporal Depth) and higher-order structure, provides a concrete empirical target for testing the framework's claims. Its proposed validation strategy—including exploratory and confirmatory factor analysis, convergent and discriminant validity assessment, and tests of incremental and criterion validity—provides a clear program for empirical investigation. Should the IPO Scale perform as predicted, it would constitute significant evidence for the framework.

Crucially, the framework has specified the conditions under which it would be abandoned. If, after rigorous programmatic research, the IPO factor structure fails to cohere, IPO shows no association with well-being, IPO proves redundant with existing measures, clinical populations show no IPO deficits, therapeutic change occurs without IPO change, IPO fails to mediate established relationships, and nature-connectedness adds no unique variance—the framework should be abandoned. If, instead, the results suggest modifications to the factor structure, the causal model, or the measurement approach while preserving the core claim that integrative processing is central to mental health—the framework should be refined. These are not empty gestures; they are commitments that make the framework genuinely testable.

Moreover, the framework demonstrates theoretical utility by providing a unified lens through which to understand the mechanisms of existing evidence-based therapies. Cognitive-behavioural, acceptance-based, psychodynamic, humanistic, systemic, and mindfulness approaches all, from this perspective, promote mental health by fostering integrative processing—whether by restructuring repulsive cognitions, building psychological flexibility, providing corrective relational experiences, or strengthening the observer capacity that enables conscious choice. The emotional capture analysis of Chapter 3 provides the mechanistic account of why these diverse approaches converge: all interrupt the meaning-capture loops in which repulsive processing becomes self-reinforcing.

The value of this framework, should its predictions be confirmed and its falsification conditions not be met, lies in its capacity to connect the phenomenology of mental health and illness to the broader architecture of a relational universe. It offers an account in which human psychological functioning is not an isolated domain governed by arbitrary norms but is continuous with the organizational principles observable across the sciences. Such an account, if empirically supported, would not only advance psychological science but would also provide individuals with a deeply grounded understanding of what it means to be mentally healthy and why it matters. The work of empirical validation—of developing the IPO Scale, testing its predictions, and subjecting the framework to the genuine risk of falsification—remains to be done. But the conceptual specification provided here, together with the explicit falsification conditions that distinguish scientific commitment from metaphysical immunization, is the necessary first step.

Chapter 6: Objections and Replies

6.1. Introduction

The preceding chapters have advanced a systematic argument: the normativity required for a scientifically robust definition of mental health can be objectively grounded in the teleological structure of living systems, without appeal to cultural consensus or purely evolutionary criteria. Mental health, on this account, consists in systemic alignment—the reliable tendency of the decision-making system to produce reactions that sustain and generate integrative wholes (Chapter 4). Mental illness, correspondingly, is repulsive processing—a pattern of decision-making that degrades the cooperative connections upon which both individual and collective flourishing depend, often maintained through the emotional capture dynamics described in Chapter 3 (3.8). The framework generates specific, falsifiable predictions operationalized through the Integrative Processing Orientation construct (Chapter 5).

Any framework with such ambitions invites vigorous challenge. This chapter anticipates and responds to the most serious objections that might be raised against the theory. These objections span six domains: (1) the naturalistic fallacy and the is-ought gap; (2) the problem of identifying objective teleology in a post-Darwinian universe; (3) the risk of reductionism and the neglect of human autonomy; (4) the framework's potential for pathologizing difference and justifying oppression; (5) internal tensions and boundary-condition problems; and (6) questions about empirical testability and parsimony. Before addressing these objections, however, the chapter establishes that the framework earns its place within a crowded theoretical landscape. Section 6.2 examines three influential alternative approaches to the normative foundations of mental health—Fulford's values-based practice, Bolton and Banner's deliberative framework, and Boorse's biostatistical theory—and demonstrates why each, despite its contributions, leaves the normative crisis unresolved. Each objection in the subsequent sections is stated in its strongest form before a reply is developed.

The chapter then presents a detailed case study—the experience of gender dysphoria in a non-affirming environment—demonstrating step by step how the framework's criteria yield a determinate, non-arbitrary judgment that is superior to both the harmful dysfunction analysis and pure cultural consensus. The aim is not to claim that the framework is invulnerable to criticism—no philosophical theory of mental health has ever achieved such a status—but to demonstrate that it can withstand the most pressing challenges while retaining its distinctive commitments and providing practical guidance in contested cases.

6.2. Alternative Frameworks and Why They Are Insufficient

The framework advanced in this paper is not proposed in a vacuum. Several sophisticated alternatives already attempt to address the normative foundations of mental health. Demonstrating that the present framework earns its place requires showing that these alternatives, despite their genuine contributions, leave the normative crisis unresolved in ways that the systemic alignment framework does not. This

section examines three prominent positions: Fulford's values-based practice, Bolton and Banner's deliberative model, and Boorse's biostatistical theory.

6.2.1. Fulford's Values-Based Practice

K.W.M. Fulford's values-based practice (VBP) represents one of the most sustained attempts to address the role of values in psychiatric classification. Writing from within the analytic philosophy tradition, Fulford argues that the failure of earlier attempts to define mental disorder in purely factual terms reveals something fundamental about the concepts themselves: "illness" and "disorder" are irreducibly value-laden concepts. This is not a defect to be eliminated but a feature to be managed. VBP accordingly provides a clinical framework for negotiating diverse values in psychiatric practice, emphasizing that values are not merely patient preferences but can be held by clinicians, families, and communities. The framework draws on philosophical work by R.M. Hare and J.L. Austin to argue that medical concepts function as "action-guiding" terms whose meaning cannot be exhausted by descriptive criteria.

Fulford's alternative to the harmful dysfunction analysis is a "reverse" strategy: rather than attempting to constrain values with an objective dysfunction criterion, VBP acknowledges that values are ineliminable and develops procedures for handling value diversity transparently. The approach emphasizes mutual respect, careful attention to the values actually in play in clinical encounters, and training clinicians to recognize when disagreements arise from value differences rather than factual disputes. VBP has been influential in UK mental health policy and has generated practical training programs.

Why it is insufficient. VBP's great strength—its honest acknowledgment of value pluralism—is also its limitation. By treating values as given starting points to be negotiated rather than as potentially answerable to objective features of human flourishing, VBP provides no basis for distinguishing between values that genuinely promote mental health and values that undermine it. In the language of the present framework, VBP cannot distinguish between genuinely integrative wholes and pseudo-wholes. A family system that values conformity over authentic self-expression would, under VBP, have its values treated as one legitimate perspective to be negotiated alongside the patient's. But this is precisely the problem: some values are incompatible with human flourishing, and a framework that treats all values as negotiable starting points cannot say why.

Consider the gender dysphoria case analysed in detail below (6.9). In a non-affirming community, the family holds values that reject Alex's gender identity; Alex holds values that affirm it. VBP can facilitate negotiation between these perspectives, but it cannot say—except by importing further values—that the family's values are contributing to a pseudo-integration that undermines Alex's flourishing. The framework's procedural virtues become substantive liabilities when the values in conflict include values that are themselves pathogenic. The systemic alignment framework, by contrast, provides criteria for identifying which values promote genuine integration and which sustain pseudo-wholes at the expense of their members' flourishing. This is not a return to authoritarian value imposition; it is a recognition that values, like beliefs, can be more or less adequate to the realities of human life.

6.2.2. Bolton and Banner's Deliberative Framework

Derek Bolton and Natalie Banner (2012, and subsequent work) have developed a nuanced position that shares ground with both Fulford and Wakefield. They argue that the concept of mental disorder inherently involves both factual and normative elements, and that the task is not to eliminate normativity but to ensure that normative judgments are made through appropriate deliberative procedures. Drawing on deliberative democracy theory and the capabilities approach, they propose that decisions about diagnostic boundaries should be informed by empirical evidence but ultimately made through inclusive deliberation that gives voice to all affected parties—patients, clinicians, families, and communities.

Bolton and Banner's framework improves on VBP in two respects. First, it acknowledges that not all values are equal: deliberative procedures can, in principle, identify which value judgments are better supported by evidence and reasoning. Second, it connects psychiatric classification to broader questions of human well-being and social justice, recognizing that diagnostic decisions have profound consequences for people's lives. Their work represents a serious attempt to navigate between the Scylla of value-free scientism and the Charybdis of untethered cultural relativism.

Why it is insufficient. The deliberative framework's primary limitation is that its proceduralism cannot supply the substantive criteria needed to resolve hard cases. Deliberation can refine, challenge, and adjudicate value judgments, but only if participants share some common understanding of what human well-being consists in. When that common understanding is precisely what is contested, deliberation may simply reproduce the conflict at a more articulate level. The gender dysphoria case illustrates the difficulty: inclusive deliberation among Alex, Alex's non-affirming family, affirming clinicians, and community members with divergent views does not guarantee a determinate outcome, because the participants disagree not merely about the application of shared values but about the nature of the good for a person in Alex's situation. The deliberative framework provides a procedure but not a criterion.

The systemic alignment framework addresses this limitation by providing substantive, publicly debatable criteria for evaluating outcomes. The question is not merely "What do the affected parties, under ideal deliberative conditions, agree upon?" but "Which outcome enables Alex to function as a well-functioning instance of the kind of being a human is—a constitutively relational creature whose flourishing depends on authentic integration into genuine wholes?" The framework's criteria (mutual benefit, emergent capacity, dynamic stability) provide traction on this question that deliberation alone cannot. Deliberative procedures remain valuable for applying these criteria in specific cases, but they require substantive normative content to guide them—content that the systemic alignment framework supplies.

6.2.3. Boorse's Biostatistical Theory

Christopher Boorse's biostatistical theory (BST), first proposed in the 1970s and refined over subsequent decades, represents the most influential attempt to define health and disease in purely descriptive, value-free terms. On Boorse's account, health is statistical normality of function: a part or process is

healthy when it makes a statistically typical contribution to the organism's survival and reproduction, relative to a reference class defined by species, age, and sex. Disease is the absence of health—a statistically atypical functional deficit that impairs the organism's capacity relative to its reference class. The BST explicitly excludes value judgments from the definition of disease. Whether a condition is a disease is, on this view, a purely factual question; whether it is bad or warrants treatment is a separate, evaluative question.

The BST has considerable virtues. It provides a clear, operationalizable definition that aligns with much medical practice. It explains why statistical normality is relevant to health without reducing health to mere normality (the reference class specification provides the normative standard). And it is genuinely naturalistic, requiring no appeal to teleology, evolutionary function, or cultural values. If successful, the BST would solve the normative crisis by showing that the crisis is illusory: the definition of disorder is value-free; only decisions about treatment require values.

Why it is insufficient. The BST faces well-known objections, several of which are fatal in the mental health context.

First, the reference class problem. Boorse's theory requires specifying a reference class (species, age, sex) against which statistical normality is measured. But the specification of reference classes is itself value-laden. Why is "age" a legitimate reference class but "profession" or "lifestyle" not? A professional musician's hearing capacity is statistically abnormal for their age but typical for their profession. The BST must either exclude profession as a reference class (which seems arbitrary) or include it (which introduces an indefinite number of reference classes, making "disease" radically relative). This problem is particularly acute for mental health, where the relevant reference class—what it is to be a functioning human being—is precisely what is at issue. The BST assumes that species-typical functioning is the relevant standard, but this begs the question against frameworks that ground normativity in the human life-form rather than statistical distributions.

Second, the BST cannot distinguish between statistical abnormalities that are pathologies and those that are benign variations. Homosexuality was statistically atypical for Boorse's reference class but is not a pathology. High intelligence is statistically atypical but is not a disease. Boorse can respond that these are not functional deficits—they do not impair the organism's capacity for survival and reproduction. But this pushes the problem to the concept of "functional deficit." What counts as a functional deficit depends on what we take the relevant functions to be. If reproduction is the relevant function, then voluntary celibacy involves a functional deficit (a functioning reproductive system not being used for its function), but we do not classify celibacy as a disease. Boorse can refine the account to specify that diseases are internal functional deficits that reduce functional capacity below the statistically typical level. But this refinement makes the BST increasingly ad hoc and decreases its fit with actual medical practice, which treats many conditions as diseases that do not reduce survival or reproduction (e.g., many dermatological conditions, infertility in post-reproductive individuals).

Third, and most relevant to the present paper, the BST provides no positive account of mental health. Like the DSM and the harmful dysfunction analysis, the BST treats health as the absence of disease. It

cannot say what mental health positively consists in or why it matters beyond the absence of statistically atypical functional deficit. This is not merely a conceptual lacuna; it has practical consequences. Without a positive account of mental health, psychiatry cannot guide individuals toward flourishing; it can only treat dysfunction. The systemic alignment framework, by contrast, provides a positive vision of mental health as integrative processing, offering guidance not only for treatment but for cultivation.

The BST's aspiration to value-freedom, while methodologically admirable, ultimately cannot be sustained. The selection of reference classes, the identification of functions, and the determination of what counts as a deficit all require judgments that cannot be purified of evaluative content. The question is not whether values enter psychiatric classification but whether the values that enter are articulated, defended, and answerable to objective features of human nature. The systemic alignment framework makes its values explicit and grounds them in the teleological structure of living systems. This is not value-freedom, but it is value-transparency, and it provides resources for criticizing value judgments that undermine flourishing—resources the BST, by attempting to exclude values entirely, cannot supply.

6.2.4. The Space for a New Framework

The three alternatives examined above each capture part of the truth about mental health. Fulford rightly insists that values are ineliminable from psychiatric practice. Bolton and Banner rightly seek deliberative procedures that respect diverse perspectives while acknowledging that some judgments are better supported than others. Boorse rightly seeks to ground health in objective features of organisms rather than cultural consensus. But each fails where the others succeed: Fulford's framework lacks substantive criteria for evaluating values; Bolton and Banner's framework lacks the substantive content to guide deliberation in hard cases; Boorse's framework cannot sustain its aspiration to value-freedom and provides no positive account of health.

The systemic alignment framework occupies the space these alternatives leave open. It acknowledges that values are ineliminable but grounds them in objective features of living systems (contra Fulford and Boorse). It provides substantive criteria for evaluating values and guiding deliberation (contra Fulford and Bolton and Banner). And it offers a positive account of mental health, not merely the absence of disease (contra Boorse and Wakefield). Whether it succeeds where they fail is for the remainder of this chapter—and the empirical program outlined in Chapter 5—to determine.

6.3. The Naturalistic Fallacy Objection

Objection. The framework commits the naturalistic fallacy. David Hume and G.E. Moore demonstrated that one cannot derive an "ought" from an "is"—no set of purely factual premises entails a normative conclusion. Chapter 2 establishes a recurrent structural pattern: the universe exhibits cooperative integration across scales. Chapter 3 then claims that this pattern grounds a teleological conception of goodness. But this move is illicit. The fact that atoms form molecules and cells form organisms does not

entail that humans should form cooperative relationships. The cosmos is full of predation, parasitism, and entropy. Even if integration often occurs, this is an inductive observation, not a moral command. The framework attempts to sneak normative force through the back door by calling certain patterns "cooperative" and "integrative," but these terms already carry evaluative weight. The argument therefore either begs the question or commits a logical howler. Unless the framework provides a non-normative procedure for identifying which patterns count as "good," it has merely redescribed its preferred values in the language of science.

Reply. This objection correctly identifies a logical gap that any naturalistic ethics must address, but it misunderstands the structure of the teleological bridge constructed in Chapter 3. The framework does not attempt to derive an "ought" from a bare "is" by deductive inference. Rather, it proceeds by what Philippa Foot called ‘natural normativity’—the recognition that evaluative judgments are internal to the description of living systems.

Consider an analogy. A veterinarian who says "This wolf's heart is defective" is making a claim that is simultaneously descriptive and evaluative. It is descriptive because it reports a fact about the wolf's anatomy and physiology. It is evaluative because "defective" means "fails to perform the function that a wolf's heart ought to perform." Yet this evaluation is not projected onto the wolf from the veterinarian's subjective values. It is discovered by examining what a wolf's heart does in the context of the wolf's life-form—the species-typical pattern of development, sustenance, and reproduction that defines what it is to be a flourishing wolf. The standard of evaluation is an objective feature of the natural world. One can be wrong about what a wolf's heart is for. One cannot simply decide that a wolf's heart ought to pump air instead of blood and then judge defective hearts accordingly.

The framework extends this logic to the human decision-making system. Just as a wolf's heart has an objective functional end (circulating blood to support the wolf's life), the human decision-making system has an objective functional end specified by what it is to be a flourishing human being. Chapter 3 argued, drawing on the biological and social instances of the pattern described in Chapter 2, that human beings are constitutively relational creatures. Our evolutionary history, developmental trajectory, and lived experience all indicate that we become fully human only through participation in larger integrative wholes—families, communities, ecosystems—whose health is a condition of our own flourishing. The "ought" of mental health, therefore, is not an external value imposed on neutral facts. It is the unpacking of what the facts already contain: that a well-functioning human decision-making system is one that reliably produces integrative rather than repulsive outputs, where "integrative" and "repulsive" are defined by the functional relationship between the system's outputs and the cooperative wholes upon which its flourishing depends.

The objection's appeal to predation and entropy misses the mark. The framework does not deny that destructive processes exist; it claims that health for any system consists in functioning that sustains the system's organized complexity. A wolf that fails to hunt is unhealthy, even though hunting involves predation. The framework is not a global moral theory prescribing universal non-violence; it is a theory of mental health for human beings who are embedded in particular relational systems. It does not claim that nature is harmonious or that competition is always pathological. It claims that mental health

consists in processing patterns that sustain the integrative wholes upon which human flourishing depends—and that this claim is evaluative only in the same sense that veterinary medicine is evaluative: grounded in objective facts about what functioning well consists in for a particular kind of being.

Crucially, as Chapter 2 emphasized, the normative argument does not depend on the physical instances of the pattern. The fact that quarks bind into protons is illustrative context, not a premise. The normative force comes from the teleological structure of living systems, where functional evaluation is independently defensible. The objection's worry about deriving norms from physics is therefore misdirected: the framework derives norms from biology and the human life-form, not from cosmology.

6.4. The Evolutionary Teleology Objection

Objection. The framework's reliance on teleology is scientifically untenable in a post-Darwinian world. Aristotle could speak of natural ends because he believed that organisms had fixed essences and that the universe was purposefully ordered. Darwin showed that species evolve through blind variation and natural selection, not the unfolding of intrinsic purposes. The apparent "design" of organisms is an illusion produced by selection pressures. There is no telos in nature—only causal histories. The framework attempts to resurrect a pre-modern biology that has been decisively refuted. Even if one grants that organisms have "functions" in the etiological sense (effects for which traits were selected), this does not ground normative conclusions about health. A trait that was selected in ancestral environments may be maladaptive in current environments, and a "dysfunction" in the etiological sense may be perfectly adaptive. The framework's appeal to "objective functional ends" collapses once we recognize that evolution has no direction, no purpose, and no standard of "proper" functioning beyond differential reproductive success.

Reply. This objection conflates two distinct senses of teleology: cosmic teleology (the claim that the universe as a whole has a purpose or direction) and organismic teleology (the claim that living systems exhibit goal-directedness and functional organization). Darwinian theory refuted the former but not the latter. A heart pumps blood; roots absorb water; eyes detect light. These are not illusions; they are objective facts about what these structures do and why they exist in the form they do. The etiological theory of functions (selected effects) provides a naturalistic account of teleology that is fully compatible with evolution by natural selection. A trait's function is the effect for which it was selected, and this is an objective historical fact, not a projection of human values.

However, the framework does not rely exclusively—or even primarily—on the etiological account. As Chapter 3 noted, the systemic account of functions provides an alternative or complementary grounding: a trait's function is the role it plays in maintaining the current organization of the system, regardless of its evolutionary history. The heart's function is to pump blood not merely because it was selected for that purpose but because that is what it does to keep the organism alive now. This systemic sense of function is logically prior to the evolutionary sense: natural selection favoured hearts that pump blood because blood-pumping already contributed to organismic survival. The framework grounds teleology in the self-maintaining, self-organizing character of living systems—their intrinsic tendency to persist as organized

wholes. This form of teleology requires no cosmic purpose and is entirely compatible with Darwinian evolution.

Applying this to mental health: the human decision-making system has an objective systemic function—to guide the organism's interactions with its environment in ways that sustain the organism's participation in the nested integrative wholes (social, ecological, cultural) upon which its flourishing depends. Whether a particular processing pattern is "healthy" depends on whether it serves this systemic function in the individual's actual environment, not on whether it matches ancestral selection pressures. The framework therefore avoids the objection that evolutionary functions are poor guides to current health. It grounds health not in evolutionary history but in current systemic integration.

The objection's deeper worry—that evolution has no direction—is accommodated rather than denied. The framework does not claim that evolution is directed toward greater complexity or that complexity is inherently good. It claims that wherever organized complexity exists, it is sustained by cooperative integration, and that for the particular kind of system a human being is, flourishing consists in maintaining and enhancing those integrative dynamics. This is a claim about the nature of a specific kind of being, not about the direction of evolution as a whole.

6.5. The Reductionism and Autonomy Objection

Objection. The framework reduces mental health to conformity with external systems, thereby obliterating individual autonomy and the value of non-conformity. On this account, a person is mentally healthy insofar as their decision-making aligns with "integrative wholes"—families, communities, ecosystems. But this is precisely the logic that has been used to pathologize dissent, silence critics, and enforce conformity. What about the artist whose creativity depends on alienation from mainstream society? What about the activist whose repulsion from an unjust system is a sign of moral health, not illness? What about the mystic whose inner life diverges radically from communal norms? The framework seems to entail that anyone whose processing is "repulsive"—who withdraws, criticizes, or refuses to cooperate—is mentally ill. This is not a liberating vision; it is a recipe for social control dressed in ecological language.

Reply. This objection raises legitimate and important concerns, but it misreads the framework's treatment of the relationship between individual and whole. Several distinctions are crucial.

First, the framework does not equate "integrative whole" with any existing social arrangement. A family system characterized by abuse, a community structured by exploitation, an economic system that destroys ecosystems—these are not genuine integrative wholes in the framework's sense. They are pseudo-wholes: systems that maintain a kind of stability through domination rather than cooperation. As Chapter 4 specified, a genuine integrative whole is characterized by three criteria: (a) mutual benefit—the relationships within the system tend to enhance the flourishing of the constituent parts, not merely exploit them for the benefit of a subset; (b) emergent capacity—the system generates properties or capabilities that could not be achieved by the parts in isolation, and these are broadly accessible; and (c)

dynamic stability—the system maintains itself through adaptive, responsive coordination rather than through rigid control or the suppression of negative feedback. An abusive family fails the mutual benefit condition. An exploitative economic system fails the emergent capacity condition (benefits are concentrated, not distributed) and the dynamic stability condition (it depends on suppressing dissent). The framework therefore provides resources for criticizing oppressive systems, not merely conforming to them.

Second, the framework distinguishes between repulsive processing (which degrades genuine cooperative connections) and healthy differentiation (which may involve withdrawal from pseudo-wholes). The person who leaves an abusive relationship, who protests an unjust policy, who refuses to comply with a corrupt institution—these actions may appear "repulsive" from the perspective of the pseudo-whole, but from the perspective of genuine systemic health they are integrative at a higher level. The activist who disrupts an unjust system is not degrading the cooperative bonds that constitute human flourishing; they are clearing the ground for a more genuinely integrative social order. The framework's teleology is nested: what is "repulsive" relative to a pseudo-whole may be "integrative" relative to a higher-level whole that includes the flourishing of all persons. The framework's standard is not any actual community's consensus but the objective functional requirements of systems that genuinely sustain human flourishing, as determined by the three criteria.

Third, the framework does not require that individuals always participate in existing wholes as currently structured; it requires that their decision-making be oriented toward generating and sustaining integrative wholes. This generative orientation may involve critique, reform, and even revolutionary change. The mystic who withdraws from social conventions may be cultivating a form of integration with the transcendent that ultimately deepens their capacity for compassionate engagement. The alienated artist may be generating new forms of meaning that will, over time, create more integrative cultural possibilities. The framework judges processing by its tendency—what it reliably produces across contexts—not by momentary compliance with group norms.

The deeper point is that human autonomy is not an end in itself but is valuable precisely because it enables individuals to choose integrative participation rather than being forced into it. A coerced "cooperation" is not genuinely integrative because it is not freely chosen and because it typically serves the interests of the dominator rather than the mutual benefit of all. The framework therefore values the capacity for reflective choice—the ability to stand back from social pressures and evaluate which wholes are genuinely worthy of participation. This capacity is itself a product of integrative processing: it depends on the individual's embeddedness in relationships that support critical reflection and provide a secure base from which to explore alternatives. The objection's concern about autonomy is thus not opposed to the framework but internal to it: autonomy is constituted by certain forms of relational integration and is destroyed by others.

6.6. The Pathologization of Difference Objection

Objection. Even if the framework avoids the most egregious forms of social control, it still risks pathologizing legitimate forms of human difference. Introversion, high sensitivity, neurodivergence (autism, ADHD, dyslexia), and minority sexual orientations are often characterized by processing patterns that diverge from "integrative" norms. An introvert may prefer solitude to social engagement; an autistic person may find typical social interactions exhausting or meaningless; a person with ADHD may struggle with sustained attention to community-maintaining tasks. The framework threatens to classify these as "repulsive processing"—as deviations from the integrative ideal. This is the same mistake that psychiatry has made repeatedly: mistaking statistical normality for health, and mistaking cultural preferences for universal human needs. The framework's teleological language does not protect it from this error; it merely dresses the error in more sophisticated garb.

Reply. This objection is serious, and the framework must show that it has the resources to avoid the errors of past psychiatric overreach. Several lines of response are available.

First, the framework does not equate "integrative" with "sociable," "extroverted," or "typical." Integration is a functional notion, not a statistical or behavioural one. An introvert may have rich, deep relationships with a small number of people that are more integrative—in the sense of mutual benefit, emergent capacity, and dynamic stability—than an extrovert's dozens of superficial acquaintances. The framework's criterion is the quality of relational processing, not the quantity of social behaviour. Solitude can be integrative if it is chosen as a way to restore the capacity for genuine connection. The introvert who withdraws to read a book is not necessarily degrading relational bonds; they may be nourishing the inner resources that make their relationships more authentic and sustained.

Second, the framework has the resources to distinguish primary processing tendencies that are simply different ways of being human from secondary patterns that represent dysfunction. Autism provides a test case. An autistic person may process social information differently from neurotypical norms—reduced eye contact, literal interpretation of language, difficulty with unstated social rules. Does the framework classify these as repulsive processing? Not necessarily. The question is whether these processing patterns *reliably degrade* the cooperative connections upon which the individual's flourishing depends. For many autistic individuals, the answer is no: they develop deep relationships with others who appreciate their communication style, they contribute valuable cognitive diversity to groups, and they flourish when their environment accommodates their processing differences. The "disability" in autism is often not intrinsic to the autistic processing pattern but emerges from a mismatch between that pattern and a neurotypically designed environment. The framework, by focusing on actual integration in actual environments, can acknowledge this: an autistic person in a supportive environment may be highly integrative; the same person in an unsupportive environment may struggle. The pathology, on this view, is partly located in the environment's failure to accommodate diverse processing styles—an environmental repulsion that should not be misattributed to the individual. This is consistent with the social model of disability while retaining the capacity to identify genuine dysfunction when it occurs.

Third, the framework explicitly rejects the pathologization of minority sexual orientations—a historical error that motivated much of this paper's critique of the DSM in Chapter 1. The 1973 removal of

homosexuality from the DSM was justified not because the scientific facts changed but because the normative framework shifted. The systemic alignment framework provides a principled basis for this shift: same-sex relationships can be fully integrative, satisfying the criteria of mutual benefit (partners flourish), emergent capacity (the relationship enables goods neither could achieve alone), and dynamic stability (adaptive coordination over time). The claim that heterosexuality alone is "natural" or "integrative" is an empirical falsehood, not a conclusion that follows from the framework. Sexual minorities who form committed, loving relationships are engaging in paradigmatically integrative processing. The framework therefore does not pathologize difference per se; it distinguishes between differences that support integration (the vast majority of human variation) and patterns that reliably degrade it (compulsive exploitation, inability to form any enduring bonds, persistent self-destruction that severs all connection).

The deeper point is that the framework's standard is not conformity to a statistical or cultural norm but alignment with the objective functional requirements of systems that sustain human flourishing. These requirements are real but they are also general—they specify that persons need secure attachment, opportunities for contribution, and freedom from chronic threat, but they do not specify the form that secure attachment must take (nuclear family, chosen family, multi-generational household) or the content of meaningful contribution (artisanal, intellectual, caregiving, environmental). The framework is compatible with a wide range of human diversity while still providing grounds for distinguishing health from pathology in clear cases.

6.7. The Oppression Justification Objection

Objection. The framework is dangerous. By grounding mental health in "alignment with integrative wholes," it provides a philosophical justification for demanding that individuals conform to whatever the dominant group declares to be "integrative." This is not a hypothetical risk. The history of psychiatry is replete with examples of oppression justified by similar reasoning: slaves who fled were diagnosed with drapetomania; women who sought independence had hysteria; dissidents in the Soviet Union were diagnosed with sluggish schizophrenia. The framework's teleological language does not protect against these abuses; it amplifies them by giving them a cosmic sheen. Even if the framework's author intends it to support liberation, the logic of the position—that mental health consists in alignment with larger systems—inevitably privileges the perspectives of those who control those systems. The only reliable safeguard against psychiatric oppression is to minimize the role of value judgments in diagnosis, not to embed them in a grand teleological theory. Wakefield's harmful dysfunction analysis, for all its problems, at least attempts to constrain values with objective dysfunction. This framework removes even that constraint, making mental health entirely a matter of value judgments about what counts as "integrative."

Reply. This objection raises the most serious practical concern any theory of mental health must address: the potential for misuse as an instrument of oppression. The framework must show not merely that it

can be used to resist oppression but that its internal logic constrains oppressive applications more effectively than its competitors.

The framework's critical resources have already been outlined in response to the autonomy objection (6.5) and will be demonstrated in detail in the case study below (6.9). To summarize here: the framework distinguishes genuine integrative wholes (characterized by mutual benefit, emergent capacity, and dynamic stability) from pseudo-wholes (characterized by exploitation, rigid control, and suppression of feedback). This distinction is not decorative; it is central to the framework's teleology. A system that requires the subordination, suffering, or destruction of some of its parts for the benefit of others is not genuinely integrative. It is a parasitic system—one that maintains itself through extraction rather than cooperation. The framework therefore provides a basis for criticizing oppressive systems as disintegrative, not for endorsing them.

Comparison with the harmful dysfunction analysis (HDA) is instructive. The objection claims that the HDA better constrains oppression by tying disorder to objective evolutionary dysfunction. But this is precisely the problem: the HDA provides no resources for criticizing oppressive social arrangements as oppressive. According to the HDA, a slave who experiences distress at her enslavement does not have a disorder unless some internal mechanism is dysfunctional. But the mechanism for detecting threat and responding to captivity may be functioning perfectly—slavery is genuinely harmful, and distress is an appropriate response. The HDA therefore classifies the slave as healthy, which seems right, but it provides no vocabulary for condemning the system that produces her suffering. The HDA is silent on social injustice. The systemic alignment framework, by contrast, locates the pathology partly in the environment: a social order that routinely produces distress, disability, and relational destruction is a disintegrative system—a pseudo-whole—and the individuals within it may be healthy precisely because they resist it. The framework thus provides concepts for criticizing oppressive arrangements directly, not merely for assessing individual functioning.

The objection's preferred alternative—minimizing value judgments in diagnosis—is not actually available. The 1973 homosexuality vote was not a failure to minimize values; it was a recognition that values are ineliminable and that the relevant values had shifted from pathologization to acceptance. The question is not whether values will operate in psychiatric classification but which values, and whether they are articulated and defended transparently. The systemic alignment framework makes its values explicit: health is alignment with genuinely integrative wholes, defined by specific, publicly debatable criteria; oppression is disintegrative; resistance to oppression can be an expression of health. These claims are contestable, but they are not hidden. The framework invites debate about what genuine integration consists in, rather than pretending that values can be expelled from the clinic. This transparency is a safeguard, not a vulnerability.

As for historical abuses: any diagnostic concept can be weaponized. "Dysfunction" was used to justify drapetomania (slaves' desire for freedom was declared a dysfunction of the "flight mechanism"). "Harm" was used to justify the institutionalization of women who sought independence. The question is not whether a concept can be misused—all can—but whether it contains internal resources for correcting misuse. The systemic alignment framework contains such resources in its criteria for genuine integration.

A diagnosis of drapetomania would fail the framework's test: the "system" of slavery was a pseudo-whole characterized by exploitation and the suppression of feedback, and the slave's flight was integrative at the higher level of universal human flourishing. The framework does not guarantee that such diagnoses will never be made—no framework can—but it provides the conceptual tools to unmask them as abuses when they occur.

6.8. The Empirical Testability Objection

Objection. Despite Chapter 5's detailed measurement proposals, the framework is not genuinely testable in its core claims. The central constructs—"integrative whole," "repulsive processing," "emotional capture"—remain too vague to operationalize with precision, and the framework's most distinctive predictions could be reinterpreted to accommodate almost any pattern of results. What would count as a genuine failure of the theory? If some individuals high in IPO were miserable, the framework could claim they were not "truly" integrated (perhaps their prosociality was coerced or shallow). If some individuals low in IPO were thriving, the framework could claim they had found integration at a higher level (perhaps their solitary pursuit was actually serving a larger whole). This is the mark of a pseudoscientific theory: it explains everything after the fact and predicts nothing in advance. Until the framework specifies in advance what evidence would refute it, it remains a metaphysical speculation dressed in empirical clothing. Moreover, the IPO Scale's factor structure—if it fails to replicate—could be dismissed as a measurement problem rather than a theoretical one. The framework thus appears unfalsifiable at its core.

Reply. This objection charges the framework with unfalsifiability. The charge must be taken seriously, but it overstates the case. The framework makes specific, risky predictions that could be disconfirmed.

Chapter 5 articulated five core predictions, each of which could fail. Prediction 1: prosocial integration will correlate positively with well-being. This could fail: a large, well-powered study might find no relationship, or a negative relationship. Prediction 2: contribution-oriented purpose will predict well-being over and above self-referential purpose. This could fail if self-referential purpose equally predicts well-being. Prediction 3: nature-connectedness will predict well-being independently of social connection. This could fail if nature-connectedness adds no unique variance. Prediction 4: IPO will mediate the relationships between these predictors and well-being. This could fail if no mediation is found or if direction reverses. Prediction 5: clinical populations will score lower on IPO and higher on repulsive measures. This could fail if clinical populations show no difference or opposite differences.

The objection anticipates a defensive move: when faced with disconfirming evidence, the framework's proponent might claim the measured construct was not "true" integration. This is a legitimate concern, but it is addressed by specifying operational definitions in advance and committing to refining measures based on empirical findings—a standard practice in psychological science, not a mark of pseudoscience. The IPO Scale proposed in Chapter 5 includes specific item content, a hypothesized factor structure, and predicted correlations with established measures. If the factor structure fails to replicate across multiple samples, the framework is in trouble. If IPO fails to show incremental validity beyond agreeableness and

social desirability, the framework is in trouble. If longitudinal studies find that changes in well-being precede changes in IPO rather than the reverse, the framework's causal claims are challenged. These are not logical impossibilities; they are empirical possibilities that could falsify specific claims.

The framework also makes distinctive predictions that differentiate it from competitors. Self-determination theory predicts that autonomy, competence, and relatedness predict well-being—but it does not predict that ecological connection contributes independently beyond social relatedness. Evolutionary psychology predicts that prosociality can be adaptive—but it does not predict that a general factor of integrative processing spanning interpersonal, systemic, ecological, and temporal domains mediates well-being outcomes. These are distinctive predictions; their failure would count against the framework specifically.

Finally, the objection's standard of falsifiability is unreasonably high. No psychological theory specifies in advance every condition under which it would be abandoned, because theories are research programs that develop over time. The question is whether the framework generates specific, testable predictions that could lead reasonable researchers to reduce their confidence in it. The answer is yes. If, after a decade of research using validated IPO measures, meta-analyses find that IPO does not predict well-being beyond established personality traits, that it does not mediate the effects of social connection or nature contact, and that clinical populations do not differ from controls on IPO, the framework would be in serious empirical trouble. The author would be among the first to acknowledge this.

6.9. Case Study: Gender Dysphoria in a Non-Affirming Environment

The preceding sections have addressed theoretical objections and established the framework's position relative to its principal competitors. But a framework's adequacy is ultimately demonstrated in its application to contested cases. This section applies the systemic alignment framework to a single, highly contested example—the experience of gender dysphoria in a non-affirming environment—and shows, step by step, how the framework's criteria yield a determinate, non-arbitrary judgment. The analysis then compares this judgment with those generated by the harmful dysfunction analysis (HDA), by pure cultural consensus, and by the alternative frameworks examined in 6.2, demonstrating the framework's comparative advantages.

6.9.1. The Case

Consider Alex, a 19-year-old assigned female at birth who has experienced persistent gender dysphoria since early adolescence. Alex experiences significant distress related to female secondary sex characteristics, a strong and consistent identification as male, and a deep desire to live and be recognized as a man. Alex lives in a socially conservative community where gender non-conformity is strongly stigmatized. Family members have expressed that Alex's identification is sinful or delusional. Peers have engaged in harassment. Alex's distress is compounded by this social rejection but is not

reducible to it; even in private, Alex experiences a profound sense of incongruence between experienced gender and bodily characteristics.

Alex faces a decision: whether to pursue gender-affirming medical treatment (testosterone therapy, possibly surgery) or to attempt to suppress or "correct" these feelings through some form of conversion effort or continued concealment. The question for a theory of mental health is: what does health consist in for Alex? Is Alex's gender dysphoria a mental disorder? Would pursuing transition be an expression of health or pathology? Would attempting to suppress these feelings be healthy or harmful?

6.9.2. The HDA Analysis

The harmful dysfunction analysis would approach Alex's case as follows.

The dysfunction component requires identifying a psychological mechanism that is failing to perform its naturally selected function. What is the relevant mechanism? One candidate is the mechanism underlying gender identity formation. If this mechanism was selected to produce a gender identity aligned with biological sex (because such alignment maximized reproductive success in ancestral environments), then Alex's transgender identity represents a dysfunction of this mechanism. Alternatively, one might argue that the mechanism underlying gender identity was not specifically selected for alignment with biological sex, or that we lack sufficient knowledge of the mechanism to make a determination. The HDA thus either pathologizes Alex's gender identity (if it posits a dysfunction) or remains agnostic (if it claims insufficient knowledge). Either outcome is problematic: the first pathologizes a stable identity that, when supported, is compatible with high functioning; the second provides no guidance.

The harm component requires that this dysfunction cause harm as judged by cultural standards. In Alex's non-affirming environment, gender dysphoria clearly causes harm—distress, social disability, increased risk of suicide. But the HDA cannot distinguish between harm caused by the condition itself and harm caused by social rejection. If the culture were fully affirming, would Alex experience harm? The HDA's reliance on cultural values means that in a non-affirming culture, the "harm" judgment is straightforward; in an affirming culture, it might not be. The HDA thus fails to provide a stable, principled judgment. It either pathologizes Alex (if dysfunction is claimed) or remains hostage to cultural consensus about harm—the very problem the HDA was designed to solve.

6.9.3. The Cultural Consensus Analysis

Under pure cultural consensus, the judgment is straightforward but deeply problematic. In Alex's non-affirming community, the cultural consensus is that gender dysphoria is a disorder, that transition is pathological, and that Alex should attempt to align identification with biological sex. In a different cultural context—say, a progressive urban community—the consensus would be that gender dysphoria is not a disorder, that transition is a legitimate medical treatment, and that Alex should be supported in

living authentically. Under cultural consensus, mental health is simply whatever the local culture says it is. This yields contradictory judgments with no principled way to adjudicate between them. It was precisely this problem—exposed by the 1973 homosexuality vote—that motivated the search for a more principled foundation.

6.9.4. The Alternative Frameworks

The three alternatives examined in §6.2 would approach Alex's case as follows.

Values-based practice would identify the conflicting values at play—Alex's value of authentic self-expression, the family's values of religious tradition and gender conformity, the community's values of social stability—and seek to negotiate among them through mutual respect and careful attention to the values actually operative in the clinical encounter. But VBP provides no basis for determining that some of these values are incompatible with Alex's flourishing. The family's values may be sincerely held, but they demand that Alex suppress her identity—a demand that, as the framework's criteria would show, violates the mutual benefit condition and sustains a pseudo-whole at Alex's expense. VBP cannot say this without importing values beyond its procedural framework. It can facilitate negotiation; it cannot provide guidance.

Bolton and Banner's deliberative framework would bring Alex, family members, clinicians, and community representatives into inclusive deliberation, aiming to reach a shared understanding through evidence-informed dialogue. Under ideal conditions, this might produce consensus on supporting Alex's authentic identity. But in Alex's actual, non-ideal environment, deliberation among parties with fundamentally divergent understandings of the good may simply reproduce the conflict. The framework provides a procedure but no substantive criteria for evaluating outcomes. It cannot say, independently of the deliberation's outcome, that the family's demand for suppression is incompatible with Alex's flourishing. The systemic alignment framework can.

Boorse's biostatistical theory would attempt to determine whether gender dysphoria represents a statistically atypical functional deficit relative to Alex's reference class. But the selection of the reference class is precisely what is at issue: is the relevant class "human females aged 19" or "transgender males aged 19"? The former would likely yield a judgment of dysfunction; the latter would not. The BST provides no principled way to resolve this question without importing values. Moreover, even if the BST identified a dysfunction, it would remain silent on whether transition or suppression is the healthier path, since it provides no positive account of health, only a negative account of disease absence.

6.9.5. The Systemic Alignment Analysis

The systemic alignment framework approaches Alex's case through the lens of integrative versus repulsive processing, applying the criteria developed in Chapters 3 and 4.

Step 1: Identify the relevant systems and their constitutive constraints. Alex is a reflective conscious processor (Chapter 3, §3.5) whose flourishing depends on participation in multiple nested systems: individual (bodily integrity, authentic self-expression), interpersonal (family relationships, friendships, intimate partnerships), communal (participation in social institutions), and ecological (connection to the natural world). Alex's constitutive constraints include the need for bodily integrity and comfort, for authentic self-expression, for secure attachment relationships, for social recognition and participation, and for a coherent narrative identity that integrates these domains.

Step 2: Assess whether Alex's current state represents repulsive processing. Repulsive processing is defined as a pattern of decision-making that systematically degrades the cooperative connections upon which flourishing depends (Chapter 4, §4.3). Alex's gender dysphoria involves significant distress. But distress alone does not constitute repulsive processing. The question is directional: do Alex's processing patterns tend to build or degrade integrative wholes?

Consider the two paths available to Alex:

Path A: Suppression/Conversion. Alex attempts to suppress or change gender identity to align with community expectations. On this path, Alex must engage in sustained self-denial—a form of processing in which authentic self-experience is systematically overridden. The emotional capture analysis of Chapter 3 (3.8) would predict that this generates a meaning-capture loop: the need for belonging produces temporary compliance, which reduces immediate social friction, but the underlying incongruence persists and may intensify, producing shame, self-alienation, and internal fragmentation. Over time, this path degrades Alex's capacity for authentic connection across all relational domains. Relationships become performances rather than genuine encounters. The family system may appear "integrated" on the surface, but this is a pseudo-integration achieved through the suppression of a member's authentic self—a violation of the mutual benefit criterion. Alex's individual processing becomes repulsive: it systematically degrades the cooperative connection between self and authentic experience, and by extension degrades the capacity for genuine connection with others. The long-term outcomes associated with this path—elevated rates of depression, anxiety, substance abuse, and suicide—are precisely what the framework would predict for sustained repulsive processing.

Path B: Gender-Affirming Transition. Alex pursues medical and social transition to align bodily characteristics and social presentation with experienced gender identity. On this path, Alex's processing shifts from self-suppression to self-integration. The decision to transition represents, in the framework's terms, a move toward integrative processing: Alex's decision-making system begins to produce outputs (self-acceptance, authentic self-presentation, pursuit of congruence) that build the cooperative connection between self and body, self and identity, self and others. This is not repulsive but integrative: it enables Alex to participate in relationships as the person Alex is, rather than as a constructed persona. The family system may experience turbulence as members adjust, but the framework's criteria for genuine integration—mutual benefit, emergent capacity, dynamic stability—are better served by relationships grounded in truth than by those grounded in suppression. The well-documented improvements in mental health outcomes following gender-affirming treatment are precisely what the framework would predict for a shift from repulsive to integrative processing.

Step 3: Address the "pseudo-whole" problem. Alex's non-affirming community presents itself as an integrative whole to which Alex should conform. But does it satisfy the criteria for genuine integration? Under examination, it fails the mutual benefit condition: Alex's flourishing is sacrificed for the community's comfort. It fails the dynamic stability condition: stability is maintained through the suppression of difference and the exclusion of feedback (Alex's authentic experience) that would require the community to adapt. It fails the emergent capacity condition: the community's rigid gender norms constrain rather than enable the flourishing of its members. The community is thus a pseudo-whole—a system that maintains coherence through domination rather than cooperation. Alex's "repulsion" from this community's gender norms is not pathological; it is healthy differentiation from a disintegrative system, in service of alignment with a more genuinely integrative whole (a community of belonging that affirms Alex's identity).

Step 4: Determine what health consists in for Alex. The framework's answer is clear: health for Alex consists in processing that builds genuine integrative wholes, characterized by mutual benefit, emergent capacity, and dynamic stability. Alex's gender dysphoria is not itself a mental disorder; it is a condition of distress arising from incongruence between experienced identity and bodily/social reality. The distress is real and warrants compassionate response. But the pathology lies not in Alex's gender identity but in the repulsive processing that would be required to suppress it, and in the social environment that demands such suppression. The framework thus supports Path B (gender-affirming transition) as the path of health, not because it follows cultural consensus—in Alex's environment, it defies consensus—but because it represents a shift toward integrative processing: aligning decision-making with the authentic self, enabling genuine connection with others, and contributing to the emergence of more inclusive social wholes.

Step 5: Address the objection that this is value-laden. The framework's judgment is value-laden, but not arbitrarily so. It is grounded in objective facts about what human flourishing requires: authentic self-expression, secure attachment, freedom from chronic suppression, participation in cooperative wholes. These are not culturally specific values; they are constitutive constraints of the kind of being a human is, as established in Chapter 3. The judgment that suppression is repulsive and integration is healthy follows from these constraints, not from the theorist's personal preferences. The framework makes its values explicit and invites debate about them—a transparency that cultural consensus, the HDA, and the alternative frameworks examined in §6.2 all lack to varying degrees.

6.9.6. Comparative Assessment

The case study demonstrates the systemic alignment framework's comparative advantages:

Against the HDA. The HDA either pathologizes Alex's gender identity (if it posits dysfunction) or remains agnostically silent (if it claims insufficient knowledge). Both outcomes fail Alex: one actively harms, the other provides no guidance. The systemic alignment framework, by contrast, identifies the pathology in the repulsive processing required to sustain suppression and in the pseudo-whole that demands it, while

recognizing Alex's authentic identity as potentially integrative. The framework thus provides a principled basis for affirming Alex's health-seeking trajectory that the HDA cannot.

Against cultural consensus. In Alex's non-affirming environment, cultural consensus declares Alex disordered. In a progressive environment, cultural consensus declares Alex healthy. The systemic alignment framework cuts through this contradiction by providing a stable standard—alignment with genuine integrative wholes, defined by specific functional criteria—that yields the same judgment regardless of cultural context. In the non-affirming environment, the framework identifies the community as a pseudo-whole and Alex's authentic self-expression as integrative. In an affirming environment, the framework would identify the community as a genuine integrative whole and Alex's participation in it as health-sustaining. The criterion is consistent; only the environmental conditions differ.

Against the alternative frameworks. Values-based practice cannot say that the family's values are incompatible with Alex's flourishing; it can only negotiate. Bolton and Banner's deliberative framework provides a procedure but not a criterion, and in non-ideal conditions may simply reproduce the conflict. Boorse's biostatistical theory cannot resolve the reference class problem and provides no positive guidance. The systemic alignment framework, by contrast, provides substantive criteria (mutual benefit, emergent capacity, dynamic stability) that yield a determinate, principled judgment while remaining open to contestation and revision.

Against the pathologization of difference. The framework does not pathologize Alex's gender identity; it pathologizes the suppression of authentic identity and the social rejection that demands it. This is precisely the right result: affirming what is integrative about Alex's experience while identifying what is genuinely pathological about the environment. The framework's graduated account of normative standing (Chapter 3, §3.5) supports this: as a reflective conscious processor, Alex's flourishing requires coherent identity and authentic self-expression, and systems that systematically frustrate these requirements are disintegrative.

6.10. Boundary Condition Clarifications

The case study above demonstrates the framework's application to a specific contested case. Several general boundary condition issues raised in earlier objections can now be addressed with greater precision.

Healthy Differentiation vs. Repulsive Withdrawal. The gender dysphoria case illustrates the distinction. Healthy differentiation (Alex's refusal to comply with community gender norms, Alex's withdrawal from rejecting relationships) is characterized by: (a) selective disengagement from specific relationships or patterns that are genuinely harmful, combined with maintained or enhanced engagement with other integrative wholes (Alex builds new, affirming relationships); (b) agency (Alex chooses this path reflectively, not compulsively); and (c) renewed capacity for integration (Alex's transition enables deeper, more authentic connection). Repulsive withdrawal, by contrast, would involve global disengagement

from all potential integrative wholes, driven by fear or shame, entrenching patterns that prevent future connection. The framework provides criteria for distinguishing these: the direction of the processing trajectory and its effects on the individual's capacity for integration across the nested systems they inhabit.

Genuine Wholes vs. Pseudo-Wholes. The gender dysphoria case also illustrates the criteria. Alex's non-affirming community presents itself as a "whole" to which members should conform—a family system, a religious community, a local culture. But it fails the mutual benefit condition (Alex's flourishing is sacrificed for others' comfort), the emergent capacity condition (rigidity prevents the community from generating new capacities through diversity), and the dynamic stability condition (stability depends on suppressing difference and excluding feedback). It is thus a pseudo-whole. A genuinely integrative community would satisfy all three criteria: Alex's authentic self would be welcomed as contributing to the community's emergent capacities, and the community would adapt dynamically to include Alex's perspective.

When Alignment with a Whole Becomes Pathological. The framework must acknowledge that alignment with a genuine integrative whole can become pathological if it involves the sacrifice of the individual's constitutive constraints. A genuinely integrative family (mutual benefit, emergent capacity, dynamic stability) might still demand that Alex suppress identity for the sake of harmony. The framework's criterion is that genuine integration requires the flourishing of the parts, not merely the persistence of the whole. If Alex's participation requires self-suppression, the whole is failing the mutual benefit criterion with respect to Alex, regardless of how well it treats other members. Alignment with a whole becomes pathological when it prevents the individual from satisfying their constitutive constraints as the kind of being they are. This is not an exception to the framework but an application of it: health requires alignment with wholes that genuinely sustain the flourishing of their parts.

6.11. Conclusion: The Framework's Defensible Core

The objections considered in this chapter are substantial. The naturalistic fallacy challenge forces clarity about the structure of natural normativity. The evolutionary teleology objection demands a distinction between cosmic and organismic teleology. The autonomy and pathologization concerns highlight the need for careful distinctions between conformity to existing systems and genuine integration. The boundary condition problem reveals that the framework requires substantive criteria for distinguishing healthy integration from pathological fusion. The empirical testability objection presses for specificity in predictions and operational definitions. The oppression justification and relativism objections raise the deepest practical and philosophical stakes. And the alternative frameworks examined in §6.2—Fulford's values-based practice, Bolton and Banner's deliberative model, and Boorse's biostatistical theory—demonstrate that the space the present framework occupies is genuinely unfilled by existing approaches.

Yet the framework survives these challenges with its core commitments intact. The defensible core can be summarized as follows:

1. Natural normativity. Evaluative judgments about living systems are grounded in objective facts about what constitutes well-functioning for a particular kind of being. This is not a deduction of "ought" from bare "is" but an appreciation that the "is" of living systems already contains normative structure. This position has been defended against expressivist, constructivist, and value-skeptical alternatives.

2. Teleology without cosmic purpose. Organisms have objective functional ends specified by their self-maintaining, self-organizing character. This organismic teleology is fully compatible with Darwinian evolution and does not require cosmic direction.

3. Constitutive relationality. Human beings are constitutively relational creatures. Individual flourishing depends on participation in larger systems that exhibit genuine integration—characterized by mutual benefit, emergent capacity, and dynamic stability.

4. Systemic alignment as health. Mental health consists in the reliable tendency of the decision-making system to produce reactions that sustain and generate integrative wholes. This is a positive conception of health, not merely the absence of pathology.

5. Repulsive processing as illness. Mental illness consists in patterns of decision-making that systematically degrade the cooperative connections upon which flourishing depends, often maintained through emotional capture loops. This is a structural description, not a moral condemnation.

6. Falsifiable predictions. The framework generates specific predictions about the relationships between Integrative Processing Orientation, prosocial integration, purpose, nature-connectedness, and well-being, operationalized through the IPO Scale proposed in Chapter 5.

7. Critical resources. The framework provides criteria for distinguishing genuine integrative wholes from pseudo-wholes, healthy differentiation from repulsive withdrawal, and transcendent experience from psychotic disintegration. These distinctions enable critique of oppressive systems rather than mere conformity to them, as demonstrated in the gender dysphoria case study.

The framework does not claim to have resolved every conceptual problem or to have eliminated all ambiguity. No theory of mental health can achieve that, given the complexity of the phenomena. What the framework offers is a coherent, empirically grounded, practically orienting conception of mental health that addresses the normative crisis exposed by the 1973 homosexuality vote and the subsequent failures of the harmful dysfunction analysis and its principal alternatives. It provides principled guidance in contested cases where cultural consensus divides, the HDA falls silent, and competing frameworks cannot supply determinate criteria. Whether this conception is ultimately accepted will depend on its ability to guide clinical practice, generate productive research, and withstand continued philosophical scrutiny. The objections raised in this chapter are invitations to refine and deepen the framework, not fatal blows that require its abandonment. The conversation about what mental health is and ought to be is far from over; this framework aims to advance that conversation by providing what the field most urgently lacks: a conception of mental health that is simultaneously scientifically grounded, philosophically rigorous, and practically orienting in the hardest cases.

Chapter 7: Conclusion: The Architecture of a Liveable Life

The argument of this paper has proceeded through six stages, each building on the last, toward a single conclusion: mental health is not the absence of pathology but the presence of a particular kind of relational functioning—a processing orientation that sustains and generates the integrative wholes upon which human flourishing depends. This conclusion is at once ancient in its resonances and novel in its synthetic grounding. It recovers insights that have animated traditions of thought from Aristotle to the humanistic psychologists, while anchoring them in a framework that is responsive to contemporary physics, evolutionary biology, clinical psychology, and the hard lessons of psychiatry's own history.

The normative crisis that opened this inquiry was not manufactured for philosophical sport. It is a real and consequential fracture at the base of a global apparatus of diagnosis, treatment, and reimbursement. When the definition of mental disorder can shift with a committee vote—as it did in 1973, and as it continues to do with each DSM revision—the entire clinical edifice operates without a clear mandate. The paper's first chapter documented this crisis not to disparage psychiatry but to demonstrate the urgency of the conceptual work required. The dominant philosophical solution, Wakefield's harmful dysfunction analysis, was shown to be an unstable hybrid: an objective evolutionary criterion harnessed to a culturally variable harm judgment, with no account of how the two are to be integrated. The deeper problem, as the chapter argued, is that the field has treated mental health as an unexamined default—the absence of disorder—rather than as a positive condition with its own identifiable structure. The chapters that followed aimed to supply that structure.

Chapter 2 established a descriptive foundation that was deliberately ambitious in scope and deliberately modest in its claims. Across the sciences, from quantum binding to human institutions, organized complexity arises and persists through a recurrent structural pattern: individuals enter into cooperative relationships, mediated by physical carriers, to form more encompassing systems. This pattern is not a cosmic law but a recurrent isomorphism—an observable regularity that holds wherever complexity endures. The chapter was careful to distinguish this descriptive "is" from any prescriptive "ought," while noting a structural truth that would prove relevant throughout: that a system whose internal dynamics become predominantly repulsive—in which parts degrade rather than sustain the relationships that constitute the whole—is a system in the process of dissolution. This truth, established at the physical level and extended through biology and social organization, provided the structural context for the argument to come.

Chapter 3 constructed the teleological bridge that Hume's challenge demands. Drawing on the neo-Aristotelian naturalism of Philippa Foot and Michael Thompson, it argued that living systems possess objective functional ends that are discernible features of their nature, not subjective projections. A wolf's heart has the function of pumping blood; this is a fact about the wolf, not a human preference projected onto it. Health, for any living system, is functioning well relative to the life-form that defines its kind. The chapter extended this logic to the human case, arguing that the human life-form is constitutively relational. We become fully human only through participation in larger wholes—families, communities, ecosystems—whose health is a condition of our own flourishing. The graduated account of normative standing situated this claim within the broader living world without overgeneralizing; the

carrier concept anchored the language of cooperation in physical reality; the emotional capture analysis explained why repulsive processing, once established, is self-reinforcing. The "ought" of mental health, the chapter concluded, is not deduced from a bare "is" but unpacked from a richer understanding of what the "is" of human nature contains.

Chapter 4 applied this teleological framework to the specific domain of mental health, yielding the paper's central definitions. Mental health is systemic alignment: the condition in which an individual's decision-making system reliably produces thoughts, emotions, and behaviours that sustain and generate integrative wholes. Mental illness is repulsive processing: a systematic tendency of the decision-making system to produce outputs that degrade, sever, or prevent the cooperative connections upon which both individual and collective flourishing depend. The chapter located the decision-making system within a constitutive relationality that aligns the framework with 4E approaches to cognition while adding the normative resources to distinguish healthy embeddedness from pathological enmeshment. The clinical phenomenology of depression, anxiety, psychosis, and addiction was reframed through this lens, revealing a common directional tendency beneath diverse symptom profiles.

Chapter 5 demonstrated that the framework is not untestable metaphysical speculation. The Integrative Processing Orientation (IPO) construct—a proposed disposition to process experience in ways that build cooperative wholes—was specified with its four-factor structure, sample items, and a multi-phase validation strategy. The chapter articulated five core predictions: that prosocial integration, contribution-oriented purpose, nature-connectedness, and a general IPO factor would predict psychological well-being, and that clinical populations would exhibit measurably lower integrative processing. Crucially, the chapter specified seven falsification conditions—patterns of results that, if obtained, would require abandoning the framework's core claims—distinguishing these from results that would require mere modification. The chapter also demonstrated the framework's unificatory power by showing how cognitive-behavioural, acceptance-based, psychodynamic, humanistic, systemic, and mindfulness interventions can all be understood as methods for interrupting emotional capture loops and restoring integrative processing.

Chapter 6 subjected the framework to the most serious objections it faces. The naturalistic fallacy objection was met by clarifying that natural normativity is not a deduction of "ought" from "is" but a recognition that the "is" of living systems already contains normative structure. The evolutionary teleology objection was addressed by distinguishing cosmic teleology from organismic teleology, the latter being fully compatible with Darwinian biology. The autonomy, pathologization, and oppression objections were answered by the framework's critical resources: the criteria of mutual benefit, emergent capacity, and dynamic stability that distinguish genuine integrative wholes from oppressive pseudo-wholes. The gender dysphoria case study demonstrated the framework's capacity to yield a determinate, non-arbitrary judgment in a deeply contested case, outperforming the harmful dysfunction analysis, cultural consensus, values-based practice, deliberative frameworks, and biostatistical theory. The framework did not emerge from these tests unscathed—no philosophical theory of mental health ever has—but it emerged with its core commitments intact.

What, then, has been accomplished?

The most modest answer is that the paper has provided a coherent, empirically grounded, and practically orienting alternative to the dominant approaches to defining mental disorder. It has shown that the normative crisis in psychiatry is not intractable; that there exists a path between the Scylla of cultural relativism and the Charybdis of value-free scientism; and that this path can be walked with philosophical rigour and empirical responsibility. The IPO construct, should it survive the validation program outlined in Chapter 5, would provide psychiatry with something it has never possessed: a positive measure of mental health grounded in an explicit theoretical framework, capable of guiding both clinical practice and public health policy.

A less modest answer is that the paper has demonstrated something about the relationship between facts and values that extends beyond psychiatry. The is-ought gap, that perennial philosophical obstacle, is not a logical chasm that no argument can cross. It is a reflection of an impoverished conception of what the "is" contains. Once we recognize that living systems have natures—that a wolf is not merely a collection of cells but a kind of being with objective functional requirements—the gap begins to close. The "ought" is not added to the "is" from outside; it is discovered within it, as the conditions of flourishing that the being's own nature specifies. This is not to commit the naturalistic fallacy but to overcome it, by showing that Hume's challenge can be met with a richer naturalism than Hume himself could have imagined—one informed by evolutionary biology, systems theory, and the empirical study of human development.

The most ambitious answer is that the paper has offered a vision of mental health that connects the most intimate dimensions of human experience to the broadest architecture of the cosmos. The same pattern that binds quarks into protons, cells into organisms, and individuals into communities also structures the well-functioning human mind. To be mentally healthy is to be aligned with the grain of reality—not with a supernatural order or an arbitrary cultural convention, but with the relational logic that generates and sustains organized complexity at every scale. This is not a moral demand dressed in scientific language. It is an empirical claim about what flourishing consists in for the kind of being a human is: a constitutively relational creature whose good is inseparable from the good of the wholes it participates in.

Much work remains. The IPO Scale must be developed, validated, and refined. The framework's predictions must be subjected to rigorous longitudinal and experimental tests. The boundary between genuine integration and pathological fusion must be further clarified through engagement with clinical cases. The framework's implications for social policy, education, and the design of therapeutic interventions must be worked out in detail. And the philosophical conversation with expressivists, constructivists, and value-sceptics will continue, as it should, for a theory that makes such substantive claims about the human good.

But the direction of travel is clear. The normative crisis in mental health arose because psychiatry tried to define disorder without defining health, and tried to define health without defining the human being whose health is at stake. This paper has argued that the human being is not an isolated atom of desire and cognition but a node in a network of constitutive relationships, a processing system whose flourishing depends on the health of the families, communities, and ecosystems that sustain it. Mental

health, on this view, is not a private achievement but a relational condition. It is cultivated not by withdrawing into self-optimization but by orienting one's decision-making toward the generation of integrative wholes—toward actions, thoughts, and emotions that build rather than erode, connect rather than sever, sustain rather than extract.

This is not a new insight. It is found in Aristotle's conception of the political animal, in the Ubuntu philosophy of personhood through other persons, in the attachment theory of John Bowlby, and in the ecological wisdom of indigenous traditions. What is new is the attempt to ground this insight in a synthetic framework that draws equally on the empirical sciences and on rigorous philosophical argument, and to operationalize it in a construct that can be measured, tested, and applied. The work of demonstrating that this framework is true—not merely elegant, not merely resonant with ancient wisdom, but true in the sense that its predictions survive empirical scrutiny and its applications improve human lives—remains to be done. This paper has aimed to make that work possible.

The 1973 vote that removed homosexuality from the DSM was both a triumph of justice and a symptom of a deeper conceptual failure. Justice prevailed, but it prevailed through a political process rather than a scientific one, leaving the foundation of psychiatric classification as vulnerable as it had always been. The challenge for the philosophy of psychiatry has been to build a foundation that could support such corrections on principle rather than by pressure—a foundation that could distinguish genuine pathology from socially disfavoured variation without collapsing into cultural relativism or pretending to a value-freedom that cannot be sustained. This paper has attempted to provide that foundation. Whether it succeeds will be determined not by the elegance of its prose or the scope of its ambition but by the conversations it provokes, the research it generates, and the lives it helps to guide.

A definition of mental health, in the end, is not merely a theoretical construct. It is a statement about what a human life is for. To define mental health as systemic alignment—as the cognitive and behavioural alignment with the universal, systemic principle of generating and maintaining integrative wholes—is to claim that human beings find their fullest flourishing not in the maximization of individual pleasure or the elimination of all distress but in the cultivation of relationships that build the kinds of wholes from which no one is excluded and in which everyone can participate. It is to claim that the architecture of a liveable life is the same architecture that builds galaxies, cells, and communities: cooperative integration, sustained by the carriers that mediate our connections, oriented toward the health of the nested systems that make us who we are. That is a large claim, and it demands large evidence. The conversation about what mental health is and ought to be is far from over. But this framework, if it has succeeded, has shown what kind of answer the question requires.

Supplementary - The Meta-Criterion

The Meta-Criterion: Recursive Non-Contradiction in Constitutive Constraint Satisfaction

The meta-criterion holds that a claim that a system constitutes a genuine integrative whole is valid only if the claim can be validated from the standpoint of each constitutive member's processing architecture, under conditions that do not systematically disable that architecture's capacity for validation. A claim is a post-hoc rationalization for power when its acceptance depends on some members' validating capacities being systematically bypassed, suppressed, or overridden.

This formulation requires unpacking, but its core intuition is straightforward: a genuine whole cannot be built on the backs of members whose own processing systems—the very systems whose integrative functioning defines health—are prevented from functioning as the kind of processing systems they are.

The Criterion in Detail

The meta-criterion rests on three nested conditions, each derived from the framework's prior commitments:

Condition 1: Procedural Accessibility. The claim that a system satisfies mutual benefit, emergent capacity, and dynamic stability must be in principle accessible to all constitutive members of the system through the exercise of their reflective processing capacities, as specified in Chapter 3's graduated account of normative standing (3.5). For human beings, this means: the justification cannot depend essentially on premises that only certain members are permitted to entertain, on evidence that is systematically withheld from certain members, or on reasoning steps that certain members are punished for performing. "You would see that this benefits you if you had faith" is not, by itself, a valid claim under this condition, because it makes validation depend on a cognitive state whose absence is precisely what the critical member is being asked to overcome. The condition does not require that every member actually endorse the claim—people can be mistaken, inattentive, or captured by repulsive processing—but it does require that the path to endorsement not be structurally blocked.

Condition 2: Non-Self-Undermining Testimony. When a constitutive member reports that their participation in the system systematically frustrates their constitutive constraints—their needs for authentic self-expression, secure attachment, freedom from chronic threat, and so forth—this testimony cannot be dismissed solely on the grounds that the member's processing architecture is defective. To do so would be to beg the question. The framework's emotional capture analysis (Chapter 3, 3.8) acknowledges that individuals can be mistaken about their own good; repulsive processing can generate self-defeating preferences. But the fact that some self-reports are distorted does not license dismissing all self-reports that conflict with a system's self-justification. The meta-criterion therefore places the burden of proof on the system: to dismiss a member's testimony of harm, the system must demonstrate, through independent evidence and reasoning accessible to the member, that the testimony is the product of repulsive processing rather than an accurate registration of systemic failure. "You only feel

oppressed because your gender dysphoria distorts your perception" is circular if the diagnosis of dysphoria as distorting is precisely what is in question.

Condition 3: Counterfactual Stability. A claim that a system is genuinely integrative must be stable under the counterfactual of power redistribution. Consider: would the claim that Alex's suppression benefits Alex still be made by the family if the family did not hold the power to enforce that suppression? Would the family accept the same reasoning if their constitutive constraints were being overridden for the sake of a larger whole they did not recognize? The meta-criterion does not require that power be actually redistributed—that is a political demand, not a conceptual one—but it requires that the justification for the system's integrative status not depend essentially on the fact that the powerful can enforce their interpretation. A justification that collapses when one imagines the positions reversed is a rationalization, not a reason. This is not a raw appeal to a universalizability principle external to the framework. It follows from the framework's own claim that human beings share a common life-form with common constitutive constraints. If the justification for sacrificing Alex's constraints would be rejected by the sacrificers were they in Alex's position, then the justification treats Alex's constraints as less real than the sacrificers' own—which is precisely to deny that Alex is a full constitutive member of the purported whole.

What the Meta-Criterion Rules Out

Applied to the religious community case, the meta-criterion yields a clear analysis:

The claim that Alex's suppression provides "mutual benefit" via eternal salvation fails Condition 1 because the validation path depends on accepting theological premises that Alex's reflective processing, operating under conditions of authentic self-awareness, cannot endorse without self-betrayal. The "benefit" is accessible only if Alex first accepts that her experienced identity is delusional—but whether that identity is delusional is precisely what is at issue. The justification is procedurally closed to Alex as the kind of processing system Alex is.

It fails Condition 2 because Alex's testimony that suppression causes suffering and self-fragmentation is dismissed on the grounds that Alex's processing is disordered—but the disorder attribution depends on the very norms Alex is contesting. The dismissal is circular.

It fails Condition 3 because the family would not accept a parallel justification in which their deepest identity commitments were declared delusional for the sake of a larger whole they did not recognize. The justification depends on power asymmetry.

The meta-criterion therefore identifies the family's claim as a post-hoc rationalization for a pseudo-whole, not a genuine instance of mutual benefit.

What the Meta-Criterion Permits

Crucially, the meta-criterion does not rule out all claims that an individual's self-report is distorted. The framework requires the possibility that repulsive processing can produce false self-assessments—otherwise the concept of emotional capture loses its clinical utility. Consider a person with severe anorexia nervosa who reports that further weight loss is necessary for health. Here:

- Condition 1 is satisfied: the validation path (medical evidence about the effects of starvation on organ function, cardiac stability, and mortality risk) is accessible to the person through their reflective processing capacities, even if motivational barriers prevent acceptance.

- Condition 2 is addressed: the dismissal of the person's self-report is grounded in independent evidence (physiological markers, longitudinal outcomes) that does not depend on assuming the person's processing is defective.

- Condition 3 is satisfied: the claim that weight restoration benefits the person would be made even if the person held all the power in the clinical encounter; the justification does not depend on clinician authority but on biological facts accessible from any standpoint.

The meta-criterion thus distinguishes cases where overriding self-report is warranted (independent evidence, accessible reasoning, power-symmetry) from cases where it is oppressive (circular reasoning, inaccessible premises, power-dependence).

The Meta-Criterion's Grounding in the Framework.

This meta-criterion is not an ad hoc addition. It follows from the framework's own core commitments:

First, from the graduated account of normative standing (Chapter 3, 3.5): human beings are reflective conscious processors whose flourishing depends on the exercise of their validating capacities. A system that systematically disables those capacities in some members for the benefit of others is not a genuine integrative whole; it is a pseudo-whole that extracts value from some processing systems to feed others. The meta-criterion simply makes this operational.

Second, from the carrier concept (Chapter 3, 3.2): all relationships are physically realized through the exchange of mediators. In human social systems, the primary carriers are symbols—words, gestures, reasons. A system whose justifications depend on some members being unable to process the relevant symbols (because the symbols are withheld, distorted, or declared off-limits) is a system whose carrier exchanges are degraded. It is, in the framework's own terms, a system with a repulsive dynamic at the justificatory level.

Third, from the emotional capture analysis (Chapter 3, 3.8): the framework already acknowledges that processing can be systematically distorted. The meta-criterion does not deny this; it operationalizes how to distinguish genuine from merely convenient attributions of distortion. The distinction between the

anorexia case and the gender dysphoria case demonstrates that the framework has the resources to make this distinction without collapsing into either naive credulity or authoritarian dismissal.

Limitations and Honest Acknowledgment

The meta-criterion specified here is a conceptual tool, not an algorithm. In hard cases, reasonable people applying the meta-criterion may still disagree about whether Condition 2 is satisfied—whether the independent evidence for dismissing a member's testimony is genuinely independent or merely the system's norms in scientific dress. The history of psychiatry's misuse of "independence" (recall the circularity of diagnosing drapetomania) counsels humility.

What the meta-criterion provides is not certainty but a structured form of interrogation. It tells us what questions to ask: Can the justification be validated from the standpoint of all members? Is the dismissal of a member's testimony circular? Does the justification survive power redistribution? These questions do not guarantee consensus, but they make it harder for power to masquerade as reason. And that, for a framework that aims to constrain cultural arbitrariness without pretending to value-freedom, is the appropriate standard. The alternative—a meta-criterion that mechanically adjudicates all cases—would be a fantasy of algorithmic ethics that the framework's own emphasis on reflective processing should lead us to reject. The goal is not to eliminate judgment but to ensure that judgment is answerable to constraints that are not themselves the mere expression of power. The recursive non-contradiction meta-criterion, grounded in the framework's account of processing systems and their constitutive constraints, provides such a constraint.

References

Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. (1978). Patterns of attachment: A psychological study of the strange situation. Lawrence Erlbaum.

American Psychiatric Association. (1973). Diagnostic and statistical manual of mental disorders (2nd ed., 7th printing). American Psychiatric Publishing.

American Psychiatric Association. (1980). Diagnostic and statistical manual of mental disorders (3rd ed.). American Psychiatric Publishing.

American Psychiatric Association. (1994). Diagnostic and statistical manual of mental disorders (4th ed.). American Psychiatric Publishing.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). American Psychiatric Publishing.

Aristotle. (1999). Nicomachean ethics (T. Irwin, Trans., 2nd ed.). Hackett Publishing.

Aristotle. (2001). On the soul (J. A. Smith, Trans.). In R. McKeon (Ed.), The basic works of Aristotle (pp. 535–603). Modern Library.

Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of Other in the Self Scale and the structure of interpersonal closeness. Journal of Personality and Social Psychology, 63(4), 596–612.

Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. American Psychologist, 54(7), 462–479.

Barkow, J. H., Cosmides, L., & Tooby, J. (Eds.). (1992). The adapted mind: Evolutionary psychology and the generation of culture. Oxford University Press.

Baron-Cohen, S. (1995). Mindblindness: An essay on autism and theory of mind. MIT Press.

Baron-Cohen, S. (2011). The science of evil: On empathy and the origins of cruelty. Basic Books.

Barrett, L. F. (2017). How emotions are made: The secret life of the brain. Houghton Mifflin Harcourt.

Bateson, G. (1972). Steps to an ecology of mind. University of Chicago Press.

Bateson, G. (1979). Mind and nature: A necessary unity. E. P. Dutton.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. Psychological Bulletin, 117(3), 497–529.

Beck, A. T. (1976). Cognitive therapy and the emotional disorders. International Universities Press.

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). Cognitive therapy of depression. Guilford Press.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). Manual for the Beck Depression Inventory-II. Psychological Corporation.

Bentall, R. P. (2003). Madness explained: Psychosis and human nature. Allen Lane.

Blackburn, S. (1998). Ruling passions: A theory of practical reasoning. Oxford University Press.

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. Frontiers in Public Health, 6, 149.

Bolton, D. (2008). What is mental disorder? An essay in philosophy, science, and values. Oxford University Press.

Bolton, D., & Banner, N. (2012). Does mental disorder involve a loss of personal autonomy? In L. Radoilska (Ed.), Autonomy and mental disorder (pp. 77–99). Oxford University Press.

Boorse, C. (1975). On the distinction between disease and illness. Philosophy & Public Affairs, 5(1), 49–68.

Boorse, C. (1977). Health as a theoretical concept. Philosophy of Science, 44(4), 542–573.

Boorse, C. (1997). A rebuttal on health. In J. M. Humber & R. F. Almeder (Eds.), What is disease? (pp. 1–134). Humana Press.

Bowlby, J. (1969). Attachment and loss: Vol. 1. Attachment. Basic Books.

Bowlby, J. (1973). Attachment and loss: Vol. 2. Separation: Anxiety and anger. Basic Books.

Bowlby, J. (1980). Attachment and loss: Vol. 3. Loss: Sadness and depression. Basic Books.

Buss, D. M. (1995). Evolutionary psychology: A new paradigm for psychological science. Psychological Inquiry, 6(1), 1–30.

Capra, F., & Luisi, P. L. (2014). The systems view of life: A unifying vision. Cambridge University Press.

Cacioppo, J. T., & Cacioppo, S. (2014). Social relationships and health: The toxic effects of perceived social isolation. Social and Personality Psychology Compass, 8(2), 58–72.

Cacioppo, J. T., & Patrick, W. (2008). Loneliness: Human nature and the need for social connection. W. W. Norton.

Clark, A. (1997). Being there: Putting brain, body, and world together again. MIT Press.

Clark, A., & Chalmers, D. (1998). The extended mind. Analysis, 58(1), 7–19.

Cohen, S., & Wills, T. A. (1985). Stress, social support, and the buffering hypothesis. Psychological Bulletin, 98(2), 310–357.

Conrad, P. (2007). The medicalization of society: On the transformation of human conditions into treatable disorders. Johns Hopkins University Press.

Cooper, R. (2007). Psychiatry and philosophy of science. Acumen Publishing.

Costa, P. T., & McCrae, R. R. (1992). Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual. Psychological Assessment Resources.

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. Journal of Consulting Psychology, 24(4), 349–354.

Damasio, A. R. (1994). Descartes' error: Emotion, reason, and the human brain. Putnam.

Damasio, A. R. (1999). The feeling of what happens: Body and emotion in the making of consciousness. Harcourt.

Damasio, A. R. (2010). Self comes to mind: Constructing the conscious brain. Pantheon.

Darwin, C. (1859). On the origin of species by means of natural selection. John Murray.

Darwin, C. (1872). The expression of the emotions in man and animals. John Murray.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. Journal of Personality and Social Psychology, 44(1), 113–126.

Deci, E. L., & Ryan, R. M. (1985). Intrinsic motivation and self-determination in human behavior. Plenum Press.

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. Psychological Inquiry, 11(4), 227–268.

DeVellis, R. F. (2017). Scale development: Theory and applications (4th ed.). SAGE Publications.

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction with Life Scale. Journal of Personality Assessment, 49(1), 71–75.

Diener, E., Wirtz, D., Tov, W., Kim-Prieto, C., Choi, D. W., Oishi, S., & Biswas-Diener, R. (2010). New well-being measures: Short scales to assess flourishing and positive and negative feelings. Social Indicators Research, 97(2), 143–156.

Dimidjian, S., Hollon, S. D., Dobson, K. S., Schmaling, K. B., Kohlenberg, R. J., Addis, M. E., ... & Jacobson, N. S. (2006). Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the acute treatment of major depression. Journal of Consulting and Clinical Psychology, 74(4), 658–670.

Doyal, L., & Gough, I. (1991). A theory of human need. Macmillan.

Dreyfus, H. L. (2007). Why Heideggerian AI failed and how fixing it would require making it more Heideggerian. Philosophical Psychology, 20(2), 247–268.

Eisenberg, N., & Strayer, J. (Eds.). (1987). Empathy and its development. Cambridge University Press.

Engel, G. L. (1977). The need for a new medical model: A challenge for biomedicine. Science, 196(4286), 129–136.

Foot, P. (2001). Natural goodness. Oxford University Press.

Foucault, M. (1965). Madness and civilization: A history of insanity in the age of reason (R. Howard, Trans.). Pantheon Books.

Frances, A. (2013). Saving normal: An insider's revolt against out-of-control psychiatric diagnosis, DSM-5, big pharma, and the medicalization of ordinary life. William Morrow.

Frankl, V. E. (1959). Man's search for meaning. Beacon Press.

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. American Psychologist, 56(3), 218–226.

Freud, S. (1917). Mourning and melancholia. In The standard edition of the complete psychological works of Sigmund Freud (Vol. 14, pp. 237–258). Hogarth Press.

Friston, K. (2010). The free-energy principle: A unified brain theory? Nature Reviews Neuroscience, 11(2), 127–138.

Fulford, K. W. M. (1989). Moral theory and medical practice. Cambridge University Press.

Fulford, K. W. M. (2004). Facts/values: Ten principles of values-based medicine. In J. Radden (Ed.), The philosophy of psychiatry: A companion (pp. 205–234). Oxford University Press.

Gallagher, S. (2005). How the body shapes the mind. Oxford University Press.

Gibbard, A. (1990). Wise choices, apt feelings: A theory of normative judgment. Harvard University Press.

Gilbert, P. (2009). The compassionate mind: A new approach to life's challenges. Constable.

Goetz, J. L., Keltner, D., & Simon-Thomas, E. (2010). Compassion: An evolutionary analysis and empirical review. Psychological Bulletin, 136(3), 351–374.

Greenberg, L. S. (2002). Emotion-focused therapy: Coaching clients to work through their feelings. American Psychological Association.

Haidt, J. (2006). The happiness hypothesis: Finding modern truth in ancient wisdom. Basic Books.

Harlow, H. F. (1958). The nature of love. American Psychologist, 13(12), 673–685.

Haslam, N., Holland, E., & Kuppens, P. (2012). Categories versus dimensions in personality and psychopathology: A quantitative review of taxometric research. Psychological Medicine, 42(5), 903–920.

Hayes, S. C., & Smith, S. (2005). Get out of your mind and into your life: The new acceptance and commitment therapy. New Harbinger Publications.

Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (1999). Acceptance and commitment therapy: An experiential approach to behavior change. Guilford Press.

Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (2012). Acceptance and commitment therapy: The process and practice of mindful change*(2nd ed.). Guilford Press.

Hoffman, M. L. (2000). Empathy and moral development: Implications for caring and justice. Cambridge University Press.

Hofmann, S. G., Asnaani, A., Vonk, I. J., Sawyer, A. T., & Fang, A. (2012). The efficacy of cognitive behavioral therapy: A review of meta-analyses. Cognitive Therapy and Research, 36(5), 427–440.

Hofmann, S. G., & Hay, A. C. (2018). Rethinking avoidance: Toward a balanced approach to avoidance in treating anxiety disorders. Journal of Anxiety Disorders, 55, 14–21.

Holt-Lunstad, J., Smith, T. B., & Layton, J. B. (2010). Social relationships and mortality risk: A meta-analytic review. PLoS Medicine, 7(7), e1000316.

Horwitz, A. V., & Wakefield, J. C. (2007). The loss of sadness: How psychiatry transformed normal sorrow into depressive disorder. Oxford University Press.

Hume, D. (1739/2000). A treatise of human nature (D. F. Norton & M. J. Norton, Eds.). Oxford University Press.

Hursthouse, R. (1999). On virtue ethics. Oxford University Press.

Huta, V., & Waterman, A. S. (2014). Eudaimonia and its distinction from hedonia: Developing a classification and terminology for understanding conceptual and operational definitions. Journal of Happiness Studies, 15(6), 1425–1456.

Insel, T. R. (2014). The NIMH Research Domain Criteria (RDoC) project: Precision medicine for psychiatry. American Journal of Psychiatry, 171(4), 395–397.

Jablonka, E., & Lamb, M. J. (2005). Evolution in four dimensions: Genetic, epigenetic, behavioral, and symbolic variation in the history of life. MIT Press.

Kabat-Zinn, J. (1990). Full catastrophe living: Using the wisdom of your body and mind to face stress, pain, and illness. Delacorte.

Kahneman, D. (2011). Thinking, fast and slow. Farrar, Straus and Giroux.

Kashdan, T. B., & Rottenberg, J. (2010). Psychological flexibility as a fundamental aspect of health. Clinical Psychology Review, 30(7), 865–878.

Kaufman, S. B., Yaden, D. B., Hyde, E., & Tsukayama, E. (2019). The light vs. dark triad of personality: Contrasting two very different profiles of human nature. Frontiers in Psychology, 10, 467.

Keltner, D., & Haidt, J. (2003). Approaching awe, a moral, spiritual, and aesthetic emotion. Cognition and Emotion, 17(2), 297–314.

Kendler, K. S. (2012). The dappled nature of causes of psychiatric illness: Replacing the organic–functional/hardware–software dichotomy with empirically based pluralism. Molecular Psychiatry, 17(4), 377–388.

Keyes, C. L. M. (2002). The mental health continuum: From languishing to flourishing in life. Journal of Health and Social Behavior, 43(2), 207–222.

Keyes, C. L. M. (2007). Promoting and protecting mental health as flourishing: A complementary strategy for improving national mental health. American Psychologist, 62(2), 95–108.

King, M., Semlyen, J., Tai, S. S., Killaspy, H., Osborn, D., Popelyuk, D., & Nazareth, I. (2008). A systematic review of mental disorder, suicide, and deliberate self harm in lesbian, gay and bisexual people. BMC Psychiatry, 8, 70.

Korsgaard, C. M. (1996). The sources of normativity. Cambridge University Press.

Korsgaard, C. M. (2009). Self-constitution: Agency, identity, and integrity. Oxford University Press.

Krueger, R. F., Derringer, J., Markon, K. E., Watson, D., & Skodol, A. E. (2012). Initial construction of a maladaptive personality trait model and inventory for DSM-5. Psychological Medicine, 42(9), 1879–1890.

Lambert, M. J. (2013). Bergin and Garfield's handbook of psychotherapy and behavior change (6th ed.). John Wiley & Sons.

Leary, M. R., & Baumeister, R. F. (2000). The nature and function of self-esteem: Sociometer theory. Advances in Experimental Social Psychology, 32, 1–62.

Lee, R. M., & Robbins, S. B. (1995). Measuring belongingness: The Social Connectedness and the Social Assurance scales. Journal of Counseling Psychology, 42(2), 232–241.

Levinas, E. (1969). Totality and infinity: An essay on exteriority (A. Lingis, Trans.). Duquesne University Press.

Linehan, M. M. (1993). Cognitive-behavioral treatment of borderline personality disorder. Guilford Press.

Ludwig, A. M. (1975). The psychiatrist as physician. JAMA, 234(6), 603–604.

MacIntyre, A. (1981). After virtue: A study in moral theory. University of Notre Dame Press.

MacIntyre, A. (1999). Dependent rational animals: Why human beings need the virtues. Open Court.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. Psychological Review, 98(2), 224–253.

Maslow, A. H. (1968). Toward a psychology of being (2nd ed.). Van Nostrand Reinhold.

Maturana, H. R., & Varela, F. J. (1980). Autopoiesis and cognition: The realization of the living. D. Reidel Publishing.

Mayer, F. S., & Frantz, C. M. (2004). The connectedness to nature scale: A measure of individuals' feeling in community with nature. Journal of Environmental Psychology, 24(4), 503–515.

McAdams, D. P., & de St. Aubin, E. (1992). A theory of generativity and its assessment through self-report, behavioral acts, and narrative themes in autobiography. Journal of Personality and Social Psychology, 62(6), 1003–1015.

McNally, R. J. (2011). What is mental illness? Harvard University Press.

Meyer, I. H. (2003). Prejudice, social stress, and mental health in lesbian, gay, and bisexual populations: Conceptual issues and research evidence. Psychological Bulletin, 129(5), 674–697.

Mikulincer, M., & Shaver, P. R. (2007). Attachment in adulthood: Structure, dynamics, and change. Guilford Press.

Mill, J. S. (1863). Utilitarianism. Parker, Son, and Bourn.

Moore, G. E. (1903). Principia ethica. Cambridge University Press.

Murphy, D. (2006). Psychiatry in the scientific image. MIT Press.

Naess, A. (1973). The shallow and the deep, long-range ecology movement: A summary. Inquiry, 16(1–4), 95–100.

Nesse, R. M. (1990). The evolutionary functions of repression and the ego defenses. Journal of the American Academy of Psychoanalysis, 18(2), 260–285.

Nesse, R. M., & Williams, G. C. (1994). Why we get sick: The new science of Darwinian medicine. Vintage Books.

Nisbet, E. K., Zelenski, J. M., & Murphy, S. A. (2009). The Nature Relatedness Scale: Linking individuals' connection with nature to environmental concern and behavior. Environment and Behavior, 41(5), 715–740.

Norcross, J. C., & Wampold, B. E. (2011). Evidence-based therapy relationships: Research conclusions and clinical practices. Psychotherapy, 48(1), 98–102.

Nussbaum, M. C. (2000). Women and human development: The capabilities approach. Cambridge University Press.

Nussbaum, M. C. (2011). Creating capabilities: The human development approach. Harvard University Press.

O'Neill, J. (2008). Happiness and the good life. Environmental Values, 17(2), 125–144.

Oshana, M. (2006). Personal autonomy in society. Ashgate Publishing.

Panksepp, J. (1998). Affective neuroscience: The foundations of human and animal emotions. Oxford University Press.

Panksepp, J., & Biven, L. (2012). The archaeology of mind: Neuroevolutionary origins of human emotions. W. W. Norton.

Peterson, C., & Seligman, M. E. P. (2004). Character strengths and virtues: A handbook and classification. Oxford University Press.

Prigogine, I., & Stengers, I. (1984). *Order out of chaos: Man's new dialogue with nature*. Bantam Books.

Rawls, J. (1971). A theory of justice. Harvard University Press.

Rogers, C. R. (1961). On becoming a person: A therapist's view of psychotherapy. Houghton Mifflin.

Rogers, C. R. (1963). The actualizing tendency in relation to "motives" and to consciousness. In M. R. Jones (Ed.), Nebraska symposium on motivation (Vol. 11, pp. 1–24). University of Nebraska Press.

Rosch, E. (1999). Principles of categorization. In E. Margolis & S. Laurence (Eds.), Concepts: Core readings (pp. 189–206). MIT Press.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. American Psychologist, 55(1), 68–78.

Ryan, R. M., & Deci, E. L. (2001). On happiness and human potentials: A review of research on hedonic and eudaimonic well-being. Annual Review of Psychology, 52, 141–166.

Ryan, R. M., & Deci, E. L. (2017). Self-determination theory: Basic psychological needs in motivation, development, and wellness. Guilford Press.

Ryff, C. D. (1989). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. Journal of Personality and Social Psychology, 57(6), 1069–1081.

Ryff, C. D., & Singer, B. H. (2008). Know thyself and become what you are: A eudaimonic approach to psychological well-being. Journal of Happiness Studies, 9(1), 13–39.

Sadler, J. Z. (2005). Values and psychiatric diagnosis. Oxford University Press.

Schroeder, M. (2008). Being for: Evaluating the semantic program of expressivism. Oxford University Press.

Seligman, M. E. P. (2002). Authentic happiness: Using the new positive psychology to realize your potential for lasting fulfillment. Free Press.

Seligman, M. E. P. (2011). Flourish: A visionary new understanding of happiness and well-being. Free Press.

Sen, A. (1999). Development as freedom. Alfred A. Knopf.

Shapiro, S. L., Carlson, L. E., Astin, J. A., & Freedman, B. (2006). Mechanisms of mindfulness. Journal of Clinical Psychology, 62(3), 373–386.

Shedler, J. (2010). The efficacy of psychodynamic psychotherapy. American Psychologist, 65(2), 98–109.

Singer, T., & Klimecki, O. M. (2014). Empathy and compassion. Current Biology, 24(18), R875–R878.

Spitzer, R. L. (1981). The diagnostic status of homosexuality in DSM-III: A reformulation of the issues. American Journal of Psychiatry, 138(2), 210–215.

Spitzer, R. L., & Endicott, J. (1978). Medical and mental disorder: Proposed definition and criteria. In R. L. Spitzer & D. F. Klein (Eds.), Critical issues in psychiatric diagnosis (pp. 15–39). Raven Press.

Steger, M. F., Frazier, P., Oishi, S., & Kaler, M. (2006). The Meaning in Life Questionnaire: Assessing the presence of and search for meaning in life. Journal of Counseling Psychology, 53(1), 80–93.

Szasz, T. S. (1961). The myth of mental illness: Foundations of a theory of personal conduct. Harper & Row.

Tafarodi, R. W., & Swann, W. B. (2001). Two-dimensional self-esteem: Theory and measurement. Personality and Individual Differences, 31(5), 653–673.

Teasdale, J. D., Segal, Z. V., Williams, J. M. G., Ridgeway, V. A., Soulsby, J. M., & Lau, M. A. (2000). Prevention of relapse/recurrence in major depression by mindfulness-based cognitive therapy. Journal of Consulting and Clinical Psychology, 68(4), 615–623.

Thompson, M. (2008). Life and action: Elementary structures of practice and practical thought. Harvard University Press.

Thompson, E. (2007). Mind in life: Biology, phenomenology, and the sciences of mind. Harvard University Press.

Tomasello, M. (2009). Why we cooperate. MIT Press.

Tomasello, M. (2016). A natural history of human morality. Harvard University Press.

Varela, F. J., Thompson, E., & Rosch, E. (1991). The embodied mind: Cognitive science and human experience. MIT Press.

Wakefield, J. C. (1992). The concept of mental disorder: On the boundary between biological facts and social values. American Psychologist, 4(3), 373–388.

Wakefield, J. C. (1999). Evolutionary versus prototype analyses of the concept of disorder. Journal of Abnormal Psychology, 108(3), 374–399.

Wakefield, J. C. (2007). The concept of mental disorder: Diagnostic implications of the harmful dysfunction analysis. World Psychiatry, 6(3), 149–156.

Wakefield, J. C., & First, M. B. (2012). Placing symptoms in context: The role of contextual criteria in reducing false positives in DSM diagnosis. Comprehensive Psychiatry, 53(2), 130–139.

Wampold, B. E. (2015). How important are the common factors in psychotherapy? An update. World Psychiatry, 14(3), 270–277.

Wampold, B. E., & Imel, Z. E. (2015). The great psychotherapy debate: The evidence for what makes psychotherapy work (2nd ed.). Routledge.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. Journal of Personality and Social Psychology, 54(6), 1063–1070.

White, P. D., Rickards, H., & Zeman, A. Z. J. (2012). Time to end the distinction between mental and neurological illnesses. BMJ, 344, e3454.

Williams, B. (1985). Ethics and the limits of philosophy. Harvard University Press.

Wilson, E. O. (1998). Consilience: The unity of knowledge. Alfred A. Knopf.

Wilson, E. O. (2012). The social conquest of earth. W. W. Norton.

Winnicott, D. W. (1965). The maturational processes and the facilitating environment: Studies in the theory of emotional development. International Universities Press.

World Health Organization. (2004). Promoting mental health: Concepts, emerging evidence, practice. WHO Press.

World Health Organization. (2022). World mental health report: Transforming mental health for all. WHO Press.

Zachar, P. (2014). A metaphysics of psychopathology. MIT Press.

Zachar, P., & Kendler, K. S. (2007). Psychiatric disorders: A conceptual taxonomy. American Journal of Psychiatry, 164(4), 557–565.

Zelenski, J. M., & Nisbet, E. K. (2014). Happiness and feeling connected: The distinct role of nature relatedness. Environment and Behavior, 46(1), 3–23.