Studying Mental Health Problems as Systems, Not Syndromes

Eiko I. Fried
Department of Clinical Psychology, Leiden University

Abstract
Despite decades of clinical, sociopolitical, and research efforts, progress in understanding and treating mental health problems remains disappointing. I discuss two barriers that have contributed to a problematic oversimplification of mental illness. The first is diagnostic literalism, mistaking mental health problems (complex within-person processes) for the diagnoses by which they are classified (clinically useful idealizations to facilitate treatment selection and prognosis). The second is reductionism, the isolated study of individual elements of mental disorders. I propose conceptualizing people’s mental health states as outcomes emerging from complex systems of biological, psychological, and social elements and show that this systems perspective explains many robust phenomena, including variability within diagnoses, comorbidity among diagnoses, and transdiagnostic risk factors. It helps us understand diagnoses and reductionism as useful epistemological tools for describing the world, rather than ontological convictions about how the world is. It provides new lenses through which to study mental illness (e.g., attractor states, phase transitions), and new levers to treat them (e.g., early warning signals, novel treatment targets). Embracing the complexity of mental health problems requires opening our ivory towers to theories and methods from other fields with rich traditions, including network and systems sciences.

Keywords
biological reductionism, complex systems, diagnostic literalism, emergence, mental disorders, psychiatric nosology, reductionism

O chestnut tree, great rooted blossomer
Are you the leaf, the blossom or the bole?
—W. B. Yeats (“Among School Children,” 1928)

Psychiatric classifications systems like the Diagnostic and Statistical Manual of Mental Disorders (DSM; American Psychiatric Association, 2013) or the International Statistical Classification of Diseases (ICD; World Health Organization, 2019) categorize mental health problems into diagnoses such as schizophrenia and posttraumatic stress disorder (PTSD). These diagnoses are further described by criteria, often specific symptoms. The history of defining the diagnosis for major depressive disorder (MDD) helps to clarify the crucial distinction between mental health problems and diagnoses by which they are classified.

DSM’s fifth edition (DSM-5) specifies that an episode of MDD can be diagnosed if a person meets at least five of nine symptoms for 2 weeks and also shows considerable impairment of functioning. What data justify these particular criteria? In an effort to introduce the medical model to psychiatry, Feighner et al. (1972) delineated what have come to be called the Feighner criteria: observable signs and symptoms for common mental disorders. For determining the symptoms of MDD, the authors relied on a 1957 article by Cassidy and colleagues. The Feighner criteria were highly influential for DSM-3, and today, all MDD symptoms proposed originally by Cassidy, with one exception (constipation), are preserved in DSM-5, with a similar threshold for clinical significance (DSM-5: at least five of 9 symptoms; Cassidy: at least six of 10 symptoms). When Cassidy was asked in 1980 about his reasoning behind these criteria, he responded that “it sounded about right” (Kendler et al., 2010, p. 136).

Corresponding Author:
Eiko I. Fried, Department of Clinical Psychology, Leiden University
Email: eikofried@gmail.com
Similar developments can be observed for other diagnoses. Many of DSMs iterations, such as DSM-5's transition from a categorical to a dimensional model of personality disorders, were based on data. But psychiatric nosology is “a pragmatic compromise among multiple competing demands and constituencies” (Lilienfeld, 2014, p. 269) and has been driven by adherence to precedent, patient advocacy, lobbying efforts, and goals to minimize stigma and avoid excessive prevalence rates. These sociopolitical forces have shaped the DSM, as have historical forces and path dependence (Lilienfeld, Scull, 2021). Feighner, for example, could have relied on one of the many frameworks competing with Cassidy’s views, which would have eventually led to a somewhat different DSM-5. And if Wernicke (an influential competitor of one of the founding fathers of modern psychiatry, Kraepelin) had not died prematurely in a bicycle accident, psychiatric nosology as a whole might look somewhat different today (Kendler, 2016a).

Diagnostic Literalism

Mental health problems are complex, biopsychosocial processes that unfold in individuals over time. In contrast, diagnoses are categorical idealizations designed as clinically useful tools to facilitate communication, accurate prognosis, and treatment selection and planning. Mental health problems can be described as diagnoses, and such description can be immensely helpful for research and clinical practice, but mental health problems and diagnoses are not the same kind of thing. Conflating the two is called diagnostic literalism (also referred to as reification or essentialism; Adriaens & De Block, 2013; Kendler, 2016b; Kendler et al., 2011; Zachar, 2014), that is, taking diagnoses for more than they are.

The following eight observations help clarify the difference between mental health problems and diagnoses and are expected results from superimposing clinical idealizations on the complex landscape of mental disorders (Hyman, 2021).

1. For most diagnoses, DSM ignores causes and etiology.
2. Classification systems such as DSM and ICD differ considerably in their conceptualization of some diagnoses (e.g., PTSD), and there are dozens of different measurement tools to diagnose the same disorder (e.g., MDD; Fried et al., 2022).
3. Interrater reliability for some common diagnoses is low (Regier et al., 2013).
4. People with the same diagnosis often have some similarities (e.g., in terms of etiology and symptoms).
5. However, they also show considerable differences (e.g., Galatzer-Levy & Bryant, 2013).
6. There is considerable comorbidity between diagnoses, and many risk factors are transdiagnostic (i.e., shared among diagnoses; Eaton et al., 2015; Kessler et al., 2005).
7. Although diagnoses are categorical, most mental health problems are best described as lying on a dimension of severity from absent to very severe—and various thresholds where normal functioning turns into mental illness have been proposed (Haslam et al., 2012; von Glischinski et al., 2021).
8. Finally, disease pathways can be characterized by equifinality (different starting points may lead to the same diagnosis) and multifinality (similar starting points may lead to different diagnoses).

Diagnoses do not carve nature at her joints, but are pragmatic kinds: things constructed (in part on the basis of evidence) to be useful for specific purposes (Kendler et al., 2011). Elements listed in psychiatry’s DSM are different from elements listed in chemistry’s equivalent of the DSM, the periodic table. Helium and magnesium are natural kinds, unchanging entities with necessary and sufficient properties that clearly define them. Every atom with two protons is helium, and helium’s internal structure, not expert consensus, defines kind membership. Unlike MDD and schizophrenia, helium has only one definition, can be diagnosed with perfect reliability, and can be clearly distinguished from other kinds, such as magnesium.

That does not mean that psychiatric diagnoses have no utility. Similar to MDD and PTSD, long-term high blood pressure is associated with numerous adverse outcomes. But because blood pressure lies on a continuum, the definition of “abnormal” blood pressure remains somewhat arbitrary and is the result of expert consensus taking into account nonscientific factors (Unger et al., 2020), just as the criteria for MDD and PTSD are. Even the decision of the International Astronomical Union to remove Pluto from the list of planets was based in part on somewhat arbitrary definitions of what a planet ought to be, influenced by nonscientific factors, and heavily opposed by some astronomers (Zachar & Kendler, 2012).

Unfortunately, clinical psychology and psychiatry have devoted most of their resources to studying the diagnostic labels that summarize the complex mental health states of people, rather than how biopsychosocial processes give rise to mental health problems. The literature is dominated by case-control studies, in which a healthy control group is compared with a group diagnosed with one specific mental disorder; such studies have resulted
Reductionism

Another roadblock to better understanding, predicting, and treating mental health problems has been the isolated study of particular parts of mental disorders. A bicycle is a simple mechanical system consisting of parts that relate to each other: A pedal moves a cogwheel, which moves a chain, which moves another cogwheel, and so on (A → B → C → D). To fully understand how this system functions at the macro level, it is sufficient to investigate its constituent components at the micro level. This approach can be understood as reductionism: figuring out the properties of the whole given the properties of its parts (Pessoa, 2022, Chapter 8). It works well in this example because characteristics of bicycle parts remain unaltered when investigated in isolation. Although reductionism has been astonishingly successful in science, its limits became clear in the 20th century when scholars tried to understand increasingly complex systems, such as the stock market, the weather, the Internet (Barabási, 2012; Von Bertalanffy, 1972), or—as I argue here—mental disorders. This is because complex systems contain interdependent elements whose properties depend on each other, and thus call for the study of system parts and the relationships among those parts (Fried et al., 2022).

Further, the biopsychosocial nature of mental illness requires investigating parts and relationships across levels. A particular barrier to progress has been biological or explanatory reductionism—the idea that lower levels (i.e., biology) offer explanatory power inherently superior to that of higher levels (i.e., psychology, environment). Explanatory reductionism has dominated the research landscape since the 1980s and constrained research funds, health-care policy, and delivery of clinical services (Borsboom et al., 2019; Miller, 2010; Scull, 2021). Highly influential strategic directives from the National Institute of Mental Health (NIMH) are good examples: They have stated that mental disorders are “brain disorders” (Insel & Cuthbert, 2015, p. 499), “brain-circuit disorders” (Insel & Cuthbert, 2015, p. 500), or “dysfunctions in neural circuits” (Insel et al., 2010, p. 749) that can be “identified with tools of clinical neuroscience” (Insel et al., 2010, p. 749), treated by “tuning these [neural] circuits” (Insel & Cuthbert, 2015, p. 500), and ultimately better understood through “neuroscience-based psychiatric classification” (Insel et al., 2010, p. 750). This perspective moves every aspect of mental health—psychiatric nosology, individual diagnosis, dysfunction, and treatment—to the level of the brain and is apparent in many of NIMH’s efforts, such as early work on the Research Domain Criteria (RDoC; Insel et al., 2010).

Explanatory reductionism faces a number of challenges (Borsboom et al., 2019; Eronen, 2021). First, lower levels are not by definition superior in explaining higher-level processes: “No one would seek to understand the origin of hypertension at the level of quarks” (Kendler, 2005, p. 1249). Instead, output from complex systems often requires higher-level explanations. The human heartbeat, for example, results from feedback loops among properties such as the joint activity of ion channels and cell potential, not from a central control unit (Eronen, 2021). Second, one may find biomarkers for a mental health problem, such as delusions, but not the content of the delusions (e.g., bizarre vs. grandiose), not their biological instantiation, drives feelings (e.g., fear vs. elation) and behaviors (e.g., fleeing vs. approaching). Third, complex phenotypes such as grandiose delusions are likely multiply realizable; that is, they likely differ in brain activation across (and perhaps within) individuals.

Biological psychiatry has led to considerable insights into human biology but has told us relatively little about the biology of specific diagnoses. The American Psychiatric Association recently concluded that “neuroimaging has yet to have a significant impact on the diagnosis or treatment of individual patients in clinical settings” (First et al., 2018, p. 915), and genome-wide association studies have largely resulted in transdiagnostic (rather than diagnosis-specific) hits that explain negligible amounts of variance (Scull, 2021). Even NIMH director (2002–2015) Insel, one of the most influential voices for studying the biology of mental illness, concluded that after

13 years at NIMH really pushing on the neuroscience and genetics of mental disorders, [although] I succeeded at getting lots of really cool papers published by cool scientists at fairly large costs—I think $20 billion—I don't think we moved the needle in reducing suicide, reducing hospitalizations, improving recovery for the tens of millions of people who have mental illness. (quoted in Rogers, 2017, para. 6)

This lack of progress is not due to the fact that biology is not crucially involved in mental health states—it is due to our focus on studying the biology of particular DSM labels that are likely the wrong targets, and due
Fig. 1. Conceptualizing mental health problems as systems. The schematic in (a) represents 20 hypothetical people diagnosed with post-traumatic stress disorder (PTSD; red) and major depressive disorder (MDD; blue). Their mental health states are characterized by three features, X, Y, and Z, as a simplification of a much larger, n-dimensional feature space. The two groups differ moderately on feature X, differ greatly on feature Y, and do not differ appreciably on feature Z; within-group variability on feature Z is much greater for the people with MDD than for those with PTSD. From a systems perspective, the mental health states of these people are emergent properties arising out of interactions of numerous biopsychosocial features (e.g., risk factors, etiology, thoughts, biological predispositions, social environments). Mental health states cluster to some degree because features stand in probabilistic relations to each other (e.g., severe trauma often leads to nightmares). Superimposing diagnoses on this complex landscape leads to one person per group who is most typical of the diagnosis (black circles), but also ignores meaningful interindividual differences. The diagram in (b) uses two copies of the same playing card, held between two fingers, as a model for vulnerable and resilient mental health states. The left card is in a vulnerable state and can be moved with little energy to an alternative state. The right card is in a resilient state, and forcing a phase transition to an alternative state requires considerable energy.
to studying biology in isolation (Cai et al., 2020; Eronen, 2021; Hitchcock et al., 2022).

Diagnostic literalism and explanatory reductionism have formed a vicious cycle of reification. We study the genomes of millions of participants diagnosed via DSM-5 when it is unclear why these phenotypes would be well suited for genetic discovery in the first place (Cai et al., 2020; Hitchcock et al., 2022). After identifying weak correlates, we further reify diagnoses by talking about “genes for major depression” or “brain biomarkers for PTSD.” Our natural tendency to essentialize mental disorders (Adriaens & De Block, 2013) and flawed inferences from measurement (“we have a commonly used measure for diagnosis X; hence X exists”; cf. McPherson & Armstrong, 2021) and external validation (“diagnosis X correlates with external constructs; hence diagnosis X exists”) provide fertile ground for this vicious cycle of reification.

**Mental Disorders as Complex Systems**

In the summer of 2019, a scholar I greatly admire was kind enough to lend me his bicycle for a few months, granted I take good care of it. When the bicycle broke down after 3 weeks, I was terribly worried, but reductionism came to the rescue: Bikes can be decomposed into their constituent parts, and fixing all parts at the micro level will restore function at the macro level. But mental disorders are not like bicycles—they are like many other complex systems in nature. Whether a lake is clean or turbid results from interactions of interdependent elements, such as oxygen levels, sunlight exposure, fish, pollution, and so on. Whether my mood while writing this manuscript is anxious or cheerful is the outcome of causal relations among elements of my mood system, including my personality and disposition; the previous night's sleep; my caffeine consumption; and external influences such as my email inbox. The same applies to mental health states. From a systems perspective (see Fig. 1a), such states result from interactions of numerous biological, psychological, and social features, including specific risk and protective factors, moods, thoughts, behaviors, biological predispositions, and social environments (Borsboom, 2017; Fried et al., 2022; Kendler et al., 2011; Olthof et al., 2021; Robinaugh et al., 2020).

This framework helps to explain the eight observations I described above. Some simply follow from acknowledging that diagnoses are superimposed on a complex landscape of mental health states: They may ignore important features such as etiology (Observation 1) and summarize dimensional processes as categorical (Observation 7); multiple potentially equally valid summaries and measures can coexist (Observation 2); and interrater reliability will necessarily be limited for some diagnoses (Observation 3).

This systems perspective also explains why mental health states are somewhat clustered in nature (Observation 4) rather than completely randomly distributed. Some states are much more likely than others because features are linked in probabilistic ways. For example, severe traumatic experiences often lead to nightmares, which lead to further sleep problems and then interfere with daily activities. If trauma is severe enough, most people show some level of impairment, in the same way that very few lakes with severe pollution show healthy fish populations. In other words, feature interdependence constrains the states systems can be in. Some people in the feature space will be more typical of a given diagnosis than others are. These people can be thought to be in the most likely mental health state, given that they reflect the average result of all probabilistic feature relations. Probabilistic feature relations allow for considerable interindividual differences within a given DSM diagnosis (Observation 5), as well as equifinality and multifinality (Observation 8).

This framework can also accommodate two diagnoses being both similar and different (Observations 4 and 5). In Figure 1a, PTSD shows much less variability than MDD on feature Z (e.g., the amount of adverse life events experienced in the last year). The degree of similarity and difference between diagnoses can vary across features: In the figure, PTSD and MDD differ strongly on feature Y (e.g., flashbacks are much more common in PTSD), moderately on feature X (e.g., MDD has somewhat higher comorbidity rates with generalized anxiety disorder), and not at all on feature Z. If one were to plot the mental health state of millions of people and superimpose the most prevalent DSM diagnoses, some point clouds would overlap substantially because of shared risk factors, symptoms, or pathophysiology, whereas others would be somewhat separated. This explains comorbidities and shared risk factors among DSM diagnoses (Observation 6). Splitting or lumping such diagnoses is a pragmatic decision guided by clinical utility, and I understand DSM and ICD, as well as recent frameworks, such as the Hierarchical Taxonomy of Psychopathology (HiTOP; a hierarchical, dimensional taxonomy), as competing attempts to adequately summarize this space.

From this perspective, it is unsurprising that common one-size-fits-all treatments that target diagnoses (e.g., antidepressants for MDD) have shown limited efficacy, and equally unsurprising that biological psychiatry has been largely unsuccessful in identifying specific markers for specific diagnoses. Further, symptoms lose their privileged epistemological status as the criteria for studying mental disorders, given that many other
features are also probabilistically associated with diagnoses. In other words, symptoms are not inherently superior features compared with etiology, personality, and other factors.

A Systems Perspective Offers New Lenses and Levers

A systems perspective provides new lenses through which we can study mental health problems: by using theories and methods from other fields with rich traditions, including network and systems sciences (Barabási, 2012; Von Bertalanffy, 1972). I introduce a few concepts below that may help to advance understanding of mental health problems.

In complex systems, order can emerge at the macro level as a result of local interdependence among features at the micro level. Flocking of birds is an example of an emergent phenomenon arising from self-organization (Olfhof et al., 2021). Such emergent properties cannot be fully understood by studying isolated elements of a system. Bird flocking cannot be identified at the level of individual birds, in the same way that fluidity of water is not a property of its constituent elements, hydrogen and oxygen. Even systems following very simple rules, such as prey and predator populations that simultaneously depend on each other (A ⟷ B), can show emergence (Pessoa, 2022, Chapter 8). Ignoring the interdependence of features at the micro level, such as the well-established feedback loop between fear and avoidance relevant to many anxiety disorders, undermines our ability to understand, predict, and treat mental health problems that arise from interactions among features (Robinaugh et al., 2020).

Complex systems can have two (or more) attractor states in which they can settle (Scheffer et al., 2018; van de Leemput et al., 2014). Figure 1b shows two systems, each consisting of a playing card held between two fingers: One system is vulnerable (i.e., prone to a phase transition into an alternative attractor state), whereas the other system is resilient, as indicated by the different lengths of the horizontal black arrows. Lake states (clean, turbid) and mental health states (healthy, depressed) work in similar ways. These systems experience random perturbations, such as a very hot week or a negative life event. As a result of these perturbations, a system state may move around slightly, but as long as it does not leave the safe operating space—which is much smaller for vulnerable than for resilient systems—the system stays in the current attractor state. If perturbations are severe enough to push a system out of its safe operating space, the system may transition into the alternative state via phase transition. Removing the force that caused a phase transition does not always return the system to the prior state—a phenomenon known as the hysteresis effect (Cramer et al., 2016). For example, pollution can cause a lake to tip into a turbid state, but removing the pollution will not necessarily return the lake to a healthy state, in the same way that the playing cards in Figure 1b will not return to their previous states after the forces that pushed them are removed.

A systems perspective also offers new levers. When a system become less resilient over time, this can be measured by how long it takes the system to recover from random perturbations. If we pushed against the playing cards in Figure 1b and took a slow-motion video, we would see that the card in the vulnerable system would “snap back” more slowly, because there is less “grip” on the card to keep it in its current state compared with the card in the resilient system. When systems snap back more slowly because they become more vulnerable, this is called critical slowing down, and it is one of many early warning signals that may foreshadow upcoming transitions into mental disorders (Olfhof et al., 2020; van de Leemput et al., 2014; Wichers et al., 2016). Such signals may facilitate timely prevention and intervention.

A systems view also allows researchers to utilize frameworks such as control theory (Henry et al., 2022) to simulate interventions, much as meteorologists simulate what would happen to Earth’s climate under potential interventions, such as reduced CO₂ emissions. Such simulations may reveal more optimal intervention strategies or completely novel intervention targets, or they may showcase what interventions are most effective in what kind of systems (Henry et al., 2022).

The idea of mental health states as complex systems aligns closely with how many clinicians think about and treat mental illness (Schiepek, 2009). Functional analysis is commonly employed in psychotherapy to map out causal relations of a person’s system, and clinical psychologists have long made use of the concept of network destabilization. Suppose the resilient system in Figure 1b is in a psychopathological attractor state such as PTSD: Destabilizing the attractor state (i.e., softening the grip on the playing card) will make it much easier to help the system reorganize into a healthy state (Fried et al., 2022; Hayes et al., 2015; Olfhof et al., 2021). A large-scale consortium in The Netherlands is currently investigating the efficacy of such system-based interventions (Roefs et al., 2022).

Conclusions

A systems view casts diagnoses and reductionism as useful epistemological tools for describing the world, rather than as ontological convictions about how the world is. It challenges the view of different diagnoses
as clearly separable entities with single-cause etiologies (Kendler, 2012).

Many promising efforts toward embracing the complexity of mental health problems are on the way. For example, although NIMH's RDoC initiative started as a poster child of explanatory reductionism (Insel et al., 2010), it has quickly matured into a truly integrative framework focused on neurobiological, psychological, developmental, and environmental elements of mental health systems. RDoC and other recent initiatives (e.g., Robinaugh et al., 2020; Roefs et al., 2022) bring into sharper focus open questions. What are the right levels to study in mental health systems, what are the right elements of these respective levels, and how do elements of different levels interact with each other (Eronen, 2021)? And if we accept that diagnoses are pragmatic summaries, what validators of clinical utility should be used in future classification efforts? Answering these and other questions will require building interdisciplinary bridges and opening the ivory towers of clinical elements of these respective levels, and how do elements of different levels interact with each other (Eronen, 2021)?

Recommended Reading

Eronen, M. I. (2021). (See References). A primer on problems of explanatory reductionism and how levels in psychopathology relate to each other.
Kendler, K. S. (2016a). (See References). An article challenging the idea that diagnoses carve nature at her joints and introducing the distinction between different kinds (socially constructed, natural, pragmatic) used to describe mental disorders.


Transparency

Action Editor: Robert L. Goldstone
Editor: Robert L. Goldstone

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Funding

E. I. Fried is supported by funding from the European Research Council under the European Union’s Horizon 2020 research and innovation program, Grant 940059.

ORCID iD

Eiko I. Fried https://orcid.org/0000-0001-7469-594X

Acknowledgments

I thank Rany Abend and Diane O’Leary for comments on prior versions of this manuscript.

References

First, M. B., Drevets, W. C., Carter, C., Dickstein, D. P., Kasoff, L., Kim, K. L., McConathy, J., Rauch, S., Saad, Z. S., Savitz, J.,


