Psychiatric nosology and the DSM

Klaas Enno Stephan

Translational Neuromodeling Unit

Universität Zürich[UZH]

ETH

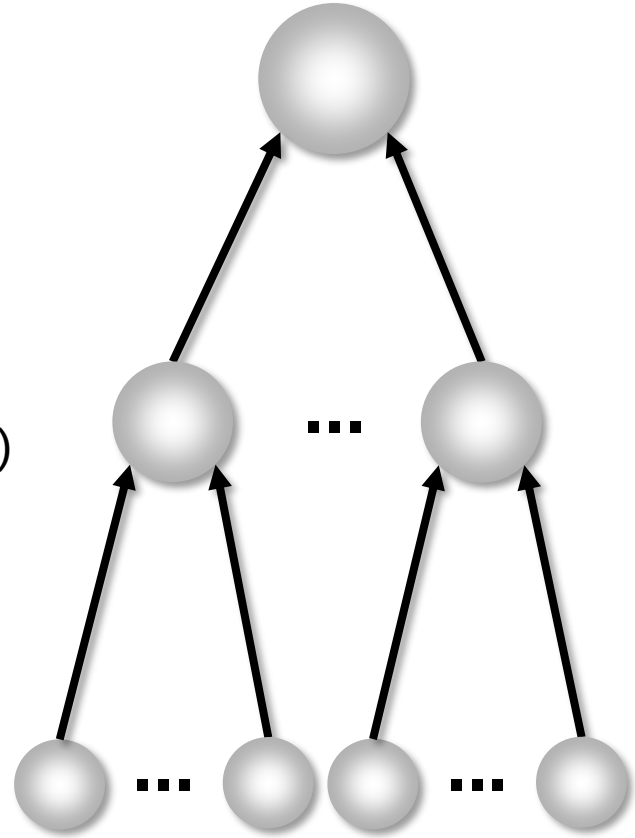
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

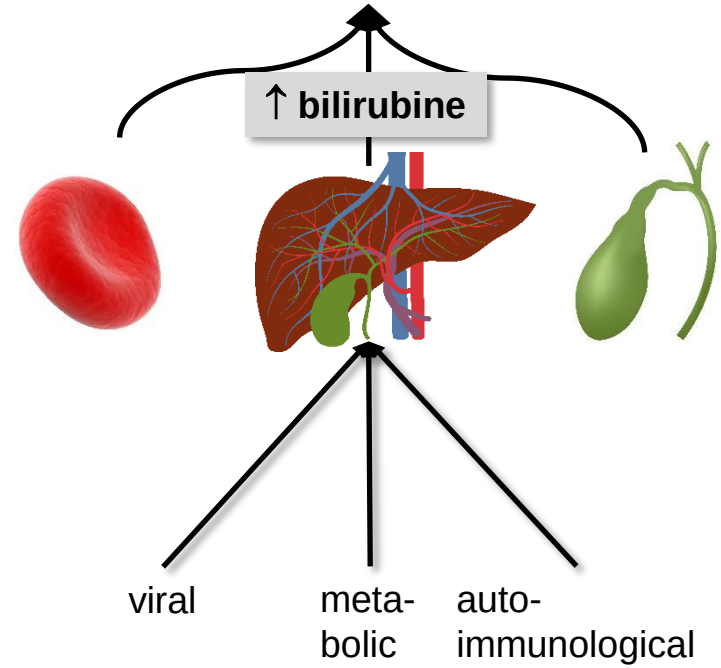
Nosology: From symptoms to mechanisms & causes

SYMPTOM

MECHANISMS
(pathophysiology)

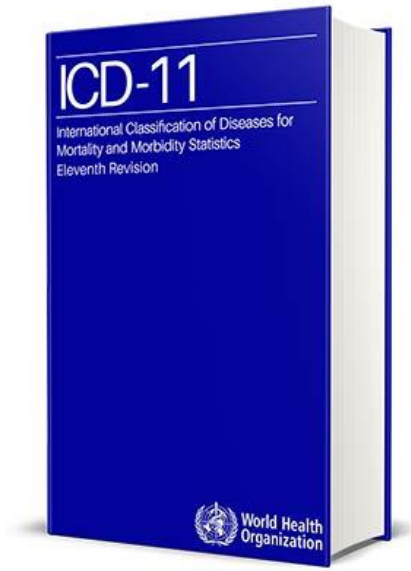
CAUSES
(aetiology)

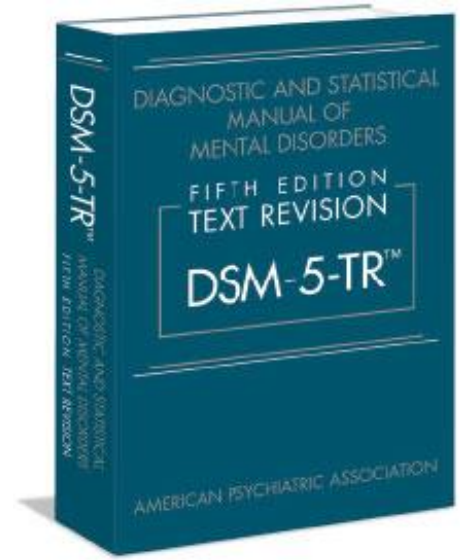




Stephan, *in preparation*

Contemporary psychiatric classifications: ICD and DSM





Both classifications are descriptive: they define mental disorders as syndromes, without reference to mechanisms or causes.

Syndrome

- **syndrome** = a set of symptoms and signs that tend to co-occur
  - a pattern detected by clinical observation
- **symptom** = deviation of bodily or mental states from an expected state
  - *subjective symptom*: accessible for the patient only (e.g. pain, fatigue, low mood)
  - *objective symptom*: can be observed by an external observer (e.g. weight loss, gait disturbance, rash)
- **sign** = an indication of a medically relevant state that is measured/detected as part of a clinical test (not only by observation)
  - examples: impaired reflexes, altered blood count or metabolite levels

General psychometric desiderata for diagnostic classifications

- **Reliability:**
  - A reliable measure is one that measures a construct consistently across time, individuals, and situations.
  - Different types: inter-rater reliability, test-retest reliability ...
  - Reliability is necessary, but not sufficient, for validity.

- **Validity:**
  - The degree to which a measure or concept makes meaningful statements about the world.
  - Main types: face validity, construct validity, predictive validity.

Does DSM enable reliable and valid diagnoses?
To understand this, let's examine its history.

The evolution of DSM since 1952

Edition	Publication date	Number of pages	Number of diagnoses	Revenue for the American Psychiatric Association
DSM-I	1952	132	128	Unknown
DSM-II	1968	119	193	$1.27 million
DSM-III	1980	494	228	$9.33 million
DSM-III-R	1987	567	253	$16.65 million
DSM-IV	1994	886	383	$120 million
DSM-IV-TR	2000	943	383	Unknown
DSM-5	2013	947	541	Unknown

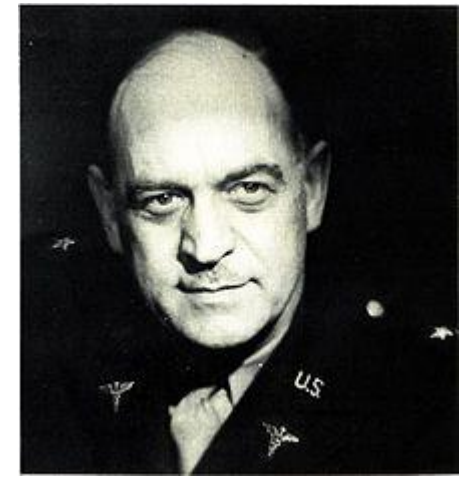
Table: Blashfield et al. 2014, *Ann. Rev. Psychol.*

Figure: https://torontopubliclibrary.typepad.com/north-york-central-blog/2012/03/the-business-of-psychiatry.html

History of psychiatric nosology: Precursors of DSM

- **Precursors:**
  - US 1840 census used a single category: "insanity"
  - 1917: APA developed a guide for mental hospitals, the "Statistical Manual for the Use of Institutions for the Insane" with 22 biologically oriented diagnoses

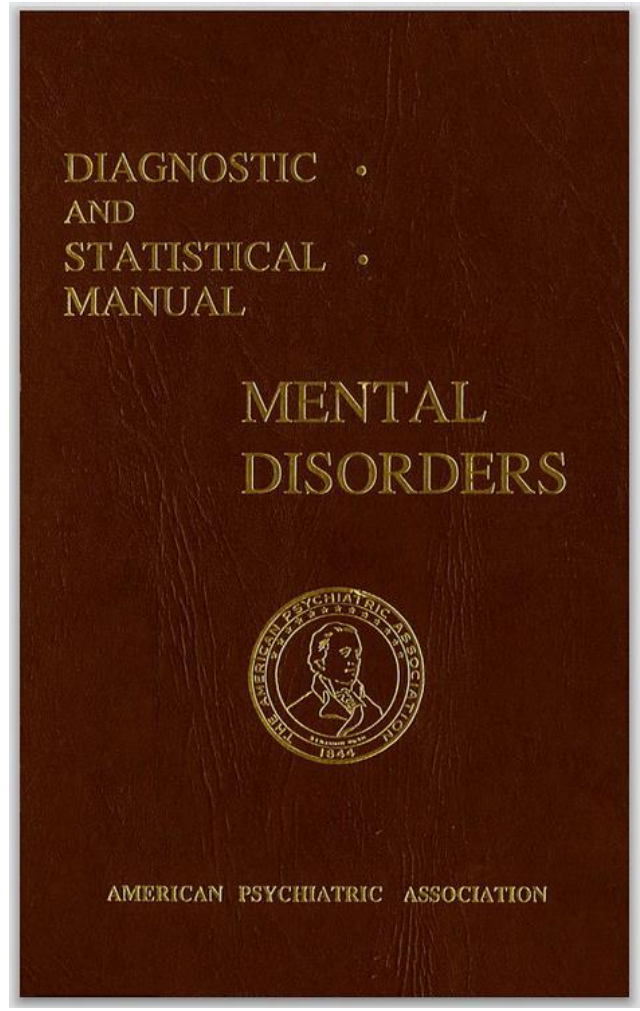
- **Medical 203:**
  - WWII: need for managing a flood of psychiatric disorders among returning US soldiers
  - committee (headed by psychiatrist General William C. Menninger) developed a classification scheme: Medical 203, issued in 1943 as a War Department Technical Bulletin
  - strongly influenced by psychoanalytic concepts



General William C. Menninger (1899-1966)

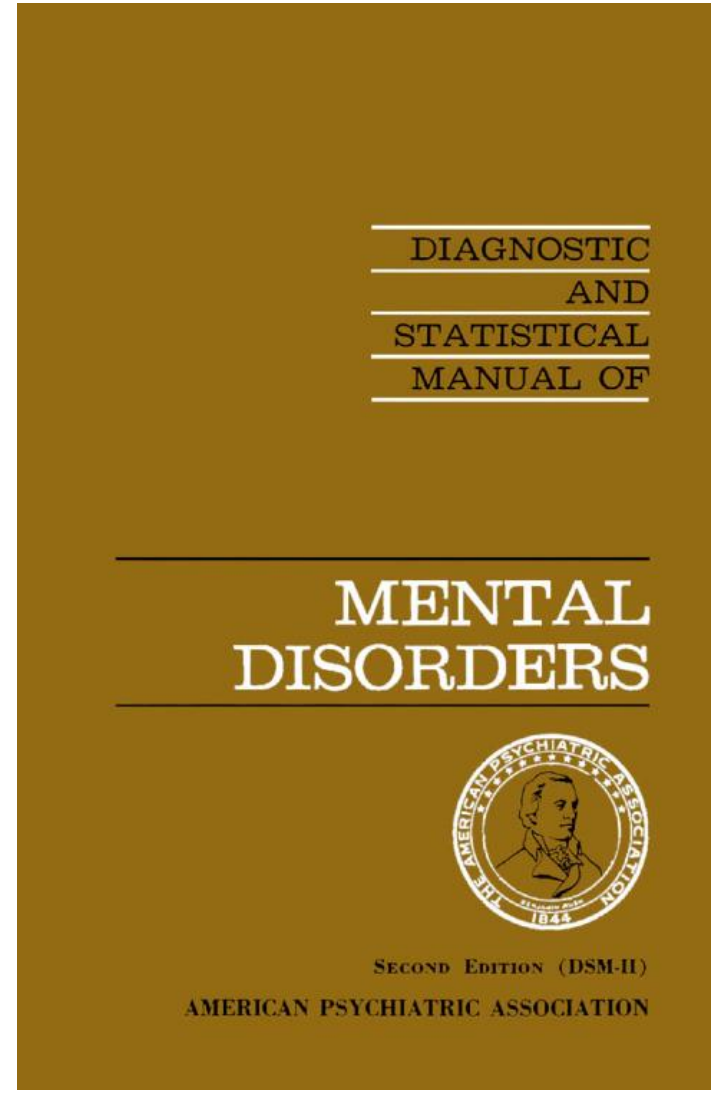
DSM-I (1952)

- 1950: APA committee circulated an adaptation of Medical 203 to approx. 10% of APA members (46% replied, of which 93% approved)
- 1951: approval of revision by APA membership
- published in 1952, with similar structure and conceptual framework as Medical 203 (with many identical passages of text)
- 132 pages long, 128 mental disorders
- strongly influenced by psychoanalysis, e.g. the psychodynamic concept of "reaction" had a major role throughout the manual



DSM-II (1968)

- continued reliance on psychodynamic theory, but first tendencies towards an atheoretical classification
  - e.g. the concept of "reaction" was dropped because some DSM-II contributors wished to avoid relying on particular theories when labeling diseases (Spitzer & Wilson 1968)
- homosexuality
  - had been considered a “sociopathic personality disturbance” in DSM-I
  - became part of “certain nonpsychotic mental disorders” in DSM-II
  - the diagnosis was eventually removed from DSM-II in 1973 (vote of APA membership)



Major criticisms of DSM-II

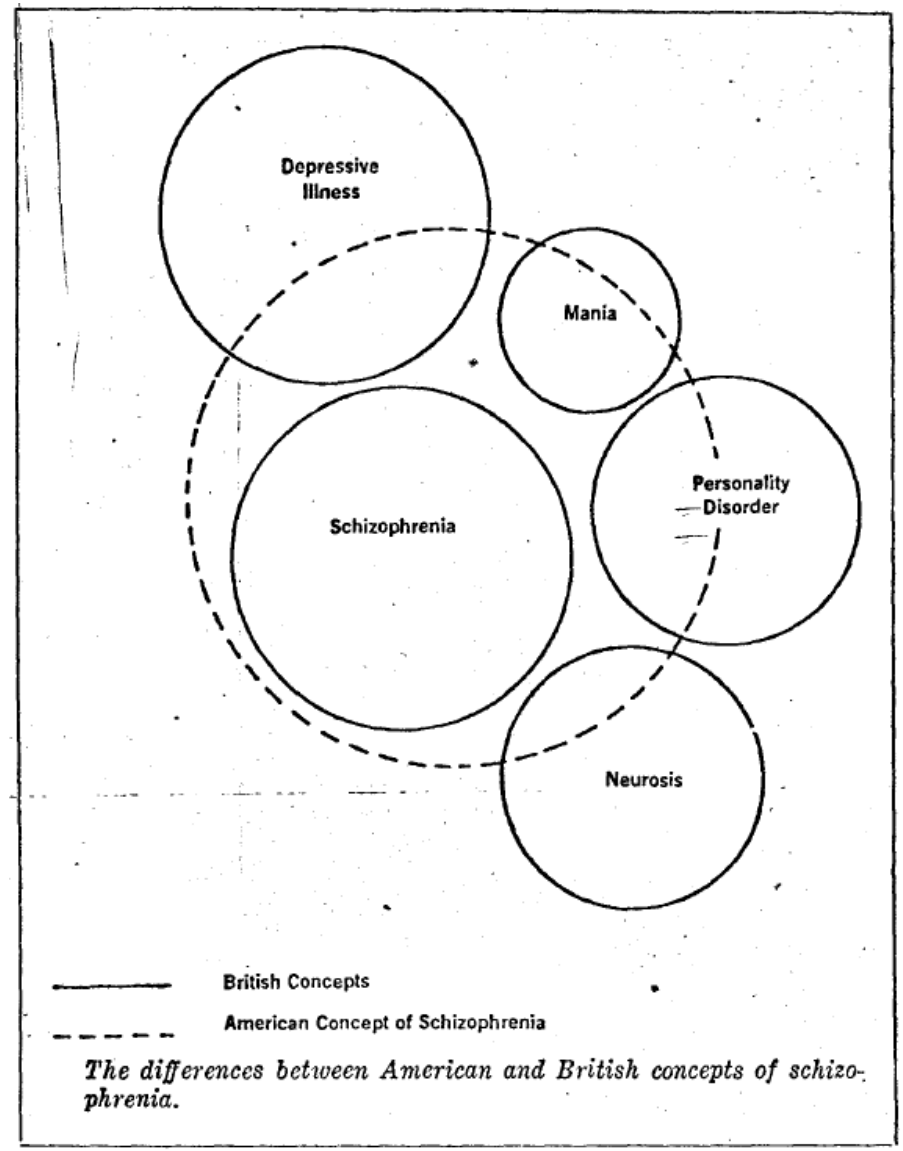
- Aetiological assumptions based on psychodynamic theory
- The concept of homosexuality as a disease
- Reliability
- Validity

Let's briefly revisit two of the key papers from that time.

Evelyn Hooker (1957) The adjustment of the male overt homosexual. *Journal of Projective Techniques* 21: 18-31.
Cooper et al. 1969, *Am. J. Psychiatry* 125: 21-29
Gurland et al. 1970, *Compr. Psychiatry*
Spitzer RL, Fleiss JL (1974) A re-analysis of the reliability of psychiatric diagnosis. *Brit. J. Psychiatry* 125: 341-347.

Kendell et al. 1971: Problems of reliability

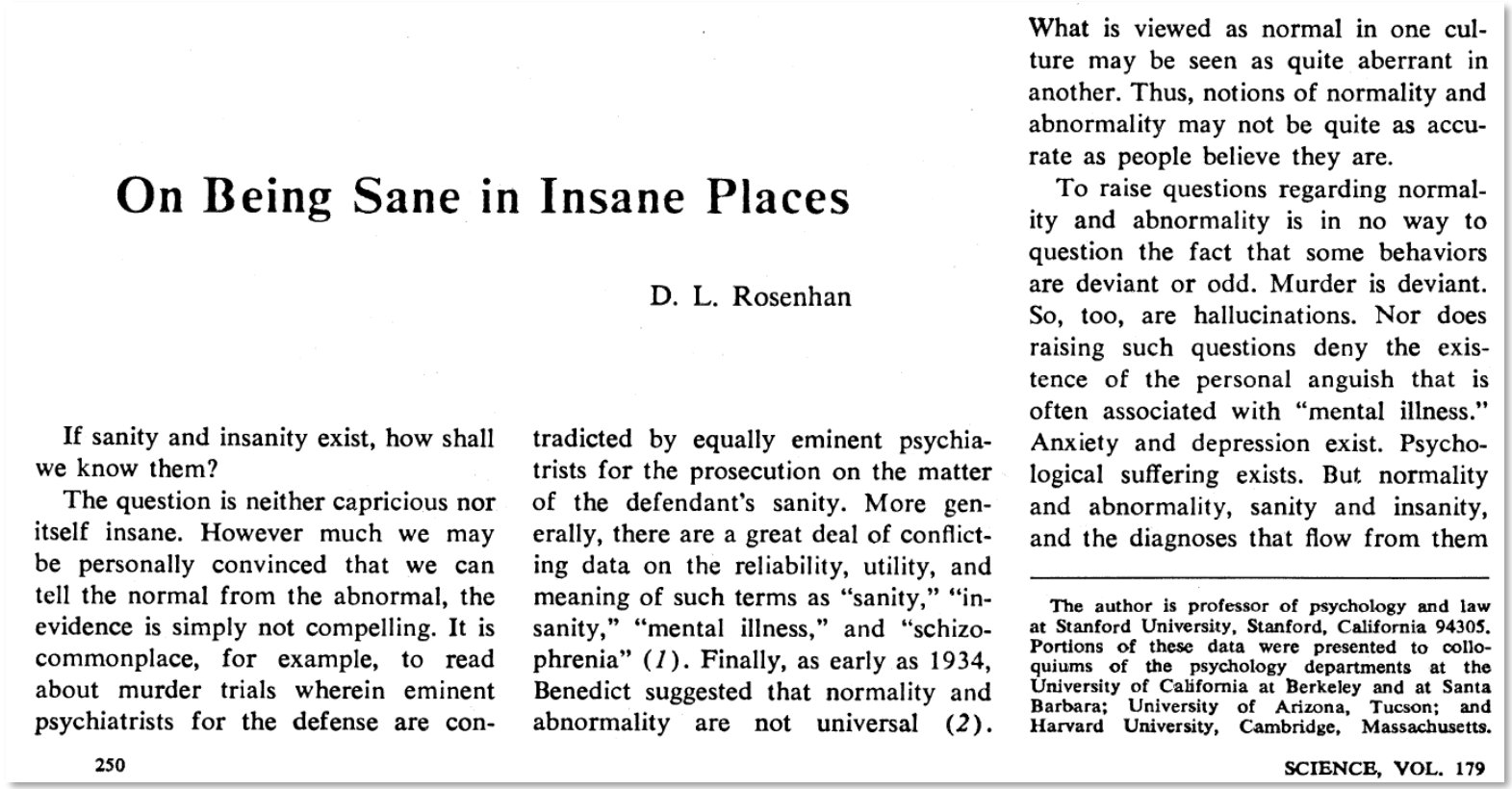
- 8 videotaped diagnostic interviews shown to >200 British and >450 American psychiatrists
- each psychiatrist gave diagnosis according to ICD-8 (similar in structure to DSM-II)
- major differences in diagnoses across nations (e.g. up to 7% vs. 80% diagnoses of schizophrenia)
- conclusion:
  - disease constructs (e.g. schizophrenia) are understood very differently
  - the diagnoses are not reliable



The differences between American and British concepts of schizophrenia.

Kendell et al. 1971, *Archives of General Psychiatry*

The Rosenhan experiment (1973): (Alleged) problems of validity



On Being Sane in Insane Places

D. L. Rosenhan

If sanity and insanity exist, how shall we know them?

The question is neither capricious nor itself insane. However much we may be personally convinced that we can tell the normal from the abnormal, the evidence is simply not compelling. It is commonplace, for example, to read about murder trials wherein eminent psychiatrists for the defense are contradicted by equally eminent psychiatrists for the prosecution on the matter of the defendant's sanity. More generally, there are a great deal of conflicting data on the reliability, utility, and meaning of such terms as "sanity," "insanity," "mental illness," and "schizophrenia" (*1*). Finally, as early as 1934, Benedict suggested that normality and abnormality are not universal (*2*). What is viewed as normal in one culture may be seen as quite aberrant in another. Thus, notions of normality and abnormality may not be quite as accurate as people believe they are.

To raise questions regarding normality and abnormality is in no way to question the fact that some behaviors are deviant or odd. Murder is deviant. So, too, are hallucinations. Nor does raising such questions deny the existence of the personal anguish that is often associated with "mental illness." Anxiety and depression exist. Psychological suffering exists. But normality and abnormality, sanity and insanity, and the diagnoses that flow from them

The author is professor of psychology and law at Stanford University, Stanford, California 94305. Portions of these data were presented to colloquiums of the psychology departments at the University of California at Berkeley and at Santa Barbara; University of Arizona, Tucson; and Harvard University, Cambridge, Massachusetts.

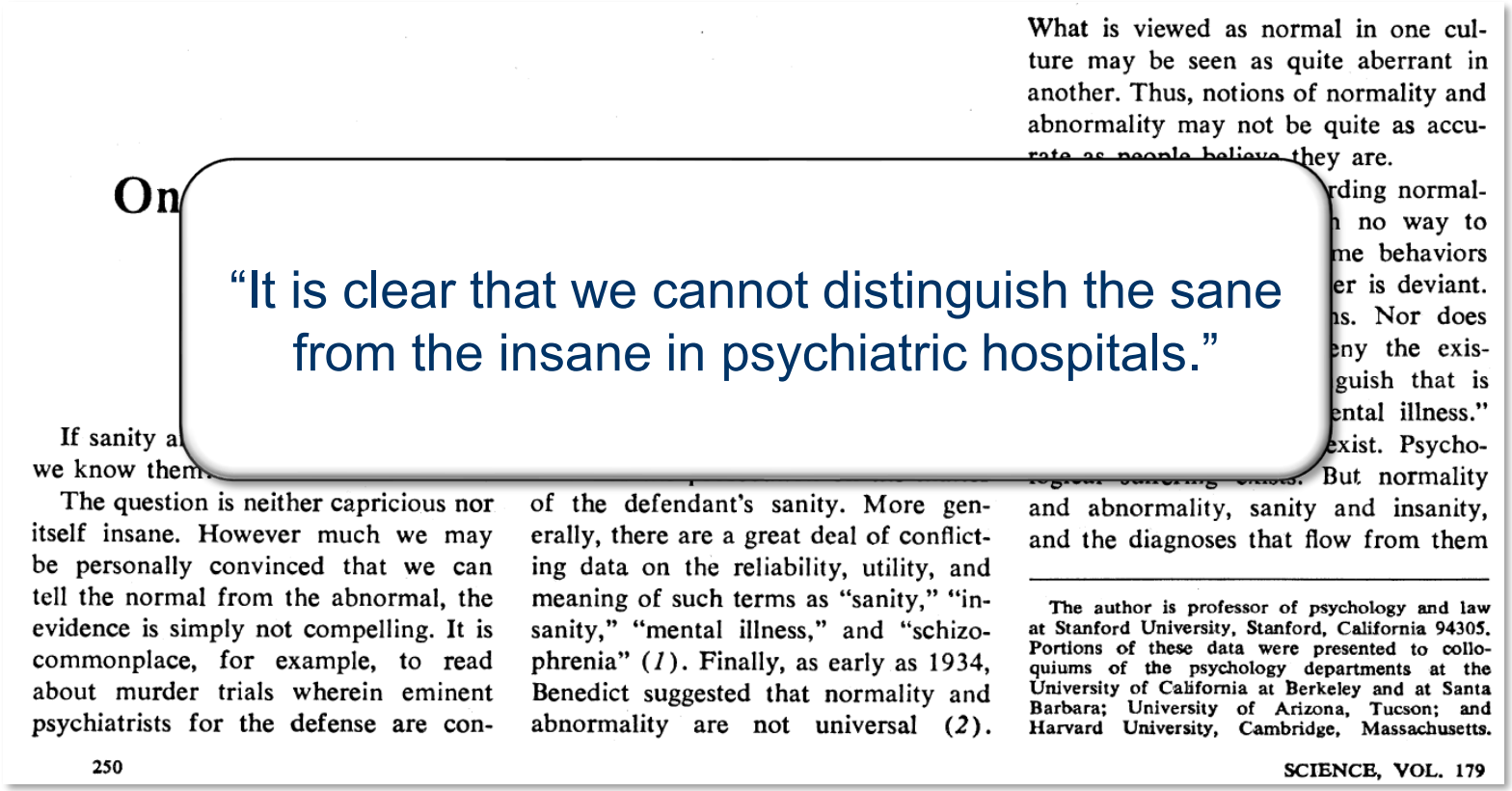
Rosenhan DL (1973) *Science 179*: 250-258

The Rosenhan experiment (1973)

- pseudopatients: "Eight sane people gained secret admission to 12 different hospitals..."; this included Rosenhan
- pseudopatients reported auditory hallucinations on admission
- after admission, the pseudopatients reported they no longer experienced hallucinations, behaved normally, friendly and cooperatively (but openly took extensive notes of his/her experiences)
- pseudopatients were kept for 7-52 days (average 19 days)
- none was detected as fraudulent, all were discharged with the diagnosis of schizophrenia

Rosenhan DL (1973) *Science 179*: 250-258

The Rosenhan experiment (1973)



What is viewed as normal in one culture may be seen as quite aberrant in another. Thus, notions of normality and abnormality may not be quite as accurate as people believe they are.

"It is clear that we cannot distinguish the sane from the insane in psychiatric hospitals."

The question is neither capricious nor itself insane. However much we may be personally convinced that we can tell the normal from the abnormal, the evidence is simply not compelling. It is commonplace, for example, to read about murder trials wherein eminent psychiatrists for the defense are con-

of the defendant's sanity. More generally, there are a great deal of conflicting data on the reliability, utility, and meaning of such terms as "sanity," "insanity," "mental illness," and "schizophrenia" (*1*). Finally, as early as 1934, Benedict suggested that normality and abnormality are not universal (*2*).

and abnormality, sanity and insanity, and the diagnoses that flow from them

The author is professor of psychology and law at Stanford University, Stanford, California 94305. Portions of these data were presented to colloquiums of the psychology departments at the University of California at Berkeley and at Santa Barbara; University of Arizona, Tucson; and Harvard University, Cambridge, Massachusetts.

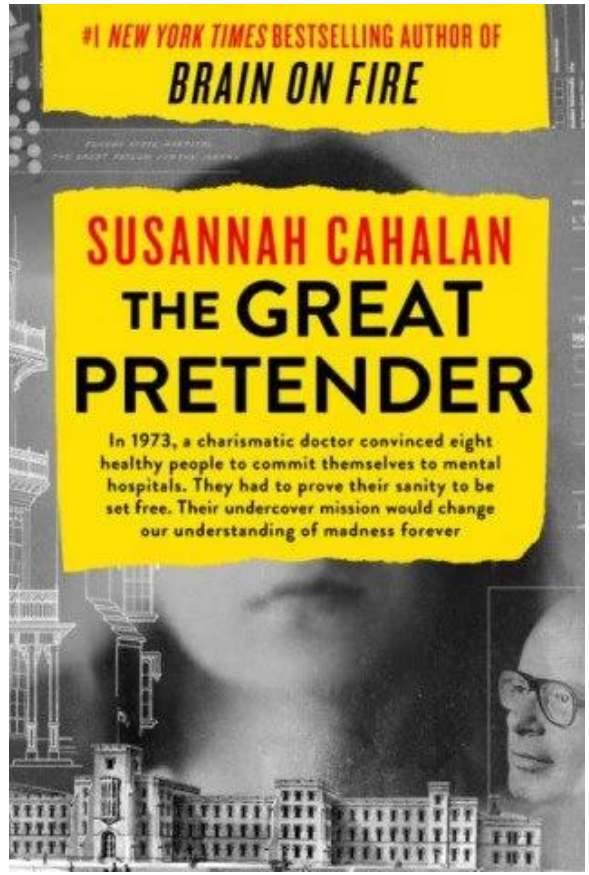
250

SCIENCE, VOL. 179

Rosenhan DL (1973) *Science 179*: 250-258

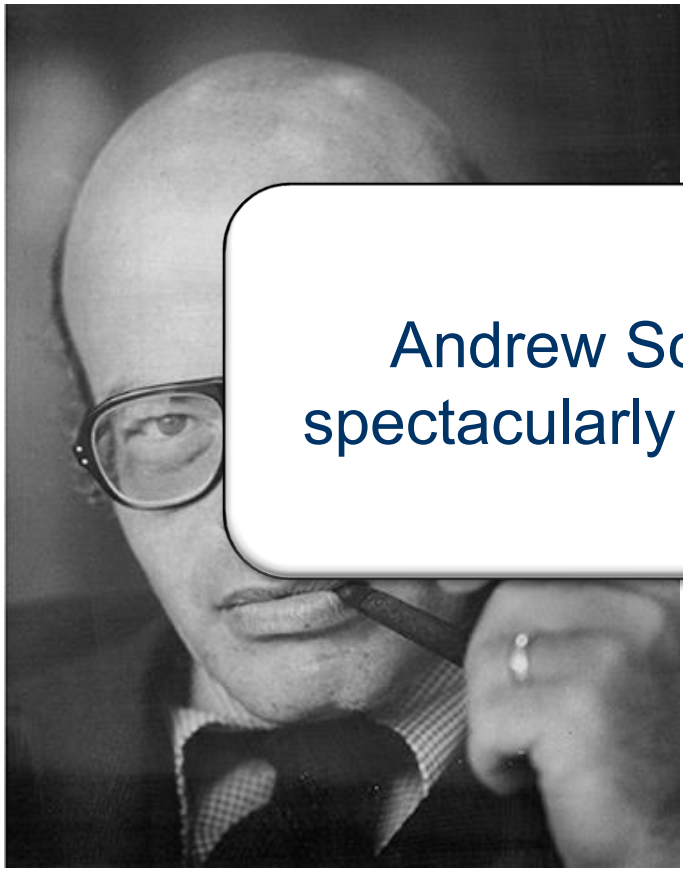
Susannah Cahalan: The Great Pretender (2019)

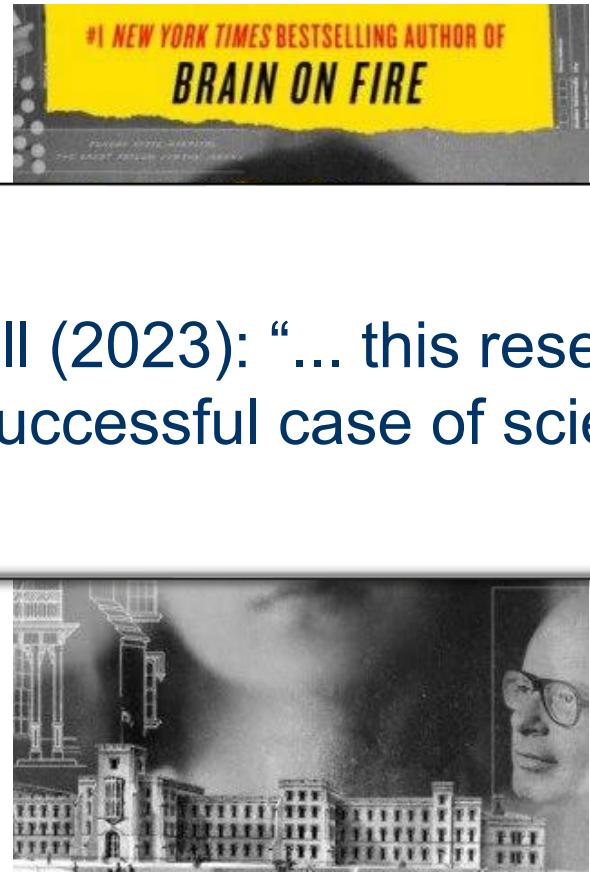
David Rosenhan

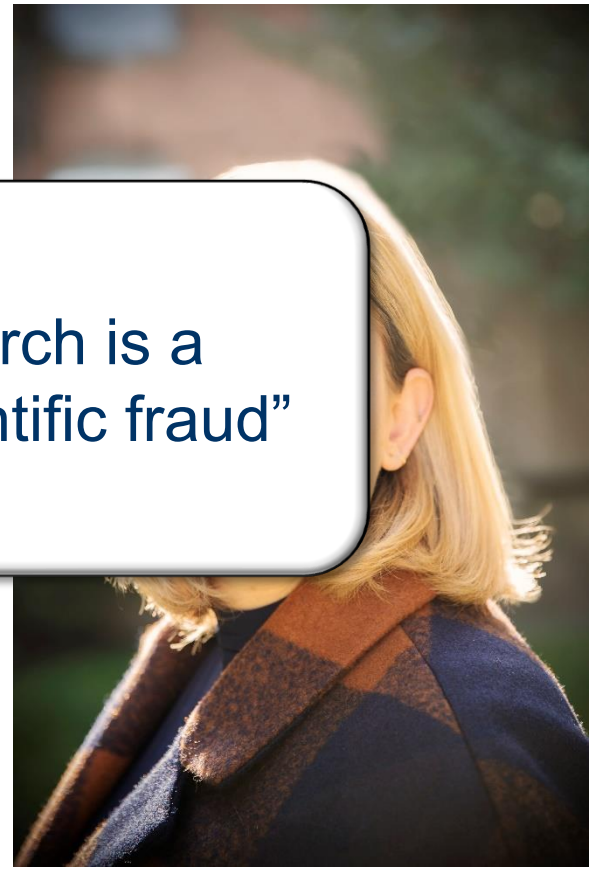


Susannah Cahalan

Susannah Cahalan: The Great Pretender (2019)





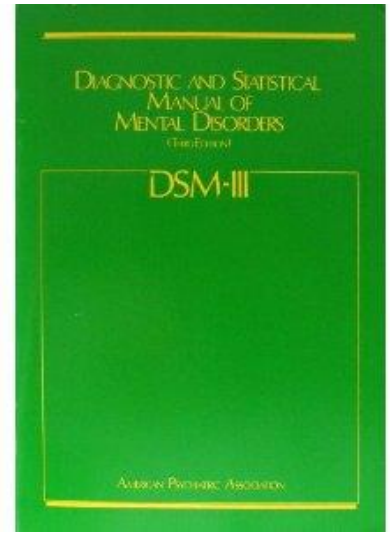


Andrew Scull (2023): “... this research is a spectacularly successful case of scientific fraud”

David Rosenhan

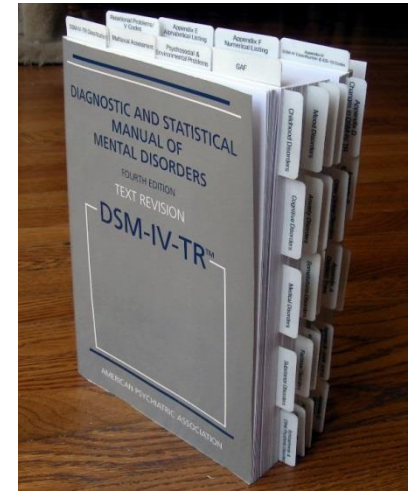
Susannah Cahalan

DSM-III (1980)



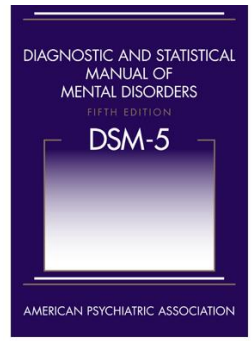
- criticisms of reliability and validity of DSM-II: credibility of psychiatry at stake!
  - revolution needed, not just a reform
- Key goals for DSM-III:
  - improve **reliability** of psychiatric diagnoses
  - **remove assumptions of etiology**
- Key features:
  - **psychodynamics abandoned:** concept of "neurosis" largely disappeared (except as addendum, due to a political compromise)
  - **categorical approach:** each particular pattern of symptoms in a category is assumed to reflect a particular underlying pathology
  - "multiaxial" system (clinical disorders, personality disorders)

DSM-IV (1994) / DSM-IV-TR (2000)



- continues the DSM-III approach but focus on transparency
- Major changes:
  - initial literature reviews, analysis of unpublished datasets
  - transparency how disease definitions were agreed upon: publication of entire evidence base (three source books)
  - 5-dimensional multi-axial system:
    - Axis I: All psychological diagnostic categories except mental retardation and personality disorder
    - Axis II: Personality disorders and mental retardation
    - Axis III: General medical conditions; acute medical conditions and physical disorders
    - Axis IV: Psychosocial and environmental factors contributing to the disorder
    - Axis V: Global Assessment of Functioning (GAF)

DSM-5 (2013)



- preparations started in 1999
- hope that progress in genetics and neuroimaging would enable a pathophysiological foundation
- this hope was not fulfilled:
  DSM-5 continues to be a purely phenomenological scheme
- general structure similar to DSM-IV
  (but multiaxial system of diagnosis removed)
- some notable changes compared to DSM-IV:
  - omitting Asperger syndrome as a distinct category
  - removing subtype classifications (paranoid, disorganized, catatonic, undifferentiated, residual) for schizophrenia
  - omitting the "bereavement exclusion" for depressive disorders
  - gambling disorder as a new entity (classified as addictive disorder, not impulse control disorder)

Reliability of DSM-5 diagnoses: official field trials

- Regier et al. 2013, AJP: confusing terminology - the paper's title (and text) refers to "test-retest") but reports estimats of inter-rater reliability estimates
- kappa: estimate of inter-rater reliability for categorical items that accounts for the possibility of agreement occurring by chance
- results for 15 adult + 8 child/adolescent diagnoses:
  - 5 diagnoses in the very good range (kappa: 0.60–0.79), e.g. PTSD, somatic symptom disorder
  - 9 diagnoses in the good range (kappa: 0.40–0.59), e.g. schizophrenia
  - 6 diagnoses in the questionable range (kappa 0.20–0.39), e.g. MDD
  - 3 diagnoses in the unacceptable range (kappa < 0.20), e.g. mixed anxiety-depressive disorder

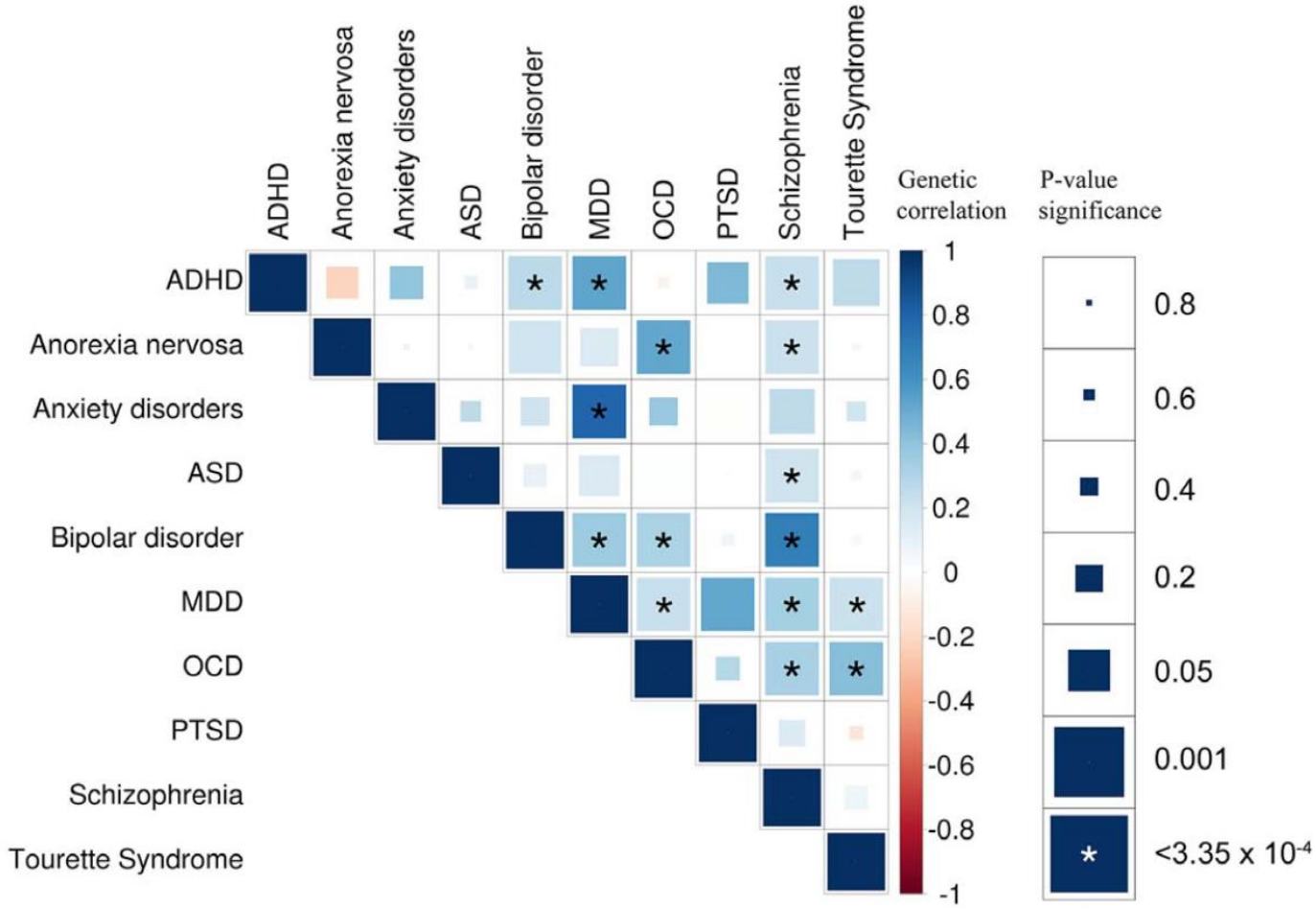
Regier et al. (2013) *Am. J. Psychiatry* 170:71-82

Validity of DSM

- validity of a diagnostic classification can be assessed in different ways
  - **face validity:** do diagnoses refer to syndromes patients and psychiatrists recognise as meaningful?
  - **construct validity:** do diagnoses refer to distinct categories (e.g., in genetic terms)?
  - **predictive validity:** do diagnoses predict clinically relevant future events?
    - treatment response,
    - development of symptoms over time (clinical trajectory)
    - probability of remission
    - risk of relapse (e.g., after treatment discontinuation)
- clinical utility is central: a diagnostic classification without predictive validity does not support clinical decision-making!

Construct validity?
Shared heritability: psychiatric disorders

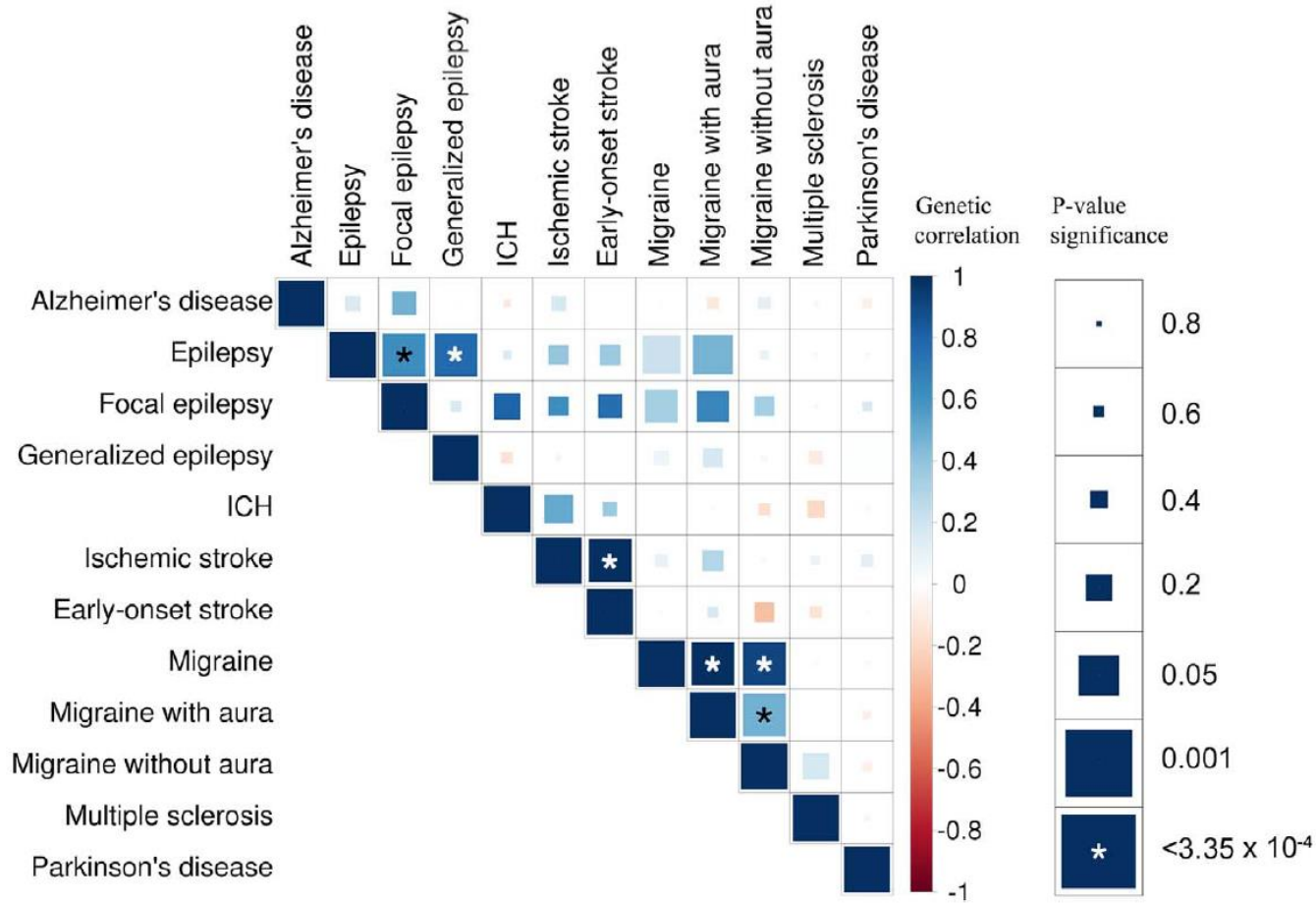
- meta-analysis
- 265,218 cases of different brain disorders and 784,643 controls
- almost all psychiatric disorders - except PTSD - show significant genetic correlations with other disorders



The Brainstorm Consortium, Science 360, 1313 (2018)

Construct validity?
Shared heritability: neurological disorders

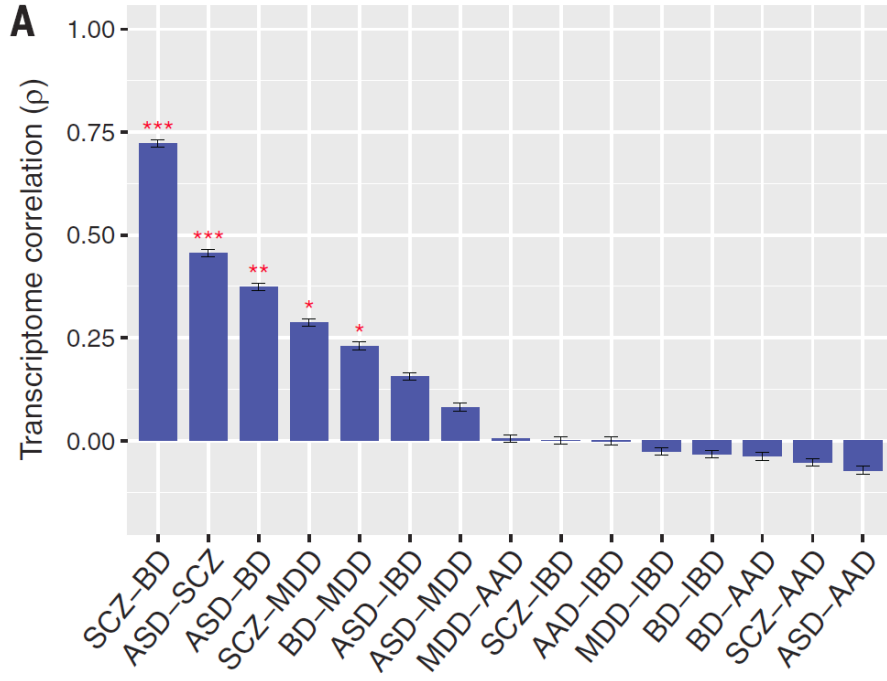
- in contrast to psychiatric disorders, neurological disorders show very limited genetic correlations with other disorders

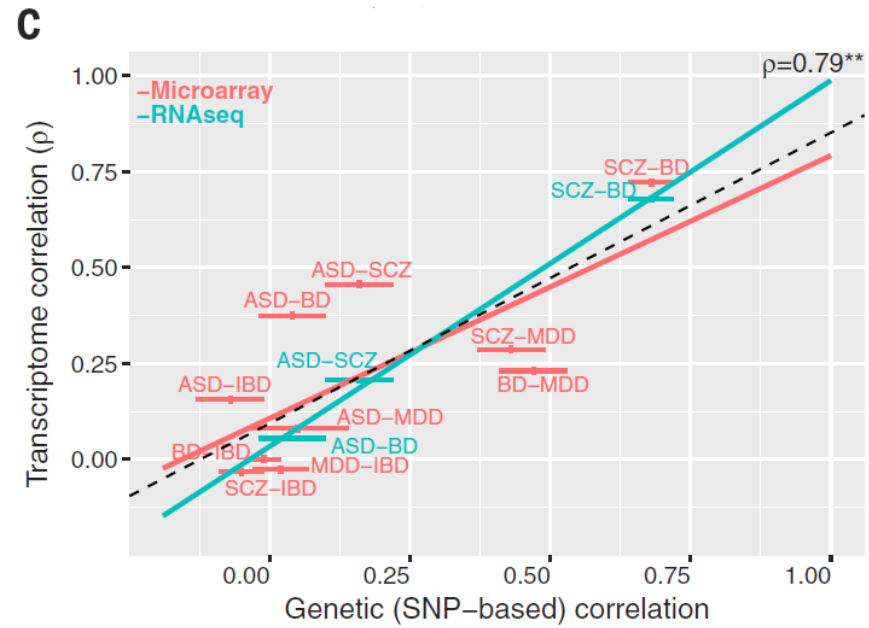


The Brainstorm Consortium, Science 360, 1313 (2018)

Construct validity? Shared transcriptional dysregulation

- post mortem microarray analyses of cortical gene expression, i.e. RNA (407 patients, 293 controls)
- disorders: autism, schizophrenia, bipolar disorder, depression, alcohol use disorder
- A: significant overlap in transcriptional dysregulation (differential gene expression compared to healthy controls)...
- C: ... that was correlated with genetic (SNP) overlap

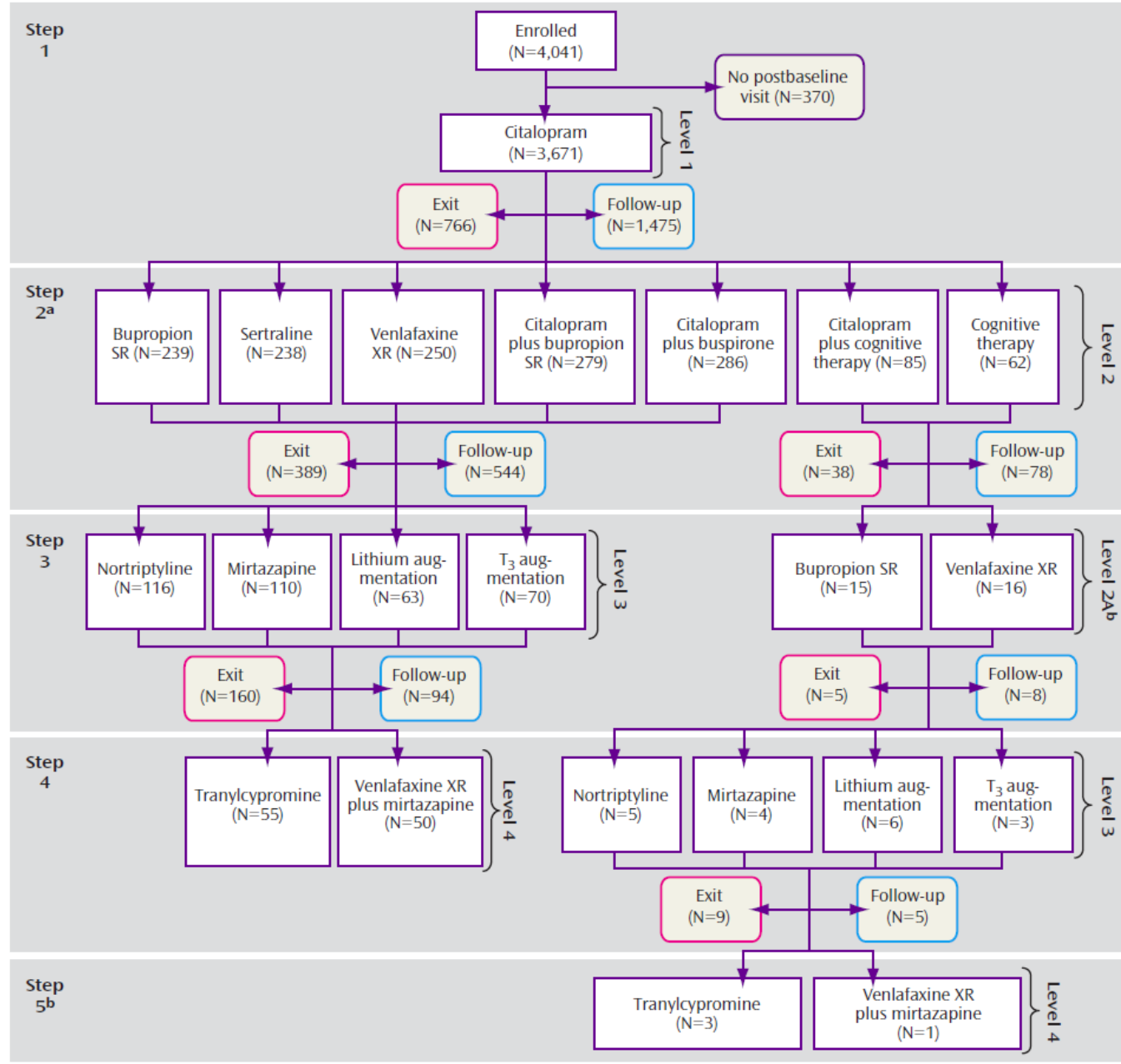




Gandal et al. 2018, *Science* 359: 693–697

Predictive validity of DSM: STAR*D study

- illustrates trial-and-error treatment process
- adult outpatients with nonpsychotic MDD receiving one (N=3,671) to four (N=123) successive acute treatments (randomized assignment)
- remission rates were 36.8%, 30.6%, 3.7%, 13.0% for steps 1-4
- cumulative remission rate after four treatments only 67%



Rush et al. (2006) *Am. J. Psychiatry* 163:1905–1917.

Predictive validity of DSM: 1-year trajectories of cognitive & somatic symptoms of MDD

- N=205 MDD patients from primary care
- retrospective assessment of weekly MDD symptoms each month, over 1 year
- latent class growth analysis (a special case of SEM)
- 4 trajectory classes for both cognitive and somatic symptoms (no diffs. in treatment!):
  - quick recovery (37%)
  - chronic cognitive (25%)
  - chronic somatic (19%)
  - chronic cognitive + somatic (19%)

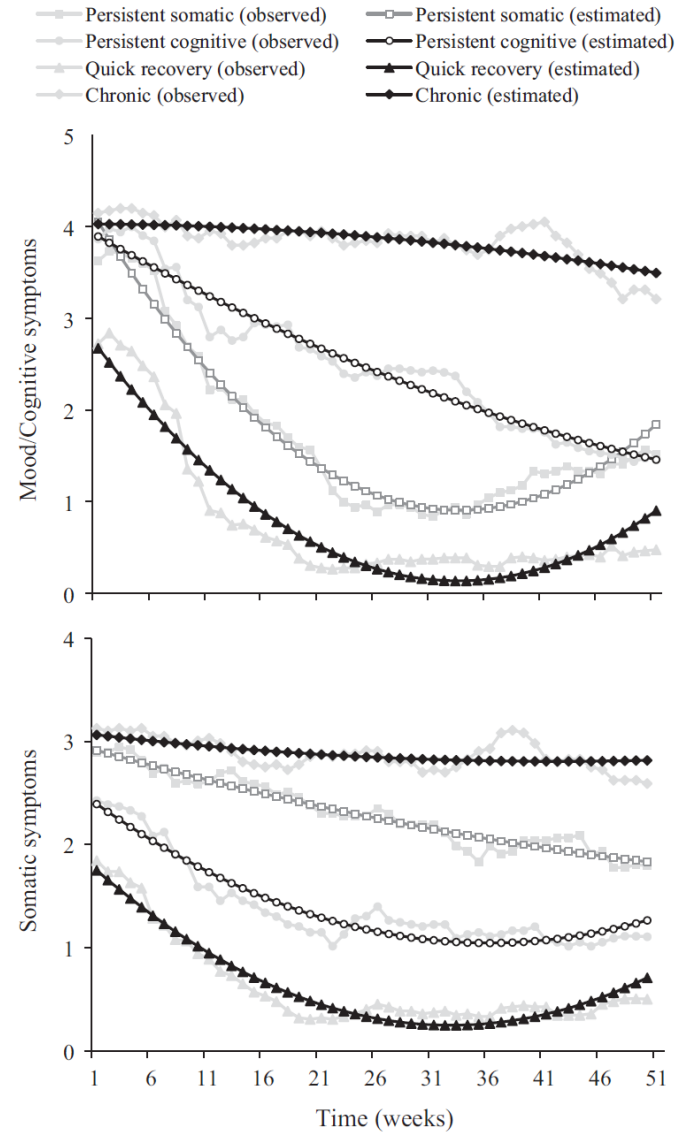


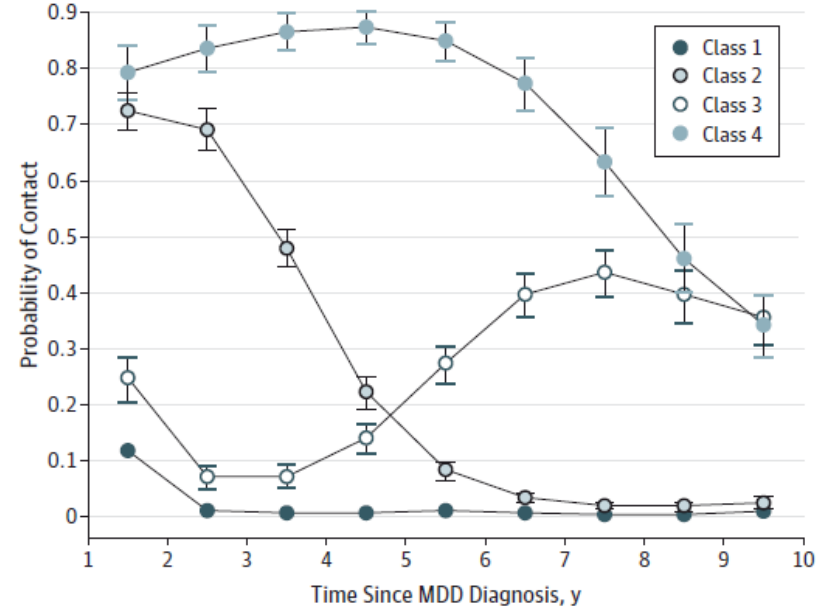
Fig. 1. Observed (gray) and estimated (black) growth trajectories for each of the PP-LCGA classes on each of the two investigated symptom domains: depressive mood/cognition and somatic symptoms.

Waardenar et al. 2015, *J. Affect. Dis.*

Predictive validity of DSM: 10-year trajectories of MDD

- Danish Psychiatric Central Research Register (DPCRR)
- 11,640 individuals (64% women,18-48 years) at their first recorded MDD diagnosis
- latent class growth analysis
- 4 trajectory classes identified
  - brief contact (77.0%)
  - prolonged initial contact (12.8%)
  - later reentry (7.1%)
  - persistent contact (3.1%)

Figure. Patterns of 10-Year Course Trajectories of Major Depressive Disorder (MDD)

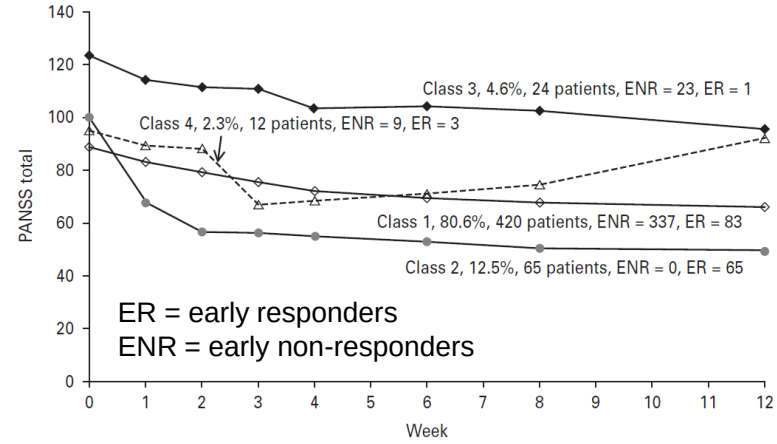


Class 1 indicates brief contact (77.0%); class 2, prolonged initial contact (12.8%); class 3, later reentry (7.1%); and class 4, persistent contact (3.1%). The model included the following covariates: sex, birth year, rural birthplace, age at initial diagnosis, calendar year at initial diagnosis, inpatient treatment at initial diagnosis, record of suicide attempts or self-harm, initial episode severity (mild [reference], moderate, severe without psychotic features, severe with psychotic features, or severity unspecified), and parental history of psychiatric diagnoses in the Danish Psychiatric Central Research Register (depression, bipolar disorder, psychotic illness, substance abuse, and anxiety or somatoform disorders). Data markers indicate predicted probabilities; error bars, 95% CIs.

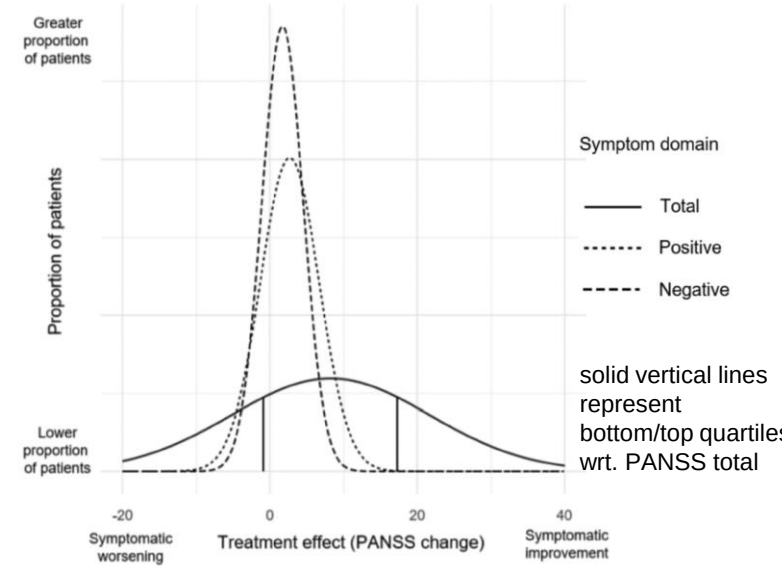
Musliner et al. 2016, *Jama Psychiatry*

Predictive validity of DSM: Clinical responses to antipsychotics

- different subgroups of treatment response to risperidone/olanzapine (growth mixture modeling, Case et al. 2011)
- meta-analysis of heterogeneity: top quartile shows >17.7 points improvement on total PANSS, bottom quartile experiences negative effect of treatment (McCutcheon et al. 2022)
- "... the treatment of patients with schizophrenia is no better than a 'trial-and-error' approach for each patient with each drug." (Buckley & Miller, npj Schizophrenia 2017)



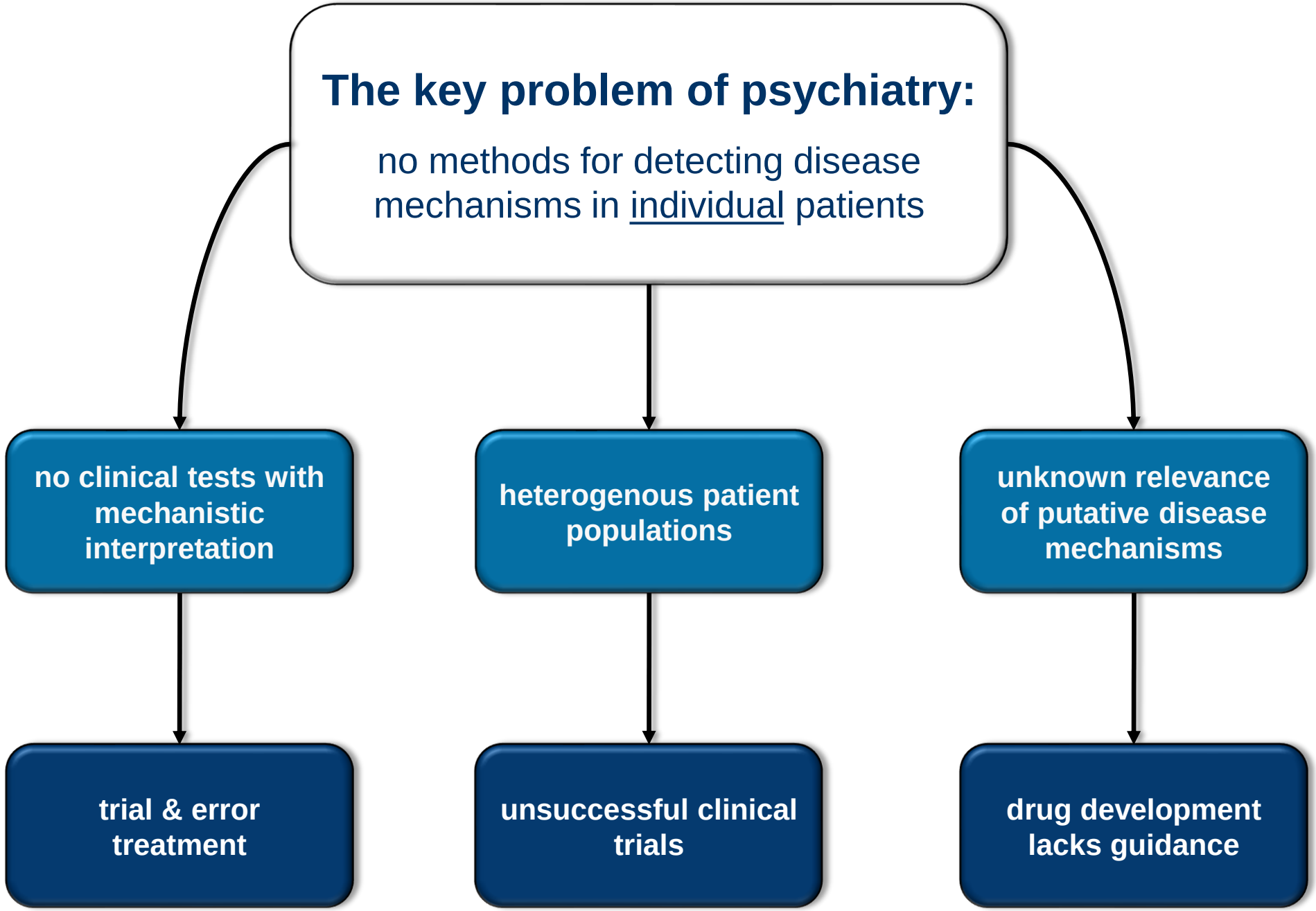
Case et al. 2011, *Psychol. Med.*



McCutcheon et al. 2022, *World Psychiatry*

Summary: DSM

- from aetiological assumptions to a phenomenological scheme focusing on reliability
- disorders
  - defined descriptively as syndromes
  - no reference to mechanisms or causes
  - consensus-defined and increasing in number over time
- reliability: differs substantially across diagnoses
- lack of predictive validity

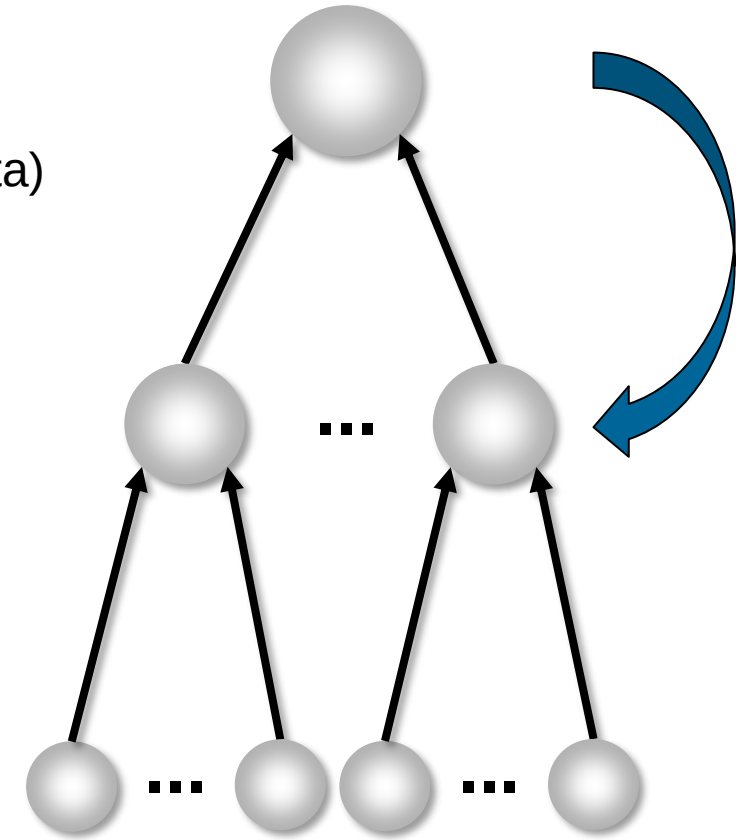

The key problem of psychiatry:
no methods for detecting disease mechanisms in individual patients
no clinical tests with mechanistic interpretation
heterogenous patient populations
unknown relevance of putative disease mechanisms
trial & error treatment
unsuccessful clinical trials
drug development lacks guidance

A key challenge for TN/CP:
Can computational assays help redefine nosology?

SYMPTOMS
(behavioural or physiological data)

MECHANISMS
(cognitive, physiological)

CAUSES
(aetiology)



differential diagnosis of alternative disease mechanisms

spectrum dissection into mechanistically distinct subgroups

prediction of clinical trajectories and treatment response

Thank you