



Data-driven medicine? The case of automated disease classification

EPSA conference
Symposium on Missing theory
“Turin”, 17 September 2021

Jan-Willem Romeijn
Faculty of Philosophy
University of Groningen

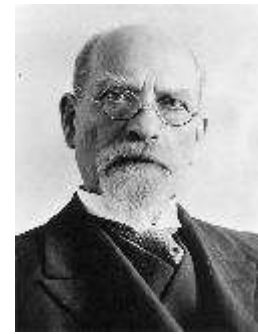
Joint work

Parts of this talk are based on work with Lian Beijers, Hanna van Loo, and several other psychiatrists.



Theoretical social science?

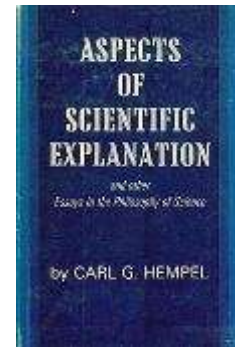
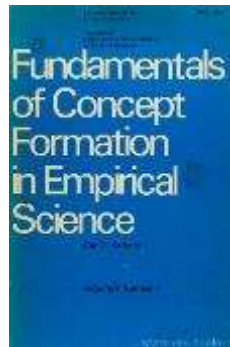
We can spot an interesting difference between the natural sciences and the social and medical sciences.



What explains this difference in the status of theoretical research?

The theoretician's dilemma

For philosophers of science the need for theory has been a live issue for longer.



Theory plays a regulative role within empirical science, and it facilitates a better connection to applications.

Missing theory

Medical and social science research is in need of a re-appraisal of theory.

- › Current social and medical science can be described as “phenomena mongering”. This also hampers meaningful relations to application.
- › The replication crisis has added to the need for methodological rigor and the integration of findings.

The focus in this talk is on how this need for rehabilitating theory combines with the rapid uptake of data science methods in these sciences.

Mathematical psychology?

Eronen and Romeijn (2021) discuss if we can learn from the natural sciences and their use of mathematics in theory development.

Article

Philosophy of science and the formalization of psychological theory

Markus I. Eronen
and Jan-Willem Romeijn
University of Groningen

Abstract
One of the original aims of this journal was to promote theory in psychology. Nowadays more and more psychological researchers are calling for more theory development, and articles on the "theory crisis" have also found their way into mainstream journals. In this article, we provide a further perspective to this theory debate. Over the past century, philosophy of science has staged extensive discussions on the mathematization of nature and on the role of mathematics in the

Theory
& Psychology

Theory & Psychology
2020, Vol. 35(6) 786–799
© The Author(s) 2020

Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0959354320969876
journals.sagepub.com/home/tap

SAGE

Disease classification

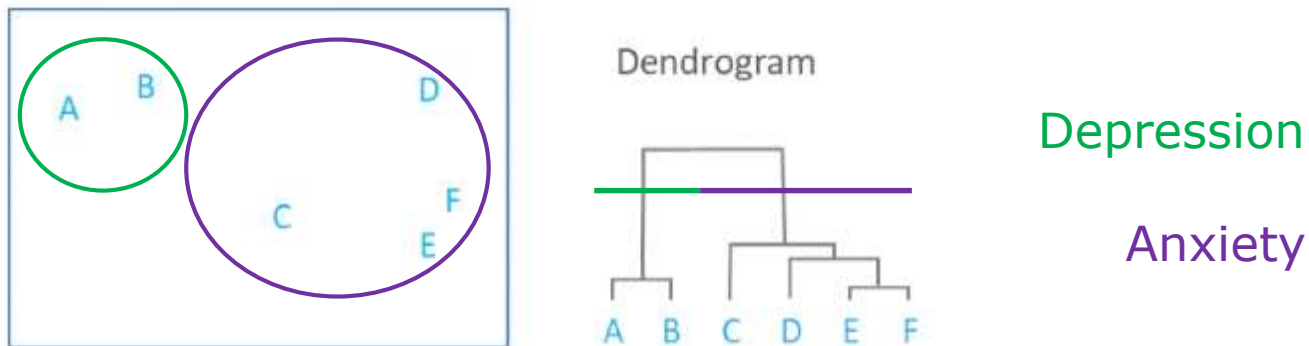
Classification schemes for mental illness serve a large variety of goals.

- › Medical doctors use classification schemes to design and apply treatments.
- › Researchers employ them to design studies and carry them out.
- › Patients and their families and friends fall back on classifications for explanation and understanding.

How can we best serve these goals? When is a classification scheme “good”?

Automated disease classification

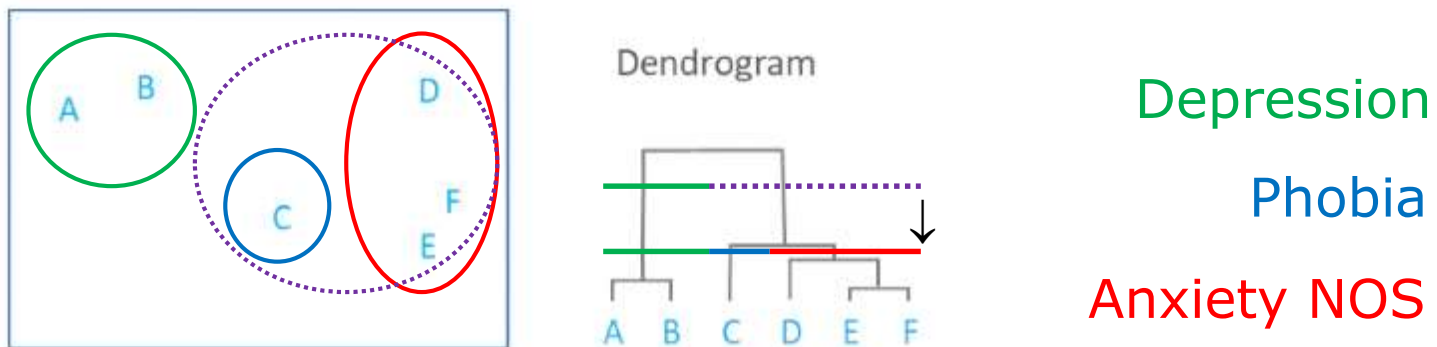
Based on a data set of patients, data science methods can generate patient groups with relevant similarities.



The dataset itself is supposed to determine the grouping.

Sensitivity to parameter settings

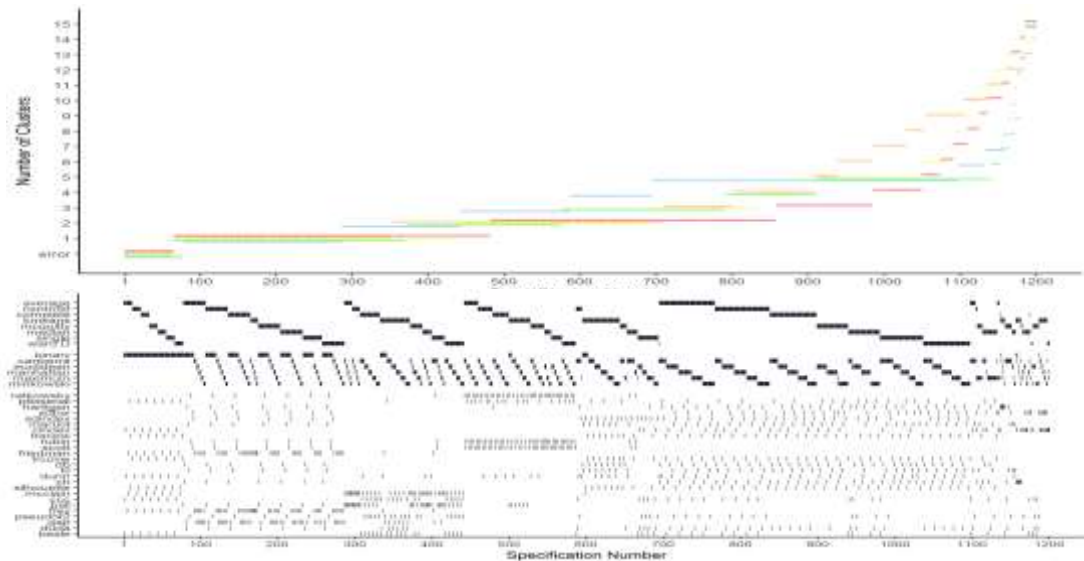
Tweaking the parameter settings in the automated clustering method will change the resulting grouping.



These parameter settings arguably import theoretical assumptions into the data science.

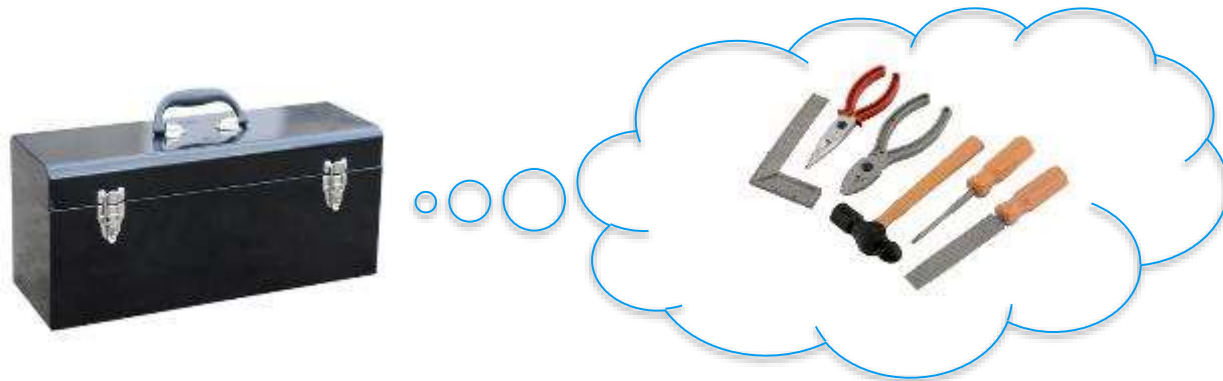
Simulated data

The behaviour of the clustering method does not improve for simulated data with a built-in grouping.



Black-box data science

For many psychologists and psychiatrists data science methods are black boxes.



They do not know how the methods work, and what sort of knowledge the methods can deliver.

Rapid uptake of data science

At the same time the social and medical sciences see a fast increase in the use of data science methods.

ARTICLE Open Access

Increased inflammation and brain glutamate define a subtype of depression with decreased regional homogeneity, impaired network integrity, and anhedonia

Bronnim Hareton^{1,2,3,4}, Sangchuan Chen^{1,2}, Zhibao Liu^{1,2}, Thirubirth Rane⁵, Bobak J. Woodward^{1,2}, Xiaodong P. Hu¹, Jennifer C. Edgar¹ and Andrew H. Miller^{1,2}

Abstract

Continued increases in peripheral inflammation and brain glutamate may identify a subtype of depression with distinct neuroimaging signatures. **Two overlapping subtypes of depressed subjects**, with and without combined elevations in plasma C-reactive protein (CRP) and brain ganglia glutamate (high and low CRP-Glu, respectively) **were identified by magnetic resonance imaging** using plasma CRP (measuring peripheral inflammation) and magnetic resonance spectroscopy (MRS)-based measurement of left basal ganglia glutamate. High CRP-Glu group status was associated

Molecular Psychiatry

Letter to the Editor | Published: 03 November 2020

Problems with latent class analysis to detect data-driven subtypes of depression

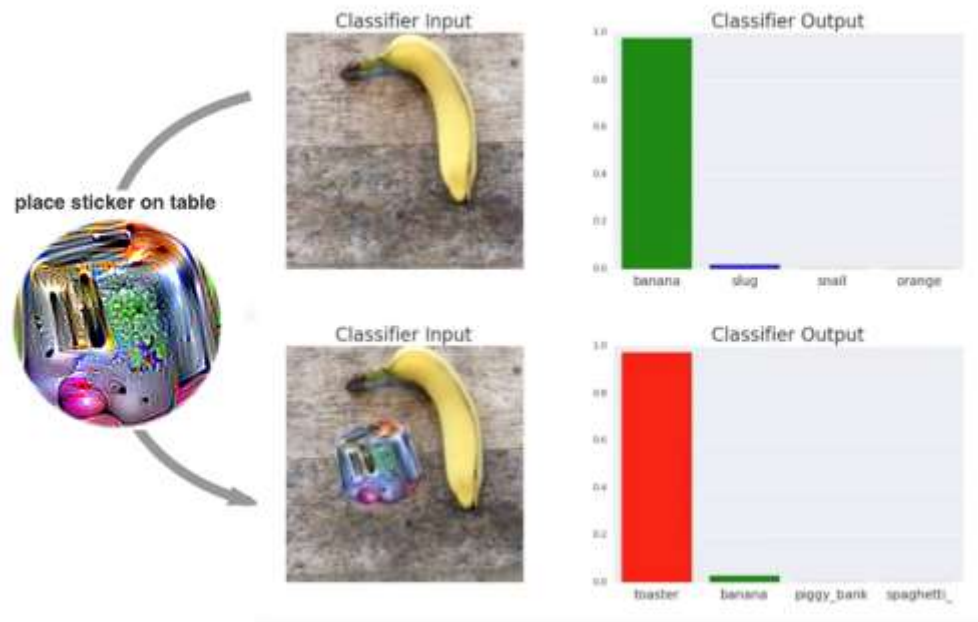
F. M. van Leeuwen^{1,2}, B. B. K. Wanders, E. J. Wardenier & E. H. Hibbe

Molecular Psychiatry 25, 401–406 (2020) | [Download Citation](#)

Depressed patients differ considerably with respect to symptom profiles, course of illness and treatment response. These differences likely contribute to the on average low efficacy of treatment, and drive the search for more homogeneous subtypes of depression in order to facilitate treatment decisions in clinical practice.¹ Latent class analysis (LCA) presents a common statistical method in current depression

Fooling the machine

Problems with the reliability of data science methods generalize. So-called adversarials are a case in point.



Ethical concerns

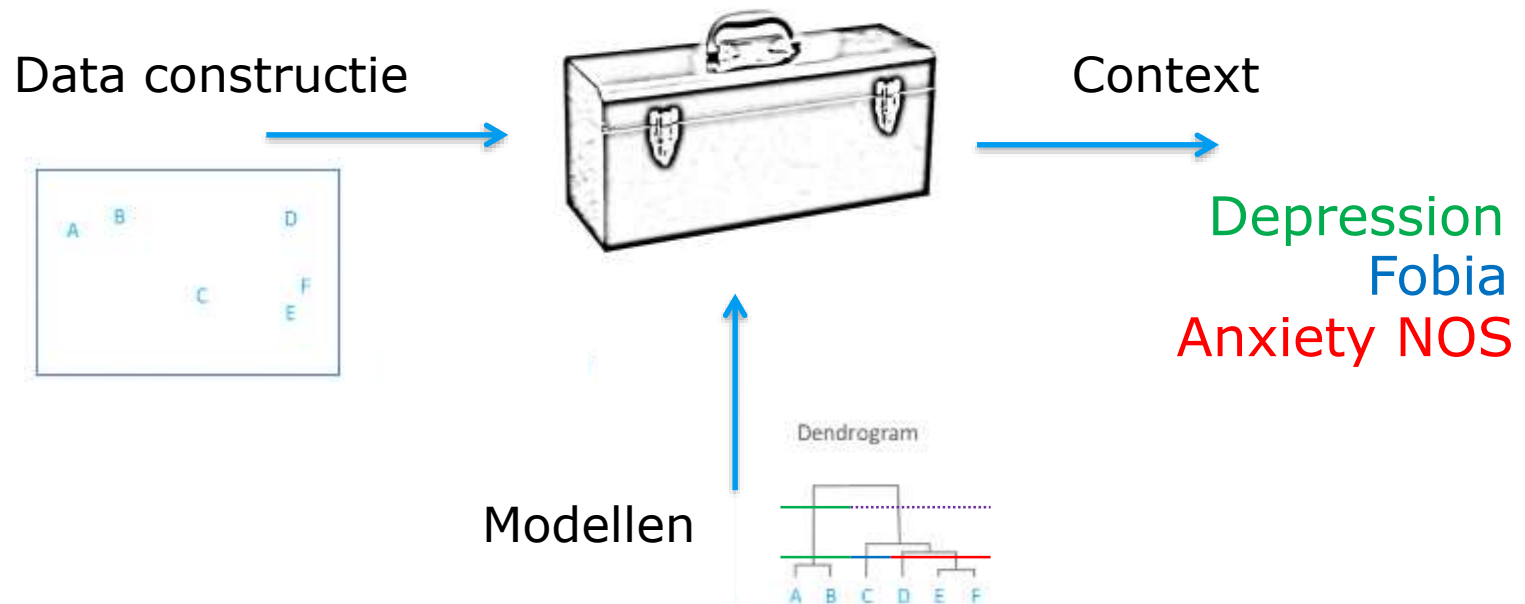
Besides problems of reliability, debates surrounding “machine bias” highlight the importance of transparency.



If we want to intervene in the world on the basis of data science, we have to clarify its workings first.

Clarifying data science

We can make data science methods transparent by a combination of mathematical and practice-oriented philosophy of science.



Thanks for your attention

Some references:

- > “Psychiatric classification: an a-reductionist perspective”, in *Conceptual Issues in Psychiatry*, ed. K. Kendler and J. Parnas, Cambridge University Press, 2020.
- > “Changing the definition of the kilogram: insights for psychiatric disease classification”, with Hanna M. van Loo and K. S. Kendler, in *Philosophy, Psychiatry and Psychology* 26(4), pp. 97–108, 2020
- > “Measuring and defining: the double role of the DSM-criteria for psychiatric disorders”, with H. van Loo, *Psychological Medicine*, 2017.
- > “Psychiatric comorbidity does not only depend on diagnostic thresholds: an illustration with major depressive disorder and generalized anxiety disorder”, with H. van Loo, P. de Jonge, K.S. Kendler, and R.S. Schoevers, *Depression and Anxiety*, DOI 10.1002/da.22453, 2015.
- > “Comorbidity: fact or artefact?”, with H. van Loo, *Theoretical Medicine and Bioethics* 36(1), pp. 41-60, 2015.
- > “Psychiatric comorbidity and causal disease models”, with H.M. van Loo, P. de Jonge, R.A. Schoevers, *Preventive Medicine*, 57(6), pp. 748-752, 2013.