

nearly identical with the sum of diagnoses that go into these respective factors, and that results hold in both general factor and correlated factor models. We see three main implications of our findings.

First, we compare two types of models: a simple sum of indicators vs. a class of highly sophisticated structural equation models that estimate a large number of parameters and impose considerable assumptions on the data, such as hierarchies in which factors are organized, or relations among factors that are constrained to zero³. Both models produce nearly identical scores for participants. If replicated in other studies, this finding suggests that the use of reflective latent variable models should be considered more carefully: what are the specific benefits of this modeling framework for the p factor literature, and do they outweigh the potential costs, such as over-parameterization and stringent assumptions imposed on the data^{3,4}? Such deliberations will benefit from explicit goals to determine whether specific statistical models are adequate in the context of a given research question. In general, scientific progress is often hampered by overreliance on any particular type of model⁸, and thinking more clearly about conceptualizations of p may offer opportunities to diversify methodology.

Second, we provide evidence that p is simply a re-expression of the sum of diagnoses that individuals experience. This is not

surprising: about 70 years ago, Cattell⁹ described scores on the general factor as “essentially the sum of the scores”, and Lahey et al¹ acknowledge the p factor is a “weighted average” of items. Our results imply that p represents severity or comorbidity, not liability, much in the same way as the sum of flu symptoms provides a rough index for severity, not liability. Whether competing accounts of p offer better explanations, such as the idea that it represents liability, requires that models be estimated on variables that actually denote liability, rather than variables denoting severity and comorbidity.

Third, if p is a mere index of the data, this suggests that the meaning of p will only be invariant across studies inasmuch as the data that go into our models are invariant across studies.

Overall, data can be brought to bear on theories when statistical models impose assumptions on the data that are in line with the theories. The p factor literature has been largely atheoretical and primarily concerned with *description* of data – a crucial first step to establish phenomena that can then be explained. But let us not lose sight of the fact that p is an effect that needs to be explained (i.e., *explanandum*), not something that does the explaining (i.e., *explanans*). It necessarily emerges from a positive manifold, and tells us nothing about the mechanisms that generated the data^{4,6}. Further, if the goal is the description of data, it is unclear why

the reflective latent variable model that is solely relied upon in the literature should be the only model suited for this goal.

Thinking more clearly about theories of p^{1,2}, and spelling out these theories precisely, will help adjudicate between different conceptual accounts of p. Criticizing and modifying theories requires that we know exactly where they start and end. Clearer theories will then facilitate choosing appropriate statistical models that can in turn guide theory reform.

Eiko I. Fried¹, Ashley L. Greene²,
Nicholas R. Eaton²

¹Unit of Clinical Psychology, Faculty of Social and Behavioral Sciences, Leiden University, Leiden, The Netherlands;

²Department of Psychology, Stony Brook University, Stony Brook, NY, USA

E.I. Fried and A.L. Greene contributed equally to this work.

1. Lahey BB, Moore TM, Kaczkurkin AN et al. *World Psychiatry* 2021;20:57-63.
2. Caspi A, Moffitt TE. *Am J Psychiatry* 2018;175:831-44.
3. Greene AL, Eaton NR, Li K et al. *J Abnorm Psychol* 2019;128:740-64.
4. van Bork R, Epskamp S, Rhemtulla M et al. *Theory Psychol* 2017;27:759-73.
5. Bender DS. *J Pers Assess* 2019;101:356-9.
6. Fried EI. <https://doi.org/10.31234/osf.io/zg84s>.
7. Greene AL, Eaton NR. *Compr Psychiatry* 2017;72:74-82.
8. Vaidyanathan U, Vrieze SI, Iacono WG. *Psychol Inq* 2015;26:209-30.
9. Cattell RB. *Factor analysis: an introduction and manual for the psychologist and social scientist*. New York: Harper, 1952.

DOI:10.1002/wps.20814

Why hierarchical dimensional approaches to classification will fail to transform diagnosis in psychiatry

At the outset, I would like to stipulate that the current DSM and ICD approaches towards diagnostic classification are not perfect. Others have elaborated on the limitations of these categorical approaches towards diagnosis¹; so I do not repeat them here. I also stipulate that there are some advantages to a dimensional conceptualization of psychopathology over a categorical one. Nonetheless, I am fairly confident that an empirically derived dimensional classification will not replace the DSM-5/ICD-11 anytime soon, if ever.

Eight potential barriers to the integra-

tion into clinical practice of one such model, the Hierarchical Taxonomy of Psychopathology (HiTOP), have been identified². Among them, are the length of clinical evaluations, billing for clinical encounters, and incorporating the model into training. The implicit message is that clinicians will require some convincing. That is, clinicians are likely to resist such a seismic change unless a compelling case is made to support the adoption of a new approach towards assessment and diagnosis. While the supporters of dimensional approaches have identified some obstacles to be over-

come to transform the categorical system to a dimensional one², there are some further important obstacles that they have not addressed, which make such a transformation highly unlikely.

Recognizing that such a change will be a challenge, Lahey et al³ note that it will be essential to demonstrate that a hierarchical dimensional diagnostic approach improves patient outcomes. If patient outcomes are not demonstrably better, it will be difficult to convince the clinical community that it is worth the effort to learn a new diagnostic language.

Clinician surveys demonstrating acceptance of a dimensional approach are not sufficient to justify a change. It will also not be enough to demonstrate that patients who are evaluated using a dimensional model improve with treatment. No doubt many patients will get better. Such a research design is analogous to an open-label medication trial. In an open-label treatment study, some patients get better, but that does not mean that the medication is effective.

To warrant an overhaul of the approach towards the assessment and diagnosis of patients needing mental health treatment, it will be necessary to conduct a randomized controlled treatment trial. Patients will be randomized to be evaluated under the categorical or dimensional approaches, treated per usual clinical practice, and then outcome assessed. I would predict that such a study will find equivalent outcomes in the two groups.

I would not expect a difference in outcome because a relatively large group of patients will respond to the non-specific aspects of treatment, no matter how they are assessed and diagnosed. In psychiatric practice, where the vast majority of patients are prescribed medication, the placebo effect accounts for much of the response to pharmacological intervention⁴. In placebo-controlled studies of major depressive disorder, bipolar disorder, schizophrenia, obsessive-compulsive disorder and anxiety disorders, it can be conservatively estimated that there is a 30% placebo response rate⁴⁻⁸. Thus, a sizable number of patients would be expected to have a positive treatment response regardless of the classification approach.

Another relatively large group of patients, albeit smaller than the placebo response group, will not respond to treatment no matter how they are assessed and diagnosed. Perhaps 20% of patients fall into this treatment resistant group.

One may quibble about the exact percentage of placebo responders and treatment resistant patients that would make up the sample in such a study. And the size of these groups will be affected, in part, by the diagnostic composition of the sample. Nonetheless, I would estimate that the response trajectory of approximately 50% of patients we treat in clinical practice is

largely pre-ordained, and the classification system will be irrelevant with regards to whether or not these patients get better.

For the remaining 50% of the patients in the sample, the question is: in how many will a different approach towards classification result in improved outcome? Given the broad-based efficacy of some medications and psychotherapeutic techniques, I suspect that the positive impact of a new classification would be modest and apply to no more than half of these patients. Thus, I would estimate that diagnostic precision has the potential to improve outcome in, at most, 25% of a sample of patients. To be sure, this is not an insignificant number of patients. However, it makes it difficult to demonstrate that a new classification approach is superior to the already established one.

Let's consider the attempt to demonstrate improved outcome based on one classification approach over another from a different perspective. How many patients will have a better outcome because they are treated differently than had they been diagnosed according to the current classification system? I would expect that the treatment of most patients would be the same regardless of the diagnostic approach. For example, selective serotonin reuptake inhibitors and serotonin-norepinephrine reuptake inhibitors will be prescribed whether patients were diagnosed with major depressive disorder or generalized anxiety disorder, or have elevated scores on an internalizing dimension with high scores on depression or fear subfactors. Because treatment will be different in only a minority of cases, it will be difficult to demonstrate that the new and improved diagnostic approach results in a better outcome.

Aside from the difficulty in convincing the mental health clinical community of the benefits of a hierarchical dimensional approach, a significant practical problem with the possible paradigm shift in psychiatric classification is the adoption of such an approach by providers who are not mental health specialists. A substantial proportion of mental health care is delivered outside of the specialty care sector. Convincing mental health professionals to change will be a big enough lift. Convincing non-mental

health professionals such as primary care providers to learn a dramatically different way to conceptualize and evaluate psychopathology seems highly unlikely. It is not tenable for different segments of the health care community to use different diagnostic approaches.

Finally, one cannot ignore the potential political forces that would oppose a change because of possible lost revenue. Might the American Psychiatric Association resist a change because of the possible loss of income accrued from the publication of the DSMs and the DSM library? Might the pharmaceutical industry oppose a change that could compromise their efforts to develop and sell new pharmaceuticals while regulatory agencies determine how to evaluate products for patients assessed under the new conceptualization of psychopathology?

In conclusion, an empirically supported system of classifying psychopathology is, of course, highly desirable. But let's not throw out the proverbial baby with the bathwater. While there may be problems with the current diagnostic systems, there is also a robust empirical literature providing evidence of validity. Despite their limitations, before I put forth the time and effort to learn and use HiTOP, or a HiTOP-like system, I will need to see data demonstrating that this will improve the care I provide to my patients. Specifically, I would need to see studies showing that more of my patients are likely to get better.

Mark Zimmerman

Department of Psychiatry and Human Behavior, Brown University School of Medicine, Rhode Island Hospital, Providence, RI, USA

1. Clark LA, Watson D, Reynolds S. *Annu Rev Psychol* 1995;46:121-53.
2. Ruggero CJ, Kotov R, Hopwood CJ et al. *J Consult Clin Psychol* 2019;87:1069-84.
3. Lahey BB, Moore TM, Kaczurkin AN et al. *World Psychiatry* 2021;20:57-63.
4. Li F, Nasir M, Olten B et al. *CNS Drugs* 2019;33:971-80.
5. Sysko R, Walsh BT. *J Clin Psychiatry* 2007;68:1213-17.
6. Leucht S. *Schizophr Res* 2017;201:315-23.
7. Furukawa TA, Cipriani A, Atkinson LZ et al. *Lancet Psychiatry* 2016;3:1059-66.
8. Kotzalidis GD, Del Casale A, Simmaco M et al. *Curr Neuropharmacol* 2019;17:741-74.

DOI:10.1002/wps.20815