

Letter

Getting Serious about Variation: Lessons for Clinical Neuroscience (A Commentary on 'The Myth of Optimality in Clinical Neuroscience')

Alexander J. Shackman^{1,2,3,6,*}
and Andrew S. Fox^{4,5,6,*}

Psychiatric illness poses a significant burden on global public health and the economy. Recent decades have witnessed the development of powerful new tools for quantifying variation in the genome and brain, leading to initial optimism that psychopathology might soon be defined and diagnosed on the basis of objective biological assays. Yet, major breakthroughs have proven elusive. In psychiatric genetics, the first wave of small-scale studies proved difficult to replicate, leading some to question whether gene hunting should be abandoned altogether [1]. The resulting period of crisis and reflection ultimately motivated the widespread adoption of more rigorous analytic approaches and the development of mega-cohorts and consortia with the tens of thousands of subjects needed to reliably detect subtle gene–disease associations [1]. This second-wave research demonstrates that the amount of disease-relevant information captured by the vast majority of individual genetic variants (loci) is vanishingly small <1% [1]. In short, while genomic approaches have proven valuable for discovering new molecular targets and validating existing therapeutics, they are not useful for routine screening or diagnosis.

Whither neuroscience? In their thoughtful review, Holmes and Patrick (H&P) make it clear that that neuroimaging research

faces a similar kind of crisis, and that there is no universal, unconditionally optimal pattern of brain structure or function [2]. These insights have enormous implications for a range of stakeholders, from researchers and clinicians to funding agencies and policy makers. Here, we highlight a few of their most important conclusions and expand on their recommendations for increasing the impact of clinical neuroscience research.

Echoing other recent commentators [3–6], H&P remind us that, as neuroimaging samples have grown larger, providing more stable and precise estimates of brain–disease associations, it has become painfully clear that popular candidate neuromarkers of psychiatric disease and risk – amygdala activity and connectivity, striatal reactivity to reward, hippocampal volume, and so on – explain a statistically significant, but small amount of disease-relevant information (e.g., risk, status, treatment response, course). Like common genetic markers, they are far too small to be useful in the clinic; they simply lack the degree of sensitivity and specificity required to benefit individual patients [5].

So, where do we go from here? H&P make several useful suggestions, emphasizing that 'it is unlikely that we will achieve a breakthrough in our understanding of how the brain's intricate functions give rise to psychiatric illness by investigating a handful of candidate biomarkers at a time'. They highlight the utility of developing more complex multivariate approaches and the necessity of very large data sets (see [Box 1](#) for detailed recommendations).

H&P also stress the need go beyond data that are simply 'Big'. Marshalling a range of evidence, they emphasize that particular behavioral, psychological, and neurobiological traits (e.g., amygdala hyper-

reactivity) are typically neither good nor bad. Most are associated with a combination of costs and benefits, with the ultimate consequences for health and disease conditional on the larger environment (e.g., presence of danger, exposure to stress or adversity) and neural context (e.g., integrity of countervailing regulatory systems; see also [7,8]). Making sense of this complexity, H&P tell us, requires data sets that are 'Broad', encompassing measures of 'brain structure and function as well as diverse clinical, demographic, behavioral, genetic phenotypes'. Doing so promises to accelerate the development of evidence-based risk calculators and stratified treatment strategies that respect the diversity of human neurobiology [9].

One point not addressed in detail by H&P is the relationship between biomarker development and mechanistic research. Consider the case of genomics. Second-wave genetics research demonstrates that the small-but-reliable associations uncovered by large-scale studies can pinpoint biological pathways with substantial 'phenotypic' consequences when manipulated with drugs or other direct perturbations. This raises the possibility that while small-but-reliable associations between neuromarkers and disease are not clinically useful, they could still prove helpful for discovering and prioritizing targets for mechanistic work or for assessing the degree to which discoveries made in animal models translate to diverse human phenotypes and contexts [10]. Such work is likely to be especially valuable when it goes beyond 'the usual suspects' and identifies previously unknown or underexplored neuromarkers that are amenable to mechanistic exploration.

With new tools for understanding biological variation comes an opportunity and a responsibility to use them to improve the lives of patients. Developing better

Box 1. Recommendations and Best Practices for Clinical Neuroscience Research

1. Dimensional phenotypes. *Diagnostic and Statistical Manual of Mental Disorders, 5th Edition* (DSM-5) diagnoses present barriers to the development of useful neuromarkers (marked heterogeneity and comorbidity, poor inter-rater reliability) [12]. Narrower dimensional phenotypes (e.g., anxiety, anhedonia) offer substantial increases in reliability and power and opportunities to understand transdiagnostic mechanisms.

2. Big and broad data. Mega-cohorts encompassing diverse measures provide increased power and more precise and realistic estimates of brain–disease associations. They afford opportunities for assessing generalizability, controlling potential confounds (e.g., motion artifact in imaging studies), and identifying disease- and risk-relevant environments.

3. Aggregate. Machine learning approaches, actuarial risk calculators, and related techniques that aggregate multiple sources of neural (e.g., activity, connectivity, and anatomy) and non-neural information are more likely to yield clinically useful tools than isolated ‘hot spots’ of brain function or structure (or focusing on the brain alone).

4. Cross-validate. Absent adequate cross-validation procedures, estimates of neuromarker performance are likely to be inflated. Separate cohorts for model training/discovery and testing/replication are the gold standard and serve as an important brake on premature application based on overly rosy preliminary results.

5. Incremental validity. Simple paper-and-pencil measures of personality and actuarial approaches that leverage readily available demographic/clinical information (e.g., smoking status) often outperform more sophisticated and expensive neuromarkers. Absent head-to-head tests of ‘incremental validity,’ the clinical value of neuroimaging remains unknown.

6. Reliability. In contrast to clinical measures, the test–retest reliability of most neuromarkers is unclear. Reliability is typically assessed in small, nonrepresentative samples, precluding definitive conclusions. We urge investigators to assess and report reliability in larger, more diverse samples.

biomarkers is important for a variety of reasons, from screening and diagnosis to treatment stratification and monitoring [3]. Developing neuromarkers that more seriously reckon with the complex interactions of human brains, contexts, and outcomes would accelerate the development of new therapeutics and the repurposing of existing ones [11]. H&P remind us that this is a major challenge, one that will require substantial time and resources, new kinds of multidisciplinary collaborations and training models, and a sober assessment of what particular kinds of

neuromarkers really can and cannot do for patients and clinicians.

Acknowledgments

We acknowledge financial support from the California National Primate Center; National Institutes of Health (DA040717 and MH107444); University of California, Davis; and University of Maryland, College Park. We thank B. Roberts and R. Poldrack for insightful discussions and L. Friedman and K. DeYoung for assistance.

¹Department of Psychology, University of Maryland, College Park, MD 20742, USA

²Neuroscience and Cognitive Science Program, University of Maryland, College Park, MD 20742, USA

³Maryland Neuroimaging Center, University of Maryland, College Park, MD 20742 USA

⁴Department of Psychology, University of California, Davis, CA 95616, USA

⁵California National Primate Research Center, University of California, Davis, CA 95616, USA

⁶Both authors contributed equally to this work; the order of authors was determined by a coin flip

*Correspondence:

shackman@umd.edu (A.J. Shackman) and
dfox@ucdavis.edu (A.S. Fox).

<https://doi.org/10.1016/j.tics.2018.02.009>

References

- Sullivan, P.F. *et al.* (2018) Psychiatric genomics: an update and an agenda. *Am. J. Psychiatry* 175, 15–27
- Holmes, A.J. and Patrick, L.M. (2018) The myth of optimality in clinical neuroscience. *Trends Cogn. Sci.* 22, 214–257
- Woo, C.W. *et al.* (2017) Building better biomarkers: brain models in translational neuroimaging. *Nat. Neurosci.* 20, 365–377
- Kapur, S. *et al.* (2012) Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? *Mol. Psychiatry* 17, 1174–1179
- Savitz, J.B. *et al.* (2013) Clinical application of brain imaging for the diagnosis of mood disorders: the current state of play. *Mol. Psychiatry* 18, 528–539
- Jollans, L. and Whelan, R. (2016) The clinical added value of imaging: a perspective from outcome prediction. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* 1, 423–432
- Kendler, K.S. and Halberstadt, L.J. (2013) The road not taken: life experiences in monozygotic twin pairs discordant for major depression. *Mol. Psychiatry* 18, 975–984
- Kendler, K.S. (2012) The dappled nature of causes of psychiatric illness: replacing the organic-functional/hardware-software dichotomy with empirically based pluralism. *Mol. Psychiatry* 17, 377–388
- Paulus, M.P. (2015) Pragmatism instead of mechanism: a call for impactful biological psychiatry. *JAMA Psychiatry* 72, 631–632
- Fox, A.S. and Shackman, A.J. (in press) The central extended amygdala in fear and anxiety: closing the gap between mechanistic and neuroimaging research. *Neurosci. Lett.* Published online November 30, 2017. <https://doi.org/10.1016/j.neulet.2017.11.056>
- Wager, T.D. and Woo, C.-W. (2015) fMRI in analgesic drug discovery. *Sci. Transl. Med.* 7, 274fs6
- Kotov, R. *et al.* (2017) The hierarchical taxonomy of psychopathology (HiTOP): a dimensional alternative to traditional nosologies. *J. Abnorm. Psychol.* 126, 454–477

Supplementary References

Alexander J. Shackman*^{1,2,3} and Andrew S. Fox*^{4,5}

¹Department of Psychology, ²Neuroscience and Cognitive Science Program, and ³Maryland Neuroimaging Center, University of Maryland, College Park, MD 20742 USA. ⁴Department of Psychology and ⁵California National Primate Research Center, University of California, Davis, CA 95616 USA

* A.J.S. and A.S.F contributed equally to this work; the order of authors was determined by a coin flip

Please Address Manuscript Correspondence To:

Dr. Alexander J. Shackman (shackman@umd.edu)

Department of Psychology, University of Maryland, College Park MD 20742 USA

Dr. Andrew S. Fox (fox.drew@gmail.com)

Department of Psychology, University of California, Davis, CA 95616 USA

ANNOTATED SUPPLEMENTARY BIBLIOGRAPHY

- Psychiatric illness is a leading burden on global public health and the economy (DiLuca & Olesen, 2014; Global Burden of Disease Collaborators, 2016).
- Initial optimism that the miseries of psychopathology might soon be defined and diagnosed on the basis of objective biological assays (Casey et al., 2013; Charney et al., 2002; Hyman, 2007).
- The second-wave of psychiatric genetics research (Bogdan, Baranger, & Agrawal, *in press*; Chabris, Lee, Cesarini, Benjamin, & Laibson, 2015; Plomin & von Stumm, 2018; Sullivan et al., 2018; Timpson, Greenwood, Soranzo, Lawson, & Richards, 2018).
- GWAS remains a critical tool for validating existing therapeutics and for discovering and prioritizing disease-relevant molecular pathways. This reflects the fact that small-but-reliable genetic associations can be used to uncover biological pathways that have large ‘phenotypic’ consequences when manipulated with drugs or other mechanistic approaches (Gandal, Leppa, Won, Parikshak, & Geschwind, 2016; Pankevich, Altevogt, Dunlop, Gage, & Hyman, 2014; Sullivan et al., 2018; Timpson et al., 2018).
- Popular candidate neuromarkers of psychiatric disease and risk explain a significant, but small amount of disease-relevant information (e.g., risk, status, treatment response, course)—far too small to be useful in the clinic (Prata, Mechelli, & Kapur, 2014; Savitz, Rauch, & Drevets, 2013). For example:
 - *Amygdala function* (Fox et al., 2015; Kaczkurkin et al., 2016; Sjouwerman, Scharfenort, & Lonsdorf, 2017; Stout, Shackman, Pedersen, Miskovich, & Larson, 2017; Swartz, Knodt, Radtke, & Hariri, 2015)
 - *Striatal reactivity to reward* (Stringaris et al., 2015)
 - *Hippocampal volume* (Logue et al., 2018)
- Paper-and-pencil measures of personality and actuarial approaches that leverage readily available demographic and clinical information often outperform candidate neuromarkers (e.g., Heinrich et al., 2016; Whelan et al., 2014), with some interesting exceptions (Jollans & Whelan, 2016; Poldrack et al., 2018).
- Best practices’ are likely to enhance the scientific and clinical value of psychiatric neuroimaging research (Gabrieli, Ghosh, & Whitfield-Gabrieli, 2015; Holmes & Yeo, 2015; Jollans & Whelan, 2016; Kapur, Phillips, & Insel, 2012; Poldrack et al., 2018; Rosenberg, Casey, & Holmes, *in press*; Savitz et al., 2013; Woo, Chang, Lindquist, & Wager, 2017; Yarkoni & Westfall, 2017).

- **Dimensional Phenotypes.** DSM-5 diagnoses present significant barriers to the development of clinically useful neuromarkers, including substantial symptom heterogeneity, frequent comorbidity, and low inter-rater reliability (i.e., uncertain ‘ground truth’) (Chmielewski, Clark, Bagby, & Watson, 2015; Fried, 2017; Galatzer-Levy & Bryant, 2013; Hasin et al., 2015; Kotov et al., 2017; Krueger et al., *in press*; Regier et al., 2013; Zald & Lahey, 2017).
- **Big Data.** Mega-cohorts provide increased power and more precise and realistic estimates of brain-disease associations. They afford important opportunities for assessing generalizability (i.e., they guard against model ‘over-fitting’); and for investigating and controlling confounding demographic and imaging variables (e.g., motion) (LeWinn, Sheridan, Keyes, Hamilton, & McLaughlin, 2017; Miller et al., 2016; Schnack & Kahn, 2016; Smith & Nichols, 2018; Varoquaux, *in press*).
- **Reliability.** In contrast to clinical measures, the reliability (e.g., test-retest stability) of most neuromarkers remains unclear (Lilienfeld & Treadway, 2016; Tomarken, 1995). Often reliability is assessed using small ($n < 30$), non-representative samples (e.g., Cannon, Cao, Mathalon, Gee, & consortium, 2018), precluding definitive conclusions about clinical utility.
- **Rigor and Reproducibility.** Understanding the neurobiological underpinnings of mental illness demands rigorous, cumulative research. Yet, the actual reproducibility of biomedical research is alarmingly low (Fox, Lapate, Davidson, & Shackman, 2018; Munafò et al., 2017). Pre-registering study hypotheses and sharing data are easy ways to enhance reproducibility, increase power, and accelerate progress.

REFERENCES

- Bogdan, R., Baranger, D. A. A., & Agrawal, A. (*in press*). Polygenic risk scores in clinical psychology: Bridging genomic risk to individual differences. *Annual review of clinical psychology*.
- Cannon, T. D., Cao, H., Mathalon, D. H., Gee, D. G., & consortium, N. (2018). Reliability of an fMRI paradigm for emotional processing in a multisite longitudinal study: Clarification and implications for statistical power. *Human Brain Mapping, 39*, 599-601.
- Casey, B. J., Craddock, N., Cuthbert, B. N., Hyman, S. E., Lee, F. S., & Ressler, K. J. (2013). DSM-5 and RDoC: progress in psychiatry research? *Nat Rev Neurosci, 14*(11), 810-814.
- Chabris, C. F., Lee, J. J., Cesarini, D., Benjamin, D. J., & Laibson, D. I. (2015). The fourth law of behavioral genetics. *Current directions in psychological science, 24*, 304-312.
- Charney, D. S., Barlow, D. H., Botteron, K., Cohen, J. D., Goldman, D., Gur, R. E., . . . Zalcman, S. J. (2002). Neuroscience research agenda to guide development of a pathophysiologically based classification system. In D. J. Kupfer, M. B. First, & D. A. Regier (Eds.), *A research agenda for DSM-V* (pp. 31-84). Washington, DC: American Psychiatric Association.
- Chmielewski, M., Clark, L. A., Bagby, R. M., & Watson, D. (2015). Method matters: Understanding diagnostic reliability in DSM-IV and DSM-5. *J Abnorm Psychol, 124*, 764-769.
- DiLuca, M., & Olesen, J. (2014). The cost of brain diseases: a burden or a challenge? *Neuron, 82*, 1205-1208.
- Fox, A. S., Lapate, R. C., Davidson, R. J., & Shackman, A. J. (2018). The nature of emotion: A research agenda for the 21st century. In A. S. Fox, R. C. Lapate, A. J. Shackman, & R. J. Davidson (Eds.), *The nature of emotion. Fundamental questions* (2nd ed.). New York: Oxford University Press.
- Fox, A. S., Oler, J. A., Shackman, A. J., Shelton, S. E., Raveendran, M., McKay, D. R., . . . Kalin, N. H. (2015). Intergenerational neural mediators of early-life anxious temperament. *Proceedings of the National Academy of Sciences USA, 112*, 9118-9122.

- Fried, E. I. (2017). Moving forward: how depression heterogeneity hinders progress in treatment and research. *Expert Rev Neurother*, *17*, 423-425.
- Gabrieli, J. D., Ghosh, S. S., & Whitfield-Gabrieli, S. (2015). Prediction as a humanitarian and pragmatic contribution from human cognitive neuroscience. *Neuron*, *85*, 11-26.
- Galatzer-Levy, I. R., & Bryant, R. A. (2013). 636,120 ways to have Posttraumatic Stress Disorder. *Perspect Psychol Sci*, *8*(6), 651-662.
- Gandal, M. J., Leppa, V., Won, H., Parikshak, N. N., & Geschwind, D. H. (2016). The road to precision psychiatry: translating genetics into disease mechanisms. *Nat Neurosci*, *19*, 1397-1407.
- Global Burden of Disease Collaborators. (2016). Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet*, *388*, 1545-1602.
- Hasin, D. S., Shmulewitz, D., Stohl, M., Greenstein, E., Aivadyan, C., Morita, K., . . . Grant, B. F. (2015). Procedural validity of the AUDADIS-5 depression, anxiety and post-traumatic stress disorder modules: Substance abusers and others in the general population. *Drug Alcohol Depend*, *152*, 246-256.
- Heinrich, A., Muller, K. U., Banaschewski, T., Barker, G. J., Bokde, A. L. W., Bromberg, U., . . . consortium, I. (2016). Prediction of alcohol drinking in adolescents: Personality-traits, behavior, brain responses, and genetic variations in the context of reward sensitivity. *Biol Psychol*, *118*, 79-87.
- Holmes, A. J., & Yeo, B. T. (2015). From phenotypic chaos to neurobiological order. *Nat Neurosci*, *18*, 1532-1534.
- Hyman, S. E. (2007). Can neuroscience be integrated into the DSM-V? *Nature Reviews Neuroscience*, *8*, 725-732.
- Jollans, L., & Whelan, R. (2016). The clinical added value of imaging: A perspective from outcome prediction. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *1*, 423-432.

- Kaczurkin, A. N., Moore, T. M., Ruparel, K., Ciric, R., Calkins, M. E., Shinohara, R. T., . . . Satterthwaite, T. D. (2016). Elevated amygdala perfusion mediates developmental sex differences in trait anxiety. *Biological psychiatry, 80*, 775-785.
- Kapur, S., Phillips, A. G., & Insel, T. R. (2012). Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? *Mol Psychiatry, 17*, 1174-1179.
- Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., . . . Zimmerman, M. (2017). The hierarchical taxonomy of psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *J Abnorm Psychol, 126*, 454-477.
- Krueger, R. F., Kotov, R., Watson, D., Forbes, M. K., Eaton, N. R., Ruggero, C. J., . . . Zimmerman, J. (*in press*). Progress in achieving empirical classification of psychopathology. *World Psychiatry*.
- LeWinn, K. Z., Sheridan, M. A., Keyes, K. M., Hamilton, A., & McLaughlin, K. A. (2017). Sample composition alters associations between age and brain structure. *Nat Commun, 8*, 874.
- Lilienfeld, S. O., & Treadway, M. T. (2016). Clashing Diagnostic Approaches: DSM-ICD Versus RDoC. *Annu Rev Clin Psychol, 12*, 435-463.
- Logue, M. W., van Rooij, S. J. H., Dennis, E. L., Davis, S. L., Hayes, J. P., Stevens, J. S., . . . Morey, R. A. (2018). Smaller hippocampal volume in posttraumatic stress disorder: A multisite ENIGMA-PGC study: Subcortical volumetry results from posttraumatic stress disorder consortia. *Biol Psychiatry, 83*, 244-253.
- Miller, K. L., Alfaro-Almagro, F., Bangerter, N. K., Thomas, D. L., Yacoub, E., Xu, J., . . . Smith, S. M. (2016). Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nat Neurosci, 19*, 1523-1536.
- Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., du Sert, N. P., . . . Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour, 1*, 21.
- Pankevich, D. E., Altevogt, B. M., Dunlop, J., Gage, F. H., & Hyman, S. E. (2014). Improving and accelerating drug development for nervous system disorders. *Neuron, 84*, 546-553.

- Plomin, R., & von Stumm, S. (2018). The new genetics of intelligence. *Nature Reviews Genetics*, *19*, 148-159.
- Poldrack, R. A., Monahan, J., Imrey, P. B., Reyna, V., Raichle, M. E., Faigman, D., & Buckholz, J. W. (2018). Predicting violent behavior: What can neuroscience add? *Trends Cogn Sci*, *22*, 111-123.
- Prata, D., Mechelli, A., & Kapur, S. (2014). Clinically meaningful biomarkers for psychosis: a systematic and quantitative review. *Neurosci Biobehav Rev*, *45*, 134-141.
- Regier, D. A., Narrow, W. E., Clarke, D. E., Kraemer, H. C., Kuramoto, S. J., Kuhl, E. A., & Kupfer, D. J. (2013). DSM-5 field trials in the United States and Canada, Part II: test-retest reliability of selected categorical diagnoses. *Am J Psychiatry*, *170*, 59-70.
- Rosenberg, M. D., Casey, B. J., & Holmes, A. J. (*in press*). Prediction complements explanation in understanding the developing brain. *Nat Commun*.
- Savitz, J. B., Rauch, S. L., & Drevets, W. C. (2013). Clinical application of brain imaging for the diagnosis of mood disorders: the current state of play. *Mol Psychiatry*, *18*, 528-539.
- Schnack, H. G., & Kahn, R. S. (2016). Detecting neuroimaging biomarkers for psychiatric disorders: Sample size matters. *Front Psychiatry*, *7*, 50.
- Sjouwerman, R., Scharfenort, R., & Lonsdorf, T. B. (2017). Individual differences in fear learning: Specificity to trait-anxiety beyond other measures of negative affect, and mediation via amygdala activation. *bioRxiv*, 233528.
- Smith, S. M., & Nichols, T. E. (2018). Statistical challenges in "Big Data" human neuroimaging. *Neuron*, *97*, 263-268.
- Stout, D. M., Shackman, A. J., Pedersen, W. S., Miskovich, T. A., & Larson, C. L. (2017). Neural circuitry governing anxious individuals' mis-allocation of working memory to threat. *Scientific Reports*, *7*, 8742.
- Stringaris, A., Vidal-Ribas Belil, P., Artiges, E., Lemaitre, H., Gollier-Briant, F., Wolke, S., . . . Consortium, I. (2015). The brain's response to reward anticipation and depression in adolescence:

Dimensionality, specificity, and longitudinal predictions in a community-based sample. *Am J Psychiatry*, *172*, 1215-1223.

Sullivan, P. F., Agrawal, A., Bulik, C. M., Andreassen, O. A., Borglum, A. D., Breen, G., . . . Psychiatric Genomics, C. (2018). Psychiatric genomics: An update and an agenda. *Am J Psychiatry*, *175*, 15-27.

Swartz, J. R., Knodt, A. R., Radtke, S. R., & Hariri, A. R. (2015). A neural biomarker of psychological vulnerability to future life stress. *Neuron*, *85*(3), 505-511.

Timpson, N. J., Greenwood, C. M. T., Soranzo, N., Lawson, D. J., & Richards, J. B. (2018). Genetic architecture: the shape of the genetic contribution to human traits and disease. *Nat Rev Genet*, *19*, 110-124.

Tomarken, A. J. (1995). A psychometric perspective on psychophysiological measures. *Psychological Assessment*, *7*, 387-395.

Varoquaux, G. (*in press*). Cross-validation failure: Small sample sizes lead to large error bars. *NeuroImage*.

Whelan, R., Watts, R., Orr, C. A., Althoff, R. R., Artiges, E., Banaschewski, T., . . . Consortium, I. (2014). Neuropsychosocial profiles of current and future adolescent alcohol misusers. *Nature*, *512*, 185-189.

Woo, C. W., Chang, L. J., Lindquist, M. A., & Wager, T. D. (2017). Building better biomarkers: brain models in translational neuroimaging. *Nat Neurosci*, *20*, 365-377.

Yarkoni, T., & Westfall, J. A. (2017). Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspect Psychol Sci*, *12*, 1100-1122.

Zald, D. H., & Lahey, B. B. (2017). Implications of the hierarchical structure of psychopathology for psychiatric neuroimaging. *Biol Psychiatry Cogn Neurosci Neuroimaging*, *2*, 310-317.