Biological Psychiatry

Realizing the Clinical Potential of Computational Psychiatry: Report From the Banbury Center Meeting, February 2019

To the Editor:

Computational psychiatry is an emerging field that examines phenomena in mental illness using formal techniques from computational neuroscience, mathematical psychology, and machine learning (1–6). These techniques can be used in a theory-driven manner to gain insight into neural or cognitive processes and in a data-driven way to identify predictive and explanatory relationships in complex datasets. The approaches complement each other: theory-driven models can be used to infer mechanisms, and the resulting measurements can be used in data-driven approaches for prediction. Recent computational studies have successfully described and measured novel mechanisms in a range of disorders (7–11), have framed disorders in new and informative ways (12), and have identified predictors of treatment response (13,14). These methods hold the potential to improve identification of relevant clinical variables and could be superior to classification based on traditional behavioral or neural data alone (15–18). However, these promising results have been slow to influence clinical practice or to improve patient outcomes.

In February 2019 a workshop was convened at the Banbury Center in Cold Spring Harbor, New York. The purpose of the meeting was to identify key developments required in the practice and infrastructure of computational psychiatry research to accelerate its ability to address real-world clinical problems in the near future. This report provides a summary of the conclusions of the meeting. At its core are suggestions to improve the measurement properties of computational assays through a rapid, iterative process that leverages coordinated waves of online and clinical testing, followed by deployment of the assays in innovative study designs to address clinically relevant questions. We particularly focus on theory-driven tasks, but where possible, the potential of data-driven approaches is highlighted. Finally, the report suggests that for the promise of computational psychiatry to be realized, the research environment must be developed to encourage large-scale, collaborative, interdisciplinary consortia. Given the focus of the report on questions of effective research practice in computational psychiatry, it is assumed that readers have some familiarity with the field. While we provide brief summaries and examples of the concepts covered in this report (Table 1 and Figure 1), readers new to the field may find it useful to review articles on the issue of clinical translation (3) or the general approaches used (1,2,5,15,16) beforehand.

We first summarize the need for computational assays with improved measurement properties and describe an iterative optimization and validation procedure by which such assays may be developed and deployed in clinically informative studies. We then consider broader adaptations to the research environment that may accelerate the translation of these techniques.

Computational Assays for Clinical Applications: What Is Missing From Current Research Practice?

Measurement. A key application of computational assays is the estimation of behavioral and cognitive variables that underlie clinical observations and measurements. Theory-driven approaches rely extensively on generative models, i.e., formal descriptions of the underlying neural and mental processes that are believed to generate observations [see (19) for an example]. Fitting generative models to observations has a number of advantages. First, it may allow identification and measurement of processes not easily captured by traditional analysis (20); second, generative models may improve the validity and reliability by which a process is measured. For instance, generative models can incorporate processes that tie different features (e.g., reaction time and choice) and modalities (e.g., behavior and physiology) together in a holistic manner (21–24). They also allow artificial data to be simulated and therefore a degree of measurement optimization to occur in silico before the assay is deployed in practice (see Table 1 for details). However, these features do not by themselves guarantee that computational assays provide reliable and valid measures of underlying processes. Rather, the measurement properties of an assay must be assessed and iteratively optimized (see Table 1 for a summary of computation-specific and general metrics of reliability and validity). Though there are notable exceptions (25–28), the issue of measurement in computational psychiatry has not yet attracted due attention. A principled and efficient process of assay development that optimizes measurement properties from the outset is a key outcome of the framework described below.

Deployment. Beyond questions of measurement, a second crucial factor in translating computational assays to clinical application is the deployment of the assays in studies that are able to address clinically useful questions. While cross-sectional designs can assess associations between symptoms and computational processes, they provide relatively limited information on the clinical utility of assays. Alternative study designs that test whether an assay provides predictive information useful to clinical decision making or the causal relationships between computationally measured processes and symptoms are likely to be especially important here. Data-driven techniques are particularly well suited to deployment in predictive studies.

Developmental Pipeline for Clinically Useful Computational Assays

Here we outline a potential framework by which promising computational assays may be efficiently developed,

https://doi.org/10.1016/j.biopsych.2019.12.026
ISSN: 0006-3223

Table 1. Key Metrics of Reliability and Validity Relevant to Computational Measures

Metric	Detailed Description
Specific Computational Measures of Reliability	
Parameter Recovery, Identifiability, and Sensitive Range	Parameter recovery is a process of validating parameterized generative models of behavior and/or neural data. It is performed in silico. A range of different parameter values are selected. These parameters are then used in the generative model to create synthetic data (at a realistic level of observation noise), which are passed back into the parameter estimation process; finally the recovered parameters are compared with the originals. The absolute difference between recovered and original parameters provides a measure of the ability of a task to estimate model parameters (if we can assume participant data can be described using a specific model), with smaller values being preferred. Parameter identifiability is a similar metric describing the degree to which model parameters exert distinct effects on the data and thus the degree to which differences in data can be confidently attributed to specific parameters. Parameter recovery and identifiability will generally not be constant over all parameter values (e.g., a very low inverse temperature parameter, which will lead a model to frequently make random choices, will impair the recovery/identifiability of the other parameters of the model), and thus it is often useful to define the sensitive range of the parameters—the range of values over which parameter recovery and identifiability are achievable.
Model Recovery	Model recovery assesses the degree to which a particular task can discriminate between different classes of generative models. This is achieved in silico by generating synthetic data using different models and then testing whether the process of model selection (see below) identifies the correct generative model. As for parameter recovery/identifiability, this can be sensitive to the range of parameters used to generate the synthetic data.
Model Selection	Where more than one model may be used to describe subject data, a process of model selection is used to select the best model. This process typically assesses the balance between the fit of the model (the degree to which the model can explain the data) and model complexity (i.e., its representational richness or flexibility to fit data in general). If two models explain the data similarly well, the simpler is preferred (Occam's razor). Taking into account the fit/complexity trade-off is important, as models with higher complexity (e.g., with more parameters) will have higher accuracy than simpler models but may be capturing measurement-specific noise (overfitting). Model selection may also concern the qualitative ability of the model to recapitulate some important features of the data. While many computational studies select a single, best model for all participants and compare model parameters between participants, it is also possible to assess whether participants differ in the model that best describes their data. The finding that data from different participants are best described by different models may in itself be interesting and may be described using an hierarchical model in which a higher level selects between separate lower level models (note that in the absence of a single model used across all participants, between-subject comparison of model parameters is not straightforward).
Common Measures of Reliability	
Test-Retest	The degree to which the measures of individuals within a group maintain a consistent relationship across time is assessed by test-retest reliability. Test-retest performance is a critical metric for tasks that are required to measure stable, traitlike, within-subject processes and for studies using correlational or longitudinal designs.
Split-Half/Interrater Reliability	Other forms of reliability such as split-half reliability or interrater reliability estimate measurement variability and may be useful in certain computational tasks.
Common Measures of Validity	
Clinical Validity	Evidence for the clinical validity of a measure is provided by associations between it and clinically important outcomes, such as symptom scores, treatment response, or illness course.
Convergent/Divergent Validity	These refer to the degree to which a measure of a construct correlates with other measures of the same construct (convergent validity) and differs from measures of other constructs (divergent validity). These metrics therefore provide an assessment of how certain we can be that we are measuring an underlying construct (convergent validity) and the degree to which our measure provides the same/different information to alternative measures (divergent validity). Questions of convergent and divergent validity have largely been overlooked in computational psychiatry. As a result, it is not clear, for example, whether learning rates for positive outcomes in the plethora of available reward learning tasks measure the same thing.
Face and Ecological Validity	This reflects the degree to which a measurement appears to subjectively measure a process (face validity) and the degree to which it captures real-life processes (ecological validity). Computational approaches are able to decompose the components of complex processes and may therefore facilitate the development of more ecologically valid measures of complex real-life interactions.
Practical Characteristics of the Measure	
Measure Duration, Complexity, and Cost	These summarize key practical costs of a measure that are essential when considering how it may be optimized for a particular study or population.
Translational Relevance of the Measure	
Cross-species Translational Potential of the Task	Depending on the specific question being addressed, the potential for a behavioral measure to be deployed in nonhuman species may be relevant for measurement selection. For example, validation of the ability of a computational assay to infer physiological mechanisms may require a degree of experimental control that cannot be achieved in humans.

See also Wilson and Collins (41) for a summary.

validated, and deployed to address clinically important questions (Figure 1). We note that the initial identification of candidate assays, informed by prior clinical, preclinical (including animal studies), and theoretical work, is crucial to this process. Rather than describing this initial identification stage, which has been a central concern of computational psychiatry (29,30), we focus on how promising assays may most efficiently be developed.

Establishing and Optimizing the Measurement Characteristics of Novel Assays

Step 1: Assay Optimization. First, the measurement factors of the assay required to address a clinical question are selected, and the structure of the assay is altered to optimize these. Table 1 outlines important metrics. The selected factors may include both specific computational properties, such as parameter identifiability (see Table 1), and practical features of an assay (e.g., duration to complete, complexity) and clinical validity (e.g., correlation with symptoms or treatment response). An objective function, a mathematical formulation that combines task metrics to produce an overall measure of performance, may be constructed to reflect the specific priorities of a research project, including factors to maximize (e.g., sensitivity to manipulations of key task variables, compliance) and minimize (e.g., task duration). The assay may then be optimized by iterative testing either in silico, using high-throughput online data collection (31), or in more deeply phenotyped clinical populations. Here, optimization occurs by systematically varying aspects of the assay's configuration (e.g., number of trials per condition, timing of stimulus presentation, reward incentives) to maximize the objective function. In some cases, this may also include hand-designed qualitative changes (e.g., to improve the task instructions used). Optimization of data-driven approaches may follow a similar trajectory with, for example, the data features being passed to a classifier that is optimized in terms of the predictive validity or the practicality of collecting the data. In effect, this step entails an expanded, recursive piloting phase during which the measurement properties of an assay are leveraged to improve its performance.

Step 2: Latent Structure Validation. Although individual model parameters may underlie specific neurocognitive processes, many clinically relevant processes are likely to consist of a latent (not directly observable) structure of relationships between multiple parameters (15,32). A useful step is therefore to describe this structure by collecting data from a range of assays within a single population of participants. Data-driven techniques such as clustering, or theory-driven techniques such as generative modeling approaches, can be used to determine the latent structure of the assays. Identified latent structures can be fed back to step 1 to inform the further development of assays, with the best performing (in terms of the metrics described in Table 1) being deployed as described below.

Deployment: Establishing the Potential of Assays as Predictors, Targets, and Mediators

Next, the potential clinical utility of assays is tested in proof-of-concept studies examining the predictive ability of the assay and/or the causal relationship between the process measured by the assay and clinically important outcomes such as symptoms.

Step 3a: Clinical Prediction and Covariation. Longitudinal observational studies may be used to assess whether an assay covaries with mental state changes or traits of interest and whether it has predictive validity, for example, by predicting response to treatment (13). The ability of cohort studies to map the development of psychiatric symptoms may be enhanced by innovative study designs such as longitudinal yet brief "natural challenge" studies (33), which make use of cohorts likely to encounter precipitative events expected to result in a change in psychiatric status (e.g., patients starting a new treatment). Prediction will typically involve a combination of theory-driven and data-driven analysis, with data-driven analyses used to establish the most powerful predictors (13,34) and to address issues of dimensionality reduction as described for latent structure validation above.

Step 3b: Causality and Treatment Targets. A second route by which computational assays may impact clinical practice is if the process measured by the assay constitutes a viable treatment target. That is, treatments may be developed specifically to alter the computationally defined process. This question hinges crucially on the causal relationship between the measured process and clinically relevant outcomes, such as symptoms or functioning. Causality is most efficiently addressed using experimental medicine designs, which manipulate the computationally measured process and then assess the consequences of the manipulation on intermediate or clinical outcomes [where this is not possible, quasi-experimental designs may also be useful (35)]. Potential manipulations may involve pharmacological, brain stimulation, cognitive, or psychotherapeutic techniques, the key issue being the ability of the intervention to engage and alter the computationally measured process.

Step 4: Clinical Efficacy. Regardless of whether the goal of using a computational assay is to predict a clinical outcome or to guide the development of a novel treatment, the efficacy of computationally informed approaches must ultimately be assessed in clinical trials. Such trials may, for example, randomly assign patients to be treated according to a predictive algorithm or standard treatment or to receive a computationally informed intervention versus a control.

In summary, these 4 steps describe a general pipeline of clinical computational assay optimization designed to yield reliable and valid assays that are deployed in clinically informative study designs.

Evolution of the Research Environment

Computational assays can be applied to preexisting datasets (36–38), and the sharing of datasets and analytic procedures is clearly of great importance. However, the process of computational assay development and deployment outlined above requires substantial structural resources well beyond those of individual laboratories. At the very least, this includes shared

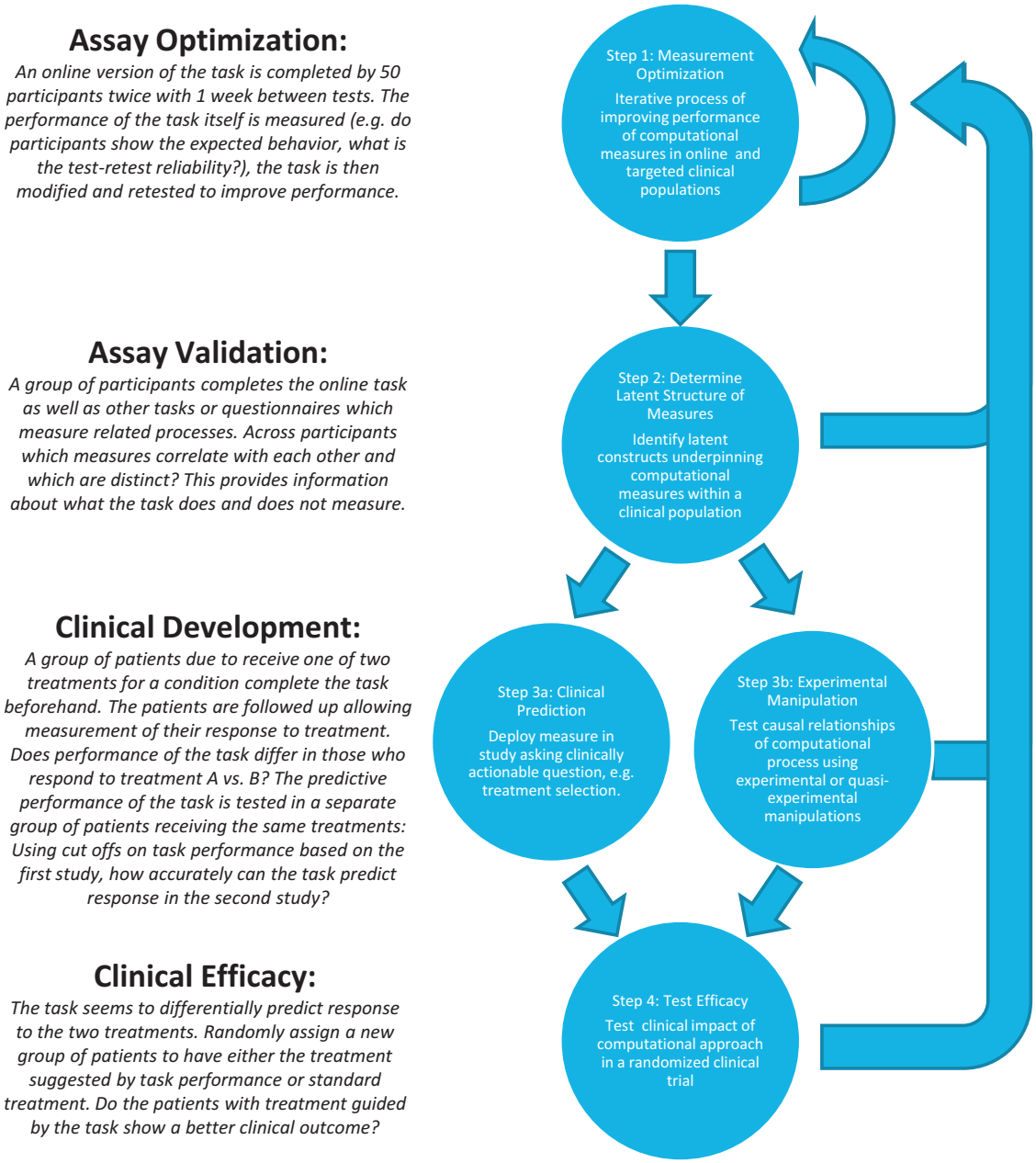

Figure 1. Suggested process by which computational measures may be optimized for deployment in clinical studies. The stages illustrated here develop the sequential roadmap outlined by Paulus *et al.* (3), who framed this process using the metaphor of drug development. We provide more detail on how each stage may be realized (see Establishing and Optimizing the Measurement Characteristics of Novel Assays) and, in this figure, present examples of studies that develop a hypothetical computational task (text in italics) through stages that are similar to phases 1–3 of the roadmap by Paulus *et al.* (3).

core infrastructure, particularly in the domain of information technology. It will necessitate common data structures, including meta-data relevant to measures, models, and populations, and common ascertainment procedures across sites that enable individual laboratories to collect high-quality behavioral and clinical data and, where relevant, physiological or biological data in a universal format (39). Curation of the data will be required to ensure that it is findable, accessible, interoperable, and reusable from the outset. Owing to rising concerns about data security on one hand and the need to provide scientists access to data on the other hand, the secure storage and aggregation of data across sites using a platform that itself may support data analysis is likely to be essential (40).

Finally, the complexity of the human mind, the diversity of processes of clinical relevance, and the range of

computational theories and interventions represent a formidable intellectual challenge. It calls for a pooling of expertise and perspectives in appropriately designed multidisciplinary consortia distributed across laboratories that have a common goal and share data and expertise. Although it is beyond the scope of this article to specify the nature and scope of such consortia, they are likely to benefit from the inclusion of, at least, expert clinicians, experimentalists, and theoreticians.

Conclusions

If computational methods are to deliver real advances for patients, we must ensure that our approaches are reliable and robust and address clinically meaningful questions. In this opinion paper, we outline processes to improve the measurement properties and deployment of computational assays and highlight the importance of interdisciplinary collaboration.

Michael Browning
Cameron S. Carter
Christopher Chatham
Hanneke Den Ouden
Claire M. Gillan
Justin T. Baker
Adam M. Chekroud
Roshan Cools
Peter Dayan
James Gold
Rita Z. Goldstein
Catherine A. Hartley
Adam Kepecs
Rebecca P. Lawson
Janaina Mourao-Miranda
Mary L. Phillips
Diego A. Pizzagalli
Albert Powers
David Rindskopf
Jonathan P. Roiser
Katharina Schmack
Daniela Schiller
Miriam Sebold
Klaas Enno Stephan
Michael J. Frank
Quentin Huys
Martin Paulus

Acknowledgments and Disclosures

This work was supported by the Medical Research Council Clinician Scientist Fellowship (Grant No. MR/N008103/1 [to MB]), National Institute for Health Research Oxford Health Biomedical Research Centre (MB), Max Planck Society (to PD), National Institutes of Health (Grant Nos. 1R01DA041528, 1R01DA047851, and 1R01AT010627 [to RZG]), National Institute on Drug Abuse Small Business Innovation Research Phase I subcontract to Soterix Medical (Grant No. 271201800035C [to RZG]), Wellcome Trust (Grant No. WT102845/Z/13/Z [to JM-M]), National Institute of Mental Health K23 Career Development Award (Grant No. K23 MH115252-01A1 [to AP]), Burroughs-Wellcome Fund Career Award for Medical Scientists (to AP), René and Susanne Braginsky Foundation (to KES), and National Institute of General Medical Sciences (Grant No. P20GM121312 [to MP]). The Banbury Center Meeting was funded by grants from the Society for Biological Psychiatry, William K. Warren Foundation, and Robert J. & Nancy D. Carney Institute for Brain Science at Brown University.

We thank John Krystal, M.D., Bruce Cuthbert, M.D., Anne Churchland, Ph.D., and Michele Ferrante, Ph.D., for invaluable discussion and advice.

MB has received travel expenses from Lundbeck for attending conferences and acted as a consultant for Janssen Research and the Centre for Human Drug Research. CC is a full-time employee of Hoffman La Roche. HDO has acted as a consultant for Lundbeck Pharmaceuticals. JTB has received consulting fees from Pear Therapeutics, BlackThorn Therapeutics, Niraxx Therapeutics, and AbleTo, Inc. AMC holds equity in Spring Care, Inc., Fitbit, Inc., and UnitedHealthcare, Inc.; is lead inventor on three patent submissions relating to treatment for major depressive disorder (U.S. Patent and Trademark Office Docket No, Y0087.70116US00, U.S. Patent and Trademark Office provisional application 62/491,660, and USPTO provisional application 62/629,041); and has consulted for Fortress Biotech, Inc., on antidepressant drug development. JG receives royalty payments from the Brief Assessment of Cognition in Schizophrenia. Over the past 3 years, MLP has received an honorarium from Sunovion Pharmaceuticals. Over the past 3 years, DAP has received consulting fees from Akili Interactive Labs, BlackThorn Therapeutics, Boehringer Ingelheim, Compass, Posit Science, and Takeda Pharmaceuticals and an honorarium from Alkermes for activities unrelated to the current article. JPR has acted as a consultant for Cambridge Cognition, Takeda, and GE Healthcare. MJF is a consultant for Hoffman La Roche. MP is an advisor to Spring Care, Inc., and he has received royalties for an article about methamphetamine in UpToDate. The other authors report no biomedical financial interests or potential conflicts of interest.

Article Information

From the Department of Psychiatry (MB), University of Oxford; Oxford Health National Health Service Foundation Trust (MB), Warneford Hospital, Oxford, United Kingdom; Departments of Psychiatry (CSC) and Psychology (CSC), University of California, Davis, Davis, California; Neuroscience and Rare Diseases (CC), Roche Pharma Research and Early Development, Roche Innovation Center, Basel, Switzerland; Donders Institute for Brain, Cognition and Behaviour (HDO, RC), Radboud University; Department of Psychiatry (RC), Radboud University Medical Center, Nijmegen, The Netherlands; School of Psychology (CMG), Trinity College Dublin, Dublin, Ireland; McLean Institute for Technology in Psychiatry (JTB) and Department of Psychiatry (DAP), McLean Hospital, Belmont; Department of Psychiatry (JTB, DAP), Harvard Medical School, Boston, Massachusetts; Spring Health (AMC), New York, New York; Max Planck Institute for Biological Cybernetics (PD), Tübingen, Germany; Maryland Psychiatric Research Center (JG), University of Maryland School of Medicine, Baltimore, Maryland; Department of Psychiatry and Neuroscience (RZG, DS), Icahn School of Medicine at Mount Sinai; Department of Psychology (CAH), New York University, New York; Cold Spring Harbor Laboratory (AK, KS), Cold Spring Harbor, New York; Department of Psychology (RPL), University of Cambridge, Cambridge; Centre for Medical Image Computing (JM-M), Department of Computer Science, Max Planck UCL Centre for Computational Psychiatry and Ageing Research (JM-M, QH), Institute of Cognitive Neuroscience (JPR), Wellcome Centre for Human Neuroimaging (KES), and Division of Psychiatry (QH), University College London, London, United Kingdom; Department of Psychiatry (MLP), University of Pittsburgh, Pittsburgh, Pennsylvania; Department of Psychiatry (AP), Yale University School of Medicine, New Haven, Connecticut; Educational Psychology (DR), Graduate School and University Center of the City University of New York, New York, New York; Department of Psychiatry and Psychotherapy (MS), Charité – Universitätsmedizin Berlin, Berlin, Germany; Translational Neuromodeling Unit (KES, QH), Institute for Biomedical Engineering, University of Zurich; Eidgenössische Technische Hochschule Zürich (KES, QH), Zurich, Switzerland; Max Planck Institute for Metabolism Research (KES), Cologne, Germany; J. & Nancy D. Carney Institute for Brain Science (MJF), Department of Cognitive, Linguistic & Psychological Sciences, Brown University, Providence, Rhode Island; and Laureate Institute for Brain Research (MP), Tulsa, Oklahoma.

MB, MJF, QH, and MP were meeting co-organizers.

Address correspondence to Michael Browning, MB.BS, Department of Psychiatry, University of Oxford, Warneford Hospital, Oxford OX3 7JX, United Kingdom; E-mail: michael.browning@psych.ox.ac.uk.

Received Aug 23, 2019; revised and accepted Dec 30, 2019.

References

1. Huys QJM, Maia TV, Paulus MP (2016): Computational psychiatry: From mechanistic insights to the development of new treatments. Biol Psychiatry Cogn Neurosci Neuroimaging 1:382–385.
2. Montague PR, Dolan RJ, Friston KJ, Dayan P (2012): Computational psychiatry. Trends Cogn Sci 16:72–80.
3. Paulus MP, Huys QJM, Maia TV (2016): A roadmap for the development of applied computational psychiatry. Biol Psychiatry Cogn Neurosci Neuroimaging 1:386–392.
4. Stephan KE, Binder EB, Breakspear M, Dayan P, Johnstone EC, Meyer-Lindenberg A, *et al.* (2016): Charting the landscape of priority problems in psychiatry, part 2: Pathogenesis and aetiology. Lancet Psychiatry 3:84–90.
5. Wang XJ, Krystal JH (2014): Computational psychiatry. Neuron 84:638–654.
6. Kishida KT, King-Casas B, Montague PR (2010): Neuroeconomic approaches to mental disorders. Neuron 67:543–554.
7. Browning M, Behrens TE, Jocham G, O'Reilly JX, Bishop SJ (2015): Anxious individuals have difficulty learning the causal statistics of aversive environments. Nat Neurosci 18:590–596.
8. Collins AGE, Albrecht MA, Waltz JA, Gold JM, Frank MJ (2017): Interactions among working memory, reinforcement learning, and effort in value-based choice: A new paradigm and selective deficits in schizophrenia. Biol Psychiatry 82:431–439.
9. Huys QJM, Pizzagalli DA, Bogdan R, Dayan P (2013): Mapping anhedonia onto reinforcement learning: A behavioural meta-analysis. Biol Mood Anxiety Disord 3:12.
10. Powers AR, Mathys C, Corlett PR (2017): Pavlovian conditioning-induced hallucinations result from overweighting of perceptual priors. Science 357:596–600.
11. Lawson RP, Mathys C, Rees G (2017): Adults with autism overestimate the volatility of the sensory environment. Nat Neurosci 20:1293–1299.
12. Braver TS, Barch DM, Cohen JD (1999): Cognition and control in schizophrenia: A computational model of dopamine and prefrontal function. Biol Psychiatry 46:312–328.
13. Chekroud AM, Zotti RJ, Shehzad Z, Gueorguieva R, Johnson MK, Trivedi MH, *et al.* (2016): Cross-trial prediction of treatment outcome in depression: A machine learning approach. Lancet Psychiatry 3:243–250.
14. Harlé KM, Stewart JL, Zhang S, Tapert SF, Yu AJ, Paulus MP (2015): Bayesian neural adjustment of inhibitory control predicts emergence of problem stimulant use. Brain 138:3413–3426.
15. Wiecki TV, Poland J, Frank MJ (2015): Model-based cognitive neuroscience approaches to computational psychiatry clustering and classification. Clin Psychol Sci 3:378–399.
16. Huys QJM, Maia TV, Frank MJ (2016): Computational psychiatry as a bridge from neuroscience to clinical applications. Nat Neurosci 19:404–413.
17. Wiecki TV, Antoniades CA, Stevenson A, Kennard C, Borowsky B, Owen G, *et al.* (2016): A computational cognitive biomarker for early-stage Huntington's disease. PLoS One 11:e0148409.
18. Brodersen KH, Schofield TM, Leff AP, Ong CS, Lomakina EI, Buhmann JM, Stephan KE (2011): Generative embedding for model-based classification of fMRI data. PLoS Comput Biol 7:e1002079.
19. Behrens TEJ, Woolrich MW, Walton ME, Rushworth MFS (2007): Learning the value of information in an uncertain world. Nat Neurosci 10:1214–1221.
20. Pulcu E, Browning M (2017): Affective bias as a rational response to the statistics of rewards and punishments [published correction appears in Elife 2017; 6:e32902. Elife 6:e27879.
21. Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K, *et al.* (2010): Research domain criteria (RDoC): Toward a new classification framework for research on mental disorders. Am J Psychiatry 167:748–751.
22. Hedge C, Powell G, Bompas A, Vivian-Griffiths S, Sumner P (2018): Low and variable correlation between reaction time costs and accuracy costs explained by accumulation models: Meta-analysis and simulations. Psychol Bull 144:1200–1227.
23. Price RB, Brown V, Siegle GJ (2019): Computational modeling applied to the dot-probe task yields improved reliability and mechanistic insights. Biol Psychiatry 85:606–612.
24. Kessels RPC (2019): Improving precision in neuropsychological assessment: Bridging the gap between classic paper-and-pencil tests and paradigms from cognitive neuroscience. Clin Neuropsychol 33:357–368.
25. Moutoussis M, Bullmore ET, Goodyer IM, Fonagy P, Jones PB, Dolan RJ, *et al.* (2018): Change, stability, and instability in the Pavlovian guidance of behaviour from adolescence to young adulthood. PLoS Comput Biol 14:e1006679.
26. Shahar N, Hauser TU, Moutoussis M, Moran R, Keramati M, consortium NSPN, Dolan RJ (2019): Improving the reliability of model-based decision-making estimates in the two-stage decision task with reaction-times and drift-diffusion modeling. PLoS Comput Biol 15: e1006803.
27. Hedge C, Powell G, Sumner P (2018): The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. Behav Res Methods 50:1166–1186.
28. Enkavi AZ, Eisenberg IW, Bissett PG, Mazza GL, MacKinnon DP, Marsch LA, Poldrack RA (2019): Large-scale analysis of test-retest reliabilities of self-regulation measures. Proc Natl Acad Sci U S A 116:5472–5477.
29. Maia TV, Huys QJM, Frank MJ (2017): Theory-based computational psychiatry. Biol Psychiatry 82:382–384.
30. Kurth-Nelson Z, O'Doherty JP, Barch DM, Denève S, Durstewitz D, Frank MJ, *et al.* (2016): Computational approaches for studying mechanisms of psychiatric disorders. In: Redish D, Gordon JA, editors. Computational Psychiatry. Cambridge, MA: MIT Press.
31. Gillan CM, Daw ND (2016): Taking psychiatry research online. Neuron 91:19–23.
32. Poldrack RA, Yarkoni T (2016): From brain maps to cognitive ontologies: Informatics and the search for mental structure. Annu Rev Psychol 67:587–612.
33. Clarke P, MacLeod CM, Shirazee N (2008): Prepared for the worst: Readiness to acquire threat bias and susceptibility to elevate trait anxiety. Emotion 8:47–57.
34. Calhoun VD, Lawrie SM, Mourao-Miranda J, Stephan KE (2017): Prediction of individual differences from neuroimaging data. Neuroimage 145:135–136.
35. Marinescu IE, Lawlor PN, Kording KP (2018): Quasi-experimental causality in neuroscience and behavioural research. Nat Hum Behav 2:891–898.
36. Etkin A, Patenaude B, Song YJC, Usherwood T, Rekshan W, Schatzberg AF, *et al.* (2015): A cognitive-emotional biomarker for predicting remission with antidepressant medications: A report from the iSPOT-D trial. Neuropsychopharmacology 40:1332–1342.
37. Trivedi MH, McGrath PJ, Fava M, Parsey RV, Kurian BT, Phillips ML, *et al.* (2016): Establishing moderators and biosignatures of antidepressant response in clinical care (EMBARC): Rationale and design. J Psychiatr Res 78:11–23.
38. Auchter AM, Hernandez Mejia M, Heyser CJ, Shilling PD, Jernigan TL, Brown SA, *et al.* (2018): A description of the ABCD organizational structure and communication framework. Dev Cogn Neurosci 32:8–15.
39. Gorgolewski KJ, Alfaro-Almagro F, Auer T, Bellec P, Capotă M, Chakravarty MM, *et al.* (2017): BIDS apps: Improving ease of use, accessibility, and reproducibility of neuroimaging data analysis methods. PLoS Comput Biol 13:e1005209.
40. Smucny J, Barch DM, Gold JM, Strauss ME, MacDonald AW, Boudewyn MA, *et al.* (2019): Cross-diagnostic analysis of cognitive control in mental illness: Insights from the CNTRACS consortium. Schizophr Res 208:377–383.
41. Wilson RC, Collins A (2019): Ten simple rules for the computational modeling of behavioral data. Elife 8:e49547.