PERSPECTIVE   OPEN

Barriers and solutions to the adoption of translational tools for computational psychiatry

David Benrimoh[1], Victoria Fisher [2], Catalina Mourgues[2], Andrew D. Sheldon[2], Ryan Smith [3] and Albert R. Powers [2✉]

Computational psychiatry is a field aimed at developing formal models of information processing in the human brain, and how alterations in this processing can lead to clinical phenomena. There has been significant progress in the development of tasks and how to model them, presenting an opportunity to incorporate computational psychiatry methodologies into large- scale research projects or into clinical practice. In this viewpoint, we explore some of the barriers to incorporation of computational psychiatry tasks and models into wider mainstream research directions. These barriers include the time required for participants to complete tasks, test-retest reliability, limited ecological validity, as well as practical concerns, such as lack of computational expertise and the expense and large sample sizes traditionally required to validate tasks and models. We then discuss solutions, such as the redesigning of tasks with a view toward feasibility, and the integration of tasks into more ecologically valid and standardized game platforms that can be more easily disseminated. Finally, we provide an example of how one task, the conditioned hallucinations task, might be translated into such a game. It is our hope that interest in the creation of more accessible and feasible computational tasks will help computational methods make more positive impacts on research as well as, eventually, clinical practice.

INTRODUCTION

A major area of both need and opportunity in the field of psychiatry is establishing the link between observed symptoms (e.g., the criteria we use to diagnose and characterize illnesses) and neurobiological findings (e.g., alterations in functional connectivity or gene expression) via alterations in established information processing carried out by the brain. The challenge posed here is that of mapping symptoms onto neurobiology in a principled manner, based on a sound understanding of what the brain is computing, how this computation is implemented neurobiologically, what specific computations are altered in disease, what changes in neurobiological processes account for these alterations, and how these alterations give rise to symptoms. This work is carried out in the hope that an improved mechanistic understanding of psychiatric illnesses will lead to new treatments and biomarkers that could be linked to treatment mechanisms of action.

The field of computational psychiatry aims to fill this need [1–4]. Computational psychiatry can be broadly divided into two main fields [5, 6]. The first focuses on prediction, using primarily data-driven methods and aimed at discovering useful models for predicting outcomes or events of interest, such as remission with antidepressant treatment [7–10]. Predictive models have begun to be implemented in clinical practice [11–13], and the major challenge for these models at present is validation, replication, optimization of clinical implementation, and interpretability [14]. The second domain focuses on theory using purely in silico models and/or models meant to explain the generation of collected data with the intention of better understanding the

pathophysiology of disease. The theory-driven domain of computational psychiatry is less often implemented in clinical and large-scale research efforts, and will be the focus of this paper; further references to computational psychiatry will therefore refer to this branch of the field, aimed at better understanding disease processes. There are many proposed computational approaches for studying relevant psychiatric problems, ranging from reinforcement learning models to hierarchical Bayesian models, amongst a number of other approaches.. However, common to each is the aim of identifying latent states that drive both normative brain function as well as symptom and disease development: as in physics, we should be able to formally describe models by which the brain processes information and then design experiments and gather data that specifically support, refute, or call for a modification of the models proposed.

The key element of computational psychiatry is that many of these models, which can be used to simulate behavior and which can be fit to observed data, contain parameters corresponding to latent states or processes that are not otherwise easily observable. These parameters can then in turn be correlated not only with behavior, but with various kinds of neural measures [15–17], ultimately linking behavior and neural implementation via computation. This approach has been applied to a number of conditions such as psychosis [15, 18–23], anxiety and depression [17, 24–26], obsessive-compulsive disorder [27], substance use [28, 29], and transdiagnostic samples [17, 30, 31] and remains an area of emphasis for major funding sources in psychiatric research, including the National Institute for Mental Health (NIMH).

[1]McGill University School of Medicine, Montreal, QC, Canada. [2]Yale University School of Medicine and the Connecticut Mental Health Center, New Haven, CT, USA. [3]Laureate Institute for Brain Research, Tulsa, OK, USA. ✉email: albert.powers@yale.edu

Indeed, computational approaches are viewed by some as the field's best option for bringing psychiatric nosology into step with that of the rest of medicine, by linking disease manifestations and distal etiologies via distinct mechanisms [2, 3].

It is worthwhile discussing one illustrative example of this approach that has recently been applied to understanding hallucinations, which may be formulated as arising because of a tendency to over-estimate the reliability of one's prior expectations (or *priors*, in Bayesian terms) during perception [15, 19, 21, 22]. With this hypothesis in mind, researchers formulated a task to create conditioned hallucinations (CH) by heightening expectations [15]. CH occur as a result of classical conditioning, where a subject is presented with a salient stimulus paired with a difficult-to-detect target (e.g., an image and a sound) at the same time in a repeated manner, such that in the presence of the salient stimulus (e.g., the image) and the absence of the target, the subject may hallucinate the target due to their strong expectation that the stimulus should be present. This task is modeled using the Hierarchical Gaussian Filter or HGF [32] which estimates parameters that can in turn be associated with neurobiological measures, such as neuroimaging data [15]. We will return to this illustrative example later in this article.

Adopting these methods into larger-scale translational research efforts would allow for the identification of a more complete set of computational 'phenotypes' than would be observed in smaller studies and would facilitate longitudinal characterization of computational parameter change over time. An improved understanding of the range and temporal dynamics of these parameters and whether they covary with clinical symptoms, functional impairment, or brain function, would allow for improved identification of mechanistically distinct patient groups both within and across diagnostic categories (e.g. see [17, 30]). This could, in turn, lead to novel therapeutic, diagnostic, and treatment monitoring approaches. These advances would only be possible in large cohort studies implementing standard batteries of computational tasks. Recent progress in computational methods has, we argue, created a novel opportunity to adopt these methods into large-scale research efforts, and perhaps eventually into clinical practice. Below, we will briefly discuss some of the barriers to widespread adoption into large scale translational research efforts that computational psychiatry faces, as well as some potential routes to overcoming them.

Barriers

We propose that practical concerns are most likely to act as barriers to the widespread implementation of computational approaches in psychiatry. Most important among these are time, ecological validity, and practical implementation concerns.

Time.   When designing or modifying a task intended to be suitable for computational modeling, the primary concern of the designer is generally the validity of the task in terms of its ability to capture relevant latent states that drive behavior. This is similar to the problem faced by many classical neuropsychological or psycho-physical tests, which often require many trials to establish reliable measures and which in turn can require, in some cases, one to multiple hours of testing, depending on the range of tasks included [33–35]. Recently-published task-model combinations in computational psychiatry take between 15 and 40 minutes for each assessment [20, 36–39]. Adding any *one* of these tasks may present a burden in the context of a larger study that must also collect various clinical and neurophysiological measures. In addition, each of these tasks is optimized for a certain set of parameters; as such, multiple tasks would likely be required for the recovery of an adequate computational phenotype of an individual. As these tasks have not yet been integrated into batteries optimized for feasibility, larger-scale research projects would be hard-pressed to include several of them into their protocols.

Ecological validity.   Another significant limitation of current tasks in computational psychiatry, as well as in more traditional neuropsychological testing, is the fact that they are not ecologically valid: behavior or experiences during a task most often do not reflect clinically relevant content domains, contexts, and/or symptoms as they are experienced in the real world [40]. For example, decisions aimed at maximizing small amounts of monetary reward or minimizing small shocks in the lab could plausibly engage very different prior expectations than decisions in real-world contexts to avoid feared situations or to maximize overall life satisfaction. Laboratory tasks are also generally designed to isolate certain behaviors or cognitive skills so that they can be effectively measured, modeled, and interpreted. Unfortunately, this ignores the fact that, during real-world functioning, a participant might employ multiple skills when solving a given problem or might need to solve different problems in sequence or in parallel; additionally, various affective or memory-based cues, or volatility in the environment, may interfere with function in a manner not apparent in the more sterile environment of a task [40–42]. It is important to note, however, that some neuropsychological and computational tasks (and their parameters) have been linked directly to patient experiences and outcomes [3, 17, 28–31]. Indeed, we have demonstrated that the parameter that denotes the relative overweighting of priors in the HGF model of the CH task correlates with recent hallucination severity [36]. Similarly, understanding how choices and decisions made during experiments relate to real-world behavior has long been a focus of fields such as psychology and economics, and methods and approaches have been developed in these fields to improve the ecological validity of tasks and to better define which aspects of it are most important in a given context [43, 44]. For example, there is evidence from some choice experiments that allowing people to have more time to make a decision (e.g. about vaccination), as they would in the real world, seems to reduce experiment related bias (in the case of vaccination, this was demonstrated by reduced propensity of those who were given time to think to fail internal validity tests on surveys about vaccines) [43, 45]. There are limitations to these findings, however. For example, the CH task may generate a parameter that correlates with recent hallucination severity, but important causal factors have not yet been included in the task or model: we are currently investigating more ecologically valid implementations of the task that incorporate affect and stress, given the importance of these in fluctuations of symptom severity in the real world [46].

However, it is important to note that the limitations inherent in less ecologically valid task implementations should not lead to the conclusion that these tasks are not valid; rather, they are abstractions that are often necessary in order to understand and operationalize latent constructs and that may, as discussed, provide clinically meaningful results. We argue simply that increasing ecological validity may be of use in widening our understanding of how these constructs may interact with others in more realistic environments and/or in those that more closely probe clinically relevant content domains. As we discuss in the case of the CH task, where we plan to add dimensions of affect and stress, more ecologically valid tasks and more complex environments in which tasks can be completed may be more informative as they include information about factors not present in less ecologically valid implementations of a task. However, those less ecologically valid implementations remain crucial; in many cases having access to the initial abstraction provided by a less ecologically valid task is necessary to interpret the more complex results of an ecologically valid version of the same task.

Implementation concerns.   In addition to the fact that a standard battery does not yet exist that would facilitate the adoption of these measures, there are several other practical barriers to

implementation. One is the fact that computational psychiatry remains a niche field, with few investigators being equipped to implement, refine, and interpret relevant models. This is problematic because, in many cases, investigators with relevant research questions, but without a computational background, would benefit from being able to implement a standardized version of a task that produces an output with a clear report of the results, but are prevented from doing so because user-friendly versions of tasks or models often do not exist. In the neuropsychological realm, this is a problem that is partially addressed by digitized batteries, where investigators who may not be experts in the development or implementation of cognitive tests can still make use of a standardized testing platform and utilize the results [34]. For those investigators who do have a computational background, or an interest in developing relevant expertise, the process of developing and iteratively validating models and tasks can also be prohibitive in terms of both time and expense, resulting from the large sample sizes required for testing.

Test-retest reliability. To be useful as clinical assessment tools, it is also imperative that computational model parameters can be measured repeatedly over time (e.g., at key points during treatment). Yet, as many of these tasks involve learning, decision strategies can change (e.g., become more habitual and with more confident priors) with repeated performance. This can result in poor test-retest reliability, raising the concern that the same latent computational process is not being captured at each timepoint of assessment. It is not yet clear which parameters should be expected to fluctuate in a state-like fashion, and over what time scales, and which parameters should be expected to remain more static and trait-like. Part of this issue arises from the lack of longitudinal studies in computational psychiatry; a key approach in the future will be careful temporal characterization of these parameters.

Solutions
We argue that each of the barriers above have arisen because the field of computational psychiatry has not yet pivoted from validation of constructs to implementation of tools. We propose the following solutions to address these barriers in turn.

Time. In order to reduce the time required to complete a given task, existing tasks should be redesigned with a focus on determination of the most efficient structure possible, to allow for reliable parameter estimation in the shortest amount of time. For example, it was recently determined that separation between hallucinators and controls occurs fairly early on in the CH task, information which is now being used to generate a shortened version of this task [36]. The CH task's relatively simple structure (i.e., establishing a prior and then gradually testing its strength) made a simple empirical review of the data sufficient to determine when the task could be reasonably truncated. However, other task designs may differ in ways that render this determination more complex. In these situations, simulations of data generated by a given task could be performed with the specific aim of determining how many trials are required to derive parameters of interest with tolerable accuracy; indeed, simulations aimed at generating data which can then be compared to the data produced by participants is regarded as being an important part of good practice and model validation in the field of computational psychiatry [3]. As these tasks are shortened, it will also be important to consider the tolerability of these tasks in aggregate, if a traditional battery structure is to be considered.

One way to improve tolerability and increase engagement would be to incorporate these tasks into the structure of a game, a strategy used in other disciplines, such as education [47]. This idea of a computational battery constructed in the form of a game is one that we will continue to develop.

It should be noted that for some tasks, shortening may not be possible or would lead to issues with reliability. This may be especially true where the required precision for a single subject is high. In these cases, it may be necessary to maintain the original task. If these tasks are limited in number, then their integration into standard batteries may yet be accommodated.

Reducing the time required to complete each task would be a useful first step. Another opportunity comes in designing novel tasks meant to model behavior driven by a richer set of computational parameters (e.g., providing separable estimates of various learning rates, prior precisions, epistemic drives, etc.), where many separate tasks may currently be necessary to gather this individual difference information. If it were demonstrated that parameter estimates for such a task were recoverable, this would potentially allow for fewer tasks while still providing the same types of information about each individual. This approach may be of use in replacing a number of single tasks that, for reasons described above, cannot be shortened. These tasks could then be tested in simulations to determine the optimal number of trials thought to be needed, and this could then in turn be tested empirically to determine if the number of shortened trials produces valid estimates compared to longer versions of the tasks. An example of shared parameter estimation would be of two tasks, one focused on perceptual judgements and the other on social judgements. While these two processes likely engage some independent processes, there are likely to be some common parameters shared between them, especially if symptoms exist that seem to affect both domains (e.g. paranoid hallucinations about a neighbor co-occuring in someone with paranoid beliefs about their neighbors). In estimating parameters using information from both tasks, it may be possible to efficiently estimate the shared parameters–and at the same time to determine which parameters are *not* shared, which in and of itself would be an interesting mechanistic finding.

The use of explicit computational models is actually helpful in this case: generative models of behavior are explicitly designed to account for and measure different latent states driving behavior that can appear similar when analyzed via descriptive summary statistics. Thus, in principle, the use of explicit models capable of taking into account multiple drivers of behavior would make for a maximally efficient route toward estimating as many latent states as possible at any given time.

Ecological validity. The idea of integrating tasks within a game that has a believable world, perhaps one modeled on the experiences of either the general public or specific groups of participants, may also have benefits with respect to ecological validity. By integrating tasks into this "gameworld", it would allow for participants to use multiple skills or be influenced by previous experiences within the game, while solving problems reminiscent of those they have to solve in clinically relevant real world contexts (e.g., probing issues of avoidance, approach-avoidance conflict, proximal vs. distal planning, exploration vs. exploitation with respect to social contingencies and their volatility, etc.). This in turn may increase the generalizability of results from these tasks to clinical outcomes of interest. There is also an opportunity here that goes beyond simply improving the ecological validity of current tasks: by creating a gameworld, participants would be able to make choices, approach problems in different ways, and generally act with greater agency than is possible in isolated computational tasks. This in turn would allow for the modeling of new parameters related to patient choice and their generation of action plans; these parameters in turn may reflect latent states more relevant for the generation and maintenance of various symptoms and syndromes [48], but which have not been previously measured in ecologically valid ways.

Implementation concerns. The generation of an integrated battery of efficient computational tasks may also help address some of the concerns around implementation. Firstly, a more comprehensive gameworld (with multiple nested perceptual and decision tasks) would provide a standardized environment, and could be engineered to elicit reports of participant behavior based on computational models programmed into the software. The nested tasks and models could also be made modular and modifiable, to support researchers with various levels of computational experience in their use of this methodology. Digital tools, such as games, are explicitly designed to be easy to disseminate and require minimal training for participants, with training often able to be delivered in the format of a tutorial experience within the game. Hosting these tools online would allow access to large populations of participants who otherwise would not be reached by current lab-based efforts. As such, the expense required for the development and validation of computational tasks and models would be significantly reduced, and their use would be feasible for a larger number of researchers with varying levels of expertise and resources. Improved education of clinicians on what computational psychiatry is, what benefits it can provide, and how relevant models function in an accessible manner are the most important components of an approach to address concerns around implementation. Adopting computational tools into large-scale clinical research projects may provide a test bed not only for computational models, but also for assessing different methods of educating clinicians around their utility.

Test-retest reliability.   One solution is that tasks be vetted for test-retest reliability prior to inclusion in a battery. In previous research, some tasks have been shown to exhibit higher reliability than others, and this can depend on the time elapsed between assessments [28, 31]. Another possible solution would be to have participants complete tasks multiple times before starting to use them for assessment, which could increase reliability if it allows participants to first settle into a stable strategy–which could better mimic the stable strategies they settle into when solving real-world contexts of clinical relevance. Yet another strategy to consider could be to continue to vary task contents, while keeping the abstract decision structure identical. This could in principle minimize changes in initial strategy if participants understood each of these tasks to be new.

It is important to understand what might be driving poor test-retest reliability in a given task. One key driver of poor reliability may simply be that the parameter being estimated is not stable—that it represents a state rather than a trait marker (or a mixture of both). For example, as discussed in the CH task, the parameter that measures overweighting of priors is generally higher in hallucinators than non-hallucinators, but it also fluctuates with the severity of recent hallucination symptoms [36]. Thus, this parameter both correlates with higher hallucination proneness (a trait) as well as recent hallucinations (a state). In these cases, the best approach would be to understand the temporal dynamics of parameters that vary naturally, and to determine what other clinical or computational parameters they vary with. Further research into novel metrics of parameter stability may help to better characterize which parameters should be thought of as more state-like and which as more trait-like. In other cases, a parameter may well be expected to be stable, but may apparently fluctuate over time because of the methodology used to approximate it; in these cases, novel approximation methods might be of use in improving the reliability of parameter estimation. Lastly, each of the parameters estimated by these models is instantiated within brain processes that are, themselves, dynamic over different time scales. A fuller understanding of these processes will help to constrain our expectations of corresponding parameters' stability and usefulness as state and trait markers of disease.

Potential interactions
Thus far, we have discussed barriers and their solutions in isolation from each other; in reality, these barriers and solutions may interact with each other in ways that are important to consider when designing novel tasks and models. Shortening a task, for example, may make it more palatable for use in a battery, yielding more data from more diverse populations; on the other hand, this may interfere with test reliability. Providing a more realistic environment may enhance ecological validity, but given that performance would then depend on that enriched environment, designers of tasks would need to validate whether parameters estimates are consistent across different virtual environments. Creating batteries of computational tests that could be used widely would also create a high burden of test validation: these tests would need to be extensively studied in different subpopulations prior to their deployment in order to avoid misinterpretation of results; this in turn may, in turn, hamper or delay implementation. Finally, the creation of a battery of tests may also create the false impression that the battery is complete or exhaustive, which in turn would perhaps make the integration of novel tasks and models in the future more difficult. Careful and nuanced management of batteries would therefore be necessary.

A summary of the barriers and solutions discussed is depicted in Fig. 1.

Conclusions and future directions
In this article, we have examined the time requirements, lack of ecological validity, test-retest reliability, and practical considerations such as the lack of widespread computational expertise and the need for expensive validation procedures have limited the utilization of computational psychiatry methodologies in large research initiatives. If these barriers can be overcome, we believe that one of the aspects of the utility of computational psychiatry— that is, an improved understanding of latent states and the derivation of parameters that can be linked to neurobiological measures in order to improve mechanistic understanding—will be more easily realized.

It is important at this point to note that the barriers and solutions discussed here were selected with the intent of facilitating the widespread adoption of computational psychiatry
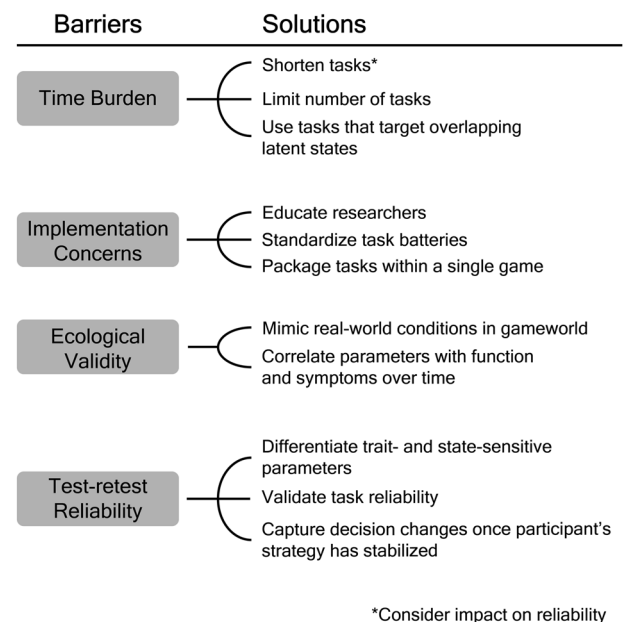

Barriers | **Solutions**

Time Burden
- Shorten tasks*
- Limit number of tasks
- Use tasks that target overlapping latent states

Implementation Concerns
- Educate researchers
- Standardize task batteries
- Package tasks within a single game

Ecological Validity
- Mimic real-world conditions in gameworld
- Correlate parameters with function and symptoms over time

Test-retest Reliability
- Differentiate trait- and state-sensitive parameters
- Validate task reliability
- Capture decision changes once participant's strategy has stabilized

*Consider impact on reliability

Fig. 1 Summary of barriers to the implementation of computational psychiatry methods into large-scale clinical research and key proposed solutions for each barrier.
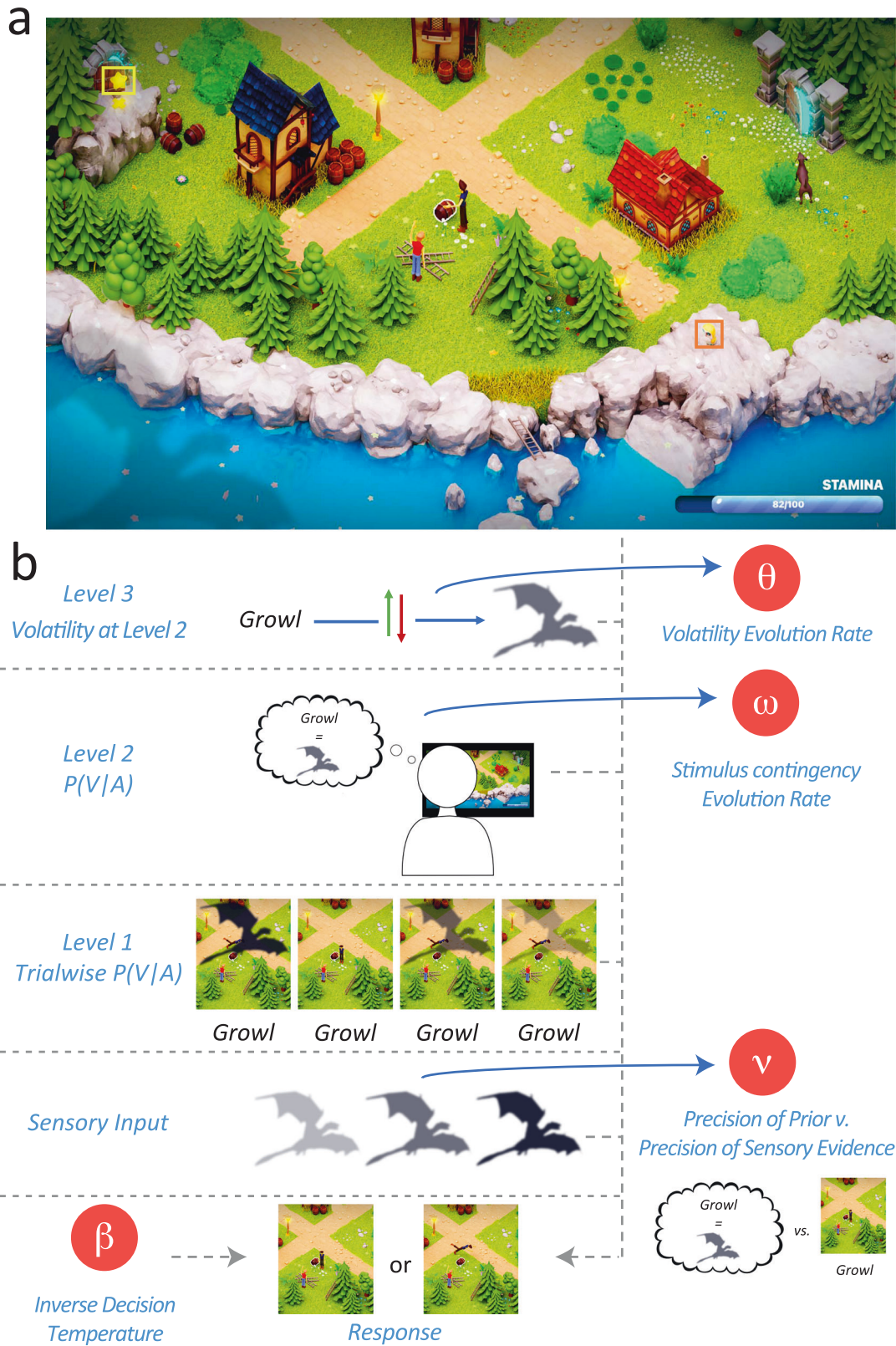
methods into large-scale research endeavors aimed at providing eventual clinical benefit. However, we believe it is important to note that this would not be the sole useful application of computational psychiatry methods. Indeed, use of these methods may facilitate mechanistic insights that could facilitate the development of novel interventions or tests that may be far simpler than the computational tests that led to their creation or discovery. Let us discuss a hypothetical example: a computational approach may identify dysfunction of a specific brain network as being key for the formation of psychotic symptoms at a particular

Fig. 2 The gameworld and the Conditioned Hallucinations (CH) task within it. a The character's avatar is located in the center of the screen in **a** which represents the gameworld. The yellow box, top left in **a**, represents a goal to be reached. The world is available for the player to explore and they may encounter computational tasks built into the environment as they seek to find paths towards the goal. Both task performance and exploration will be analyzed using computational models. As an example of this is the implementation of the visual version of the CH task: in the lower left of **a**, a dragon's shadow is present. This is part of one implementation of the CH task (see **b**). Finally, in the red box on the lower right is an owl that may serve as a target for an ongoing attention task, should it be necessary to track player attention over time. **b** In this figure, we represent the traditional hierarchical gaussian filter (HGF) model as it pertains to the Conditioned Hallucinations (CH) game in the example gameworld. The version of the CH task demonstrated here is a visual conditioned hallucination as it is more intuitive to demonstrate in a static figure, but the auditory version of the CH task can easily be implemented as well. In the game, we would modify the traditional conditioned hallucinations task such that whenever a player hears a growl and sees a dragon shadow, they should dodge it. However, if they do not see the shadow, they should not jump. As has been done successfully in other iterations of the task, the participant learns to associate the shadow with the growl and reports seeing the shadow (by dodging) even when it is absent. In the HGF, there are three levels that form an agent's perceptual model of the world in the game. Level 1 reflects the agent's belief regarding the presence/absence of the dragon shadow on any given trial (trialwise $P(V|A)$). This is reflected in their decision to dodge or not. Level 2 reflects their belief that the dragon shadow is associated with the growl (overall $P(V|A)$). Finally, level 3 represents their belief in the volatility of the association between the growl and the shadow. Based on the participant's decision to dodge or not, we can derive three important latent parameters that could act as behavior-based biomarkers of various psychiatric disorders: decision noise (β^{-1}), learning rate of the auditory-visual stimulus associations (ω), and weighting of prior expectations (ν).

pre-clinical stage. This may lead to studies using non-computational methods to probe this network, perhaps using objective measures of neural function (such as positron emission tomography or functional magnetic resonance imaging). Furthermore, if this network were amenable to modulation with neurostimulation, patients with abnormalities on testing may be candidates for preventative treatment that does not require targeting with complex computational models. This would produce clinical benefit and research methods that can be used at scale, without the computational methods being utilized beyond the provision of the initial methodological insight. In keeping with this potential approach, a recent paper demonstrated a method for turning in- silico models of medication adherence into an adherence questionnaire which could be administered without the need for a computational model [49]. We believe, based on our experiences working with computational models, that using them at scale may be extremely valuable. However, at the same time, for the reasons just discussed, we also believe that continued research into smaller scale applications and the development of more specialized models is equally important. Indeed, should the large-scale application of computational models prove too challenging at present, these smaller-scale applications will continue to provide key insights, as they have done in recent decades.

Our proposed solutions, which are by no means comprehensive or definitive, have focused on a combination of a focus on good design (i.e., re-examining tasks and shortening them when possible, vetting or adjusting tasks to ensure test-retest reliability) and an exploration of the opportunity presented by the integration of tasks into an overarching game world with nested perceptual and decision problems. Depending on its specified contents, such a gameworld could allow for greater ecological validity and the measurement of novel computational parameters, while also providing a modifiable platform that would empower researchers with varying levels of expertise and resources to begin to engage with computational psychiatry. Future work would therefore naturally be focused on the development and validation of this type of broader game environment. While this effort is in its early stages, we present here an example of how a well-validated computational task might be transported into a game world. In Fig. 2, we see both a depiction of a possible gameworld and a version of the CH task, where a player learns an association between an auditory and visual stimulus (in this case, a dragon and its growl) while navigating the gameworld, and must react in a way that may affect their gameplay (for example, if they fail to dodge when the dragon is present and growling, they may suffer a penalty). This task is modeled using the HGF, (Fig. 2b), in a manner consistent with the gameworld. Here the player is

learning the task while playing the game, and their behavior is guided by the logic of the gameworld, rather than it being dictated by the instructions of an arbitrary task. It should be noted that this example is intended for illustration and could be altered if greater relevance to everyday life is required (e.g., replacing the dragon with a more realistic threat).

In addition to constructing tasks in a gameworld, there are several other potential avenues for progress. Theoretical breakthroughs in modeling techniques (for example, computational agents capable of generating novel model architectures as they learn) may improve our ability to reproduce empirical data and, in so doing, improve their ability to explain neurobiological observations. Attempting to model large and often naturalistic data sets associated with data- driven computational psychiatry studies may lead to advances in ecological validity for theory-driven efforts and interpretability for data-driven ones.

It is our hope that this article, and the example above, sparks greater widespread motivation towards developing more accessible computational psychiatry measures, bringing them into the mainstream of psychiatric research where they are most likely to have a positive impact on patient care and preventative efforts. Indeed, beyond their use as research tools, it is our hope that something like the gameworld we have briefly illustrated here could eventually become commonplace in the assessment of, and screening for, psychiatric conditions. Furthermore, as has been demonstrated by recent approvals of digital therapeutics for psychiatric indications, these games may also have the potential to serve as accessible and personalizable vehicles for the delivery of treatments. Once computational tools have been successfully integrated into translational research and have hopefully produced novel screening and therapeutics tools, they will face a different set of barriers that will stand in the way of their seamless integration into standard clinical practice. These barriers will include regulatory considerations as well as barriers related to implementation into clinical workflows. With respect to regulation computational tools for direct implementation in the clinic will be required to meet standards for efficacy and safety relevant to medical devices [50]. Because implementation of computational tools has not yet been attempted at a large scale, the design of these studies will need to be rigorous and at times creative to manage concerns around patient, clinician, and researcher blinding as well as training users of the tool. In addition, due to the heterogeneity present in many psychiatric disorders, careful delineation of intended use populations as well as indications and use conditions for potential tools will be required to ensure tools function as intended when out of the research context. This will in turn require extensive feasibility and validation work prior to large-scale clinical studies. Once approval is granted, the clinical

utility of these tools will be determined in part by their impact on day-to-day clinical workflow. For example, if tools add time to this workflow that is not offset by utility for clinical decision-making at key junctures, they will not be adopted in a way that impacts patient care. Once tools are integrated into clinical workflows, continued monitoring of their validity in the intended use population as well as efforts to expand tools to other populations will be important. To aid this effort, the lessons learned from iterative improvement of research-based computational psychiatry tools should be applied to clinic-ready tools, allowing for progressive quality improvement and refinement of application to increasingly precise subsets of the real-world patient populations we treat. It is important to note that this iterative improvement will need to be conducted in concert with regulators, in order to ensure that changes to tools as a result of data collected after marketing continue to result in safe and effective tools for the given indication.

Lastly, overcoming these barriers will require sustained effort and resources from an already- overburdened population of academic psychiatrists. Marshaling these resources will require significant investment from both public and private sources so that incentives and expertise can align to transform these tools into market-ready, validated products. In many respects, the development of computational psychiatry tools is similar to and overlaps with the rich literature on the development and implementation of digital tools in clinical practice that is beginning to emerge [11–13, 51, 52]. This literature, which provides more detail on the barriers discussed here, as well as potential solutions, should serve as a guide for the translation of computationally-driven devices when they are ready.

REFERENCES

1. Redish AD, Gordon JA. Computational psychiatry: new perspectives on mental illness. 7 - How Could We Get Nosology from Computation? Cambridge, Massachusetts: The MIT Press; 2016.
2. Stephan KE, Mathys C. Computational approaches to psychiatry. Curr Opin Neurobiol. 2014;25:85–92.
3. Browning M, Carter CS, Chatham C, den Ouden H, Gillan CM, Baker JT, et al. Realizing the Clinical Potential of Computational Psychiatry: Report from the Banbury Center Meeting, February 2019. Biol Psychiatry 2020;88:e5–e10.
4. Peggy Series. Computational Psychiatry: A Primer. Cambridge, MA: MIT Press; 2020.
5. Bennett D, Silverstein SM, Niv Y. The two cultures of computational psychiatry. JAMA Psychiatry. 2019;76:563 https://doi.org/10.1001/jamapsychiatry.2019.0231.
6. Huys QJM, Maia TV, Frank MJ. Computational psychiatry as a bridge from neuroscience to clinical applications. Nat Neurosci. 2016;19:404–13. https://doi.org/10.1038/nn.4238.
7. Squarcina L, Villa FM, Nobile M, Grisan E, Brambilla P. Deep learning for the prediction of treatment response in depression. J Affect Disord. 2021;281:618–22. https://doi.org/10.1016/j.jad.2020.11.104.
8. Mehltretter J, Rollins C, Benrimoh D, Fratila R, Perlman K, Israel S, et al. Analysis of Features Selected by a Deep Learning Model for Differential Treatment Selection in Depression. Front Artif Intell. 2020;2. https://www.frontiersin.org/articles/10.3389/frai.2019.00031
9. Mehltretter J, Fratila R, Benrimoh D, Kapelner A, Perlman K, Snook E. et al. Differential Treatment Benefit Prediction for Treatment Selection in Depression: A Deep Learning Analysis of STAR*D and CO-MED Data. Comput Psychiatry. 2020;4:61–75. https://doi.org/10.1162/cpsy_a_00029.
10. Chekroud AM, Zotti RJ, Shehzad Z, Gueorguieva R, Johnson MK, Trivedi MH. et al. Cross-trial prediction of treatment outcome in depression: a machine learning approach. Lancet Psychiatry. 2016;3:243–50. https://doi.org/10.1016/S2215-0366(15)00471-X.
11. Tanguay-Sela M, Benrimoh D, Popescu C, Perez T, Rollins C, Snook E. et al. Evaluating the perceived utility of an artificial intelligence-powered clinical decision support system for depression treatment using a simulation center. Psychiatry Res. 2022;308:114336. https://doi.org/10.1016/j.psychres.2021.114336.
12. Popescu C, Golden G, Benrimoh D, Tanguay-Sela M, Slowey D, Lundrigan E. et al. Evaluating the Clinical Feasibility of an Artificial Intelligence-Powered, Web-Based Clinical Decision Support System for the Treatment of Depression in Adults: Longitudinal Feasibility Study. JMIR Form Res. 2021;5:e31862. https://doi.org/10.2196/31862.
13. Benrimoh D, Tanguay-Sela M, Perlman K, Israel S, Mehltretter J, Armstrong C. et al. Using a simulation centre to evaluate preliminary acceptability and impact of an artificial intelligence-powered clinical decision support system for depression treatment on the physician-patient interaction. BJPsych Open. 2021;7:e22 https://doi.org/10.1192/bjo.2020.127.
14. Benrimoh D, Israel S, Fratila R, Armstrong C, Perlman K, Rosenfeld A, et al. Editorial: ML and AI Safety, Effectiveness and Explainability in Healthcare. Front Big Data. 2021;4. https://www.frontiersin.org/articles/10.3389/fdata.2021.727856
15. Pavlovian conditioning–induced hallucinations result from overweighting of perceptual priors | Science. (n.d.). Retrieved September 27, 2022, from https://www.science-org.proxy3.library.mcgill.ca/doi/10.1126/science.aan3458
16. Schwartenbeck P, FitzGerald THB, Mathys C, Dolan R, Friston K. The dopaminergic midbrain encodes the expected certainty about desired outcomes. Cereb Cortex. 2015;25:3434–45.
17. Smith R, Kuplicki R, Feinstein J, Forthman KL, Stewart JL, Paulus MP, Khalsa SS.Investigators, T. 1000 A Bayesian computational model reveals a failure to adapt interoceptive precision estimates across depression, anxiety, eating, and substance use disorders. PLOS Comput Biol. 2020;16:e1008484. https://doi.org/10.1371/journal.pcbi.1008484.
18. Deserno L, Boehme R, Mathys C, Katthagen T, Kaminski J, Stephan KE, et al. Volatility estimates increase choice switching and relate to prefrontal activity in schizophrenia. Biol Psychiatry Cogn Neurosci Neuroimaging 2020;5:173–83. https://doi.org/10.1016/j.bpsc.2019.10.007.
19. Benrimoh DA, Parr T, Vincent P, Adams RA, Friston K. Active inference and auditory hallucinations. Comput Psychiatry. 2018;2:183–204. https://doi.org/10.1162/CPSY_a_00022.
20. Cassidy CM, Balsam PD, Weinstein JJ, Rosengard RJ, Slifstein M, Daw ND. et al. A Perceptual Inference Mechanism for Hallucinations Linked to Striatal Dopamine. Curr Biol. 2018;28:503–14.e4. https://doi.org/10.1016/j.cub.2017.12.059.
21. Benrimoh D, Parr T, Adams RA, Friston K. Hallucinations both in and out of context: An active inference account. PLoS ONE. 2019;14:e0212379. https://doi.org/10.1371/journal.pone.0212379.
22. Teufel C, Subramaniam N, Dobler V, Perez J, Finnemann J, Mehta PR, et al. Shift toward prior knowledge confers a perceptual advantage in early psychosis and psychosis-prone healthy individuals. Proc Natl Acad Sci USA 2015;112:13401–6. https://doi.org/10.1073/pnas.1503916112.
23. Fryer SL, Roach BJ, Hamilton HK, Bachman P, Belger A, Carrión RE. et al. Deficits in auditory predictive coding in individuals with the psychosis risk syndrome: prediction of conversion to psychosis. J Abnorm Psychol. 2020;129:599–611. https://doi-org.proxy3.library.mcgill.ca/10.1037/abn0000513.
24. Chen C, Takahashi T. Reward Processing in Depression. In Computational Models of Brain and Behavior. John Wiley & Sons, Ltd; 2017. p. 57–71. https://doi.org/10.1002/9781119159193.ch5
25. Huys QJ, Pizzagalli DA, Bogdan R, Dayan P. Mapping anhedonia onto reinforcement learning: A behavioural meta-analysis. Biol Mood Anxiety Disord. 2013;3:12. https://doi.org/10.1186/2045-5380-3-12.
26. Smith R, Taylor S, Wilson RC, Chuning AE, Persich MR, Wang S, et al. Lower levels of directed exploration and reflective thinking are associated with greater anxiety and depression. Front Psychiatry. 2022;12:782136.
27. Szaliznyó K, Silverstein DN. Computational Predictions for OCD Pathophysiology and Treatment: A Review. Front Psychiatry. 2021;12. https://www.frontiersin.org/articles/10.3389/fpsyt.2021.687062
28. Smith R, Taylor S, Stewart JL, Guinjoan SM, Ironside M, Kirlic N, Tulsa 1000 Investigators. Slower Learning Rates from Negative Outcomes in Substance Use Disorder over a 1-Year Period and Their Potential Predictive Utility. Comput Psychiatry. 2022;6:117–141.
29. Smith R, Schwartenbeck P, Stewart JL, Kuplicki R, Ekhtiari H, Paulus MP, Tulsa 1000 Investigators. Imprecise action selection in substance use disorder: evidence for active learning impairments when solving the explore-exploit dilemma. Drug Alcohol Depend. 2020;215:108208.
30. Smith R, Kirlic N, Stewart JL, Touthang J, Kuplicki R, Khalsa SS, et al. Greater decision uncertainty characterizes a transdiagnostic patient sample during approach-avoidance conflict: a computational modelling approach. J Psychiatry Neurosci. 2021;46:E74–E87.
31. Smith R, Kirlic N, Stewart JL, Touthang J, Kuplicki R, McDermott TJ, et al. Long-term stability of computational parameters during approach- avoidance conflict in a transdiagnostic psychiatric patient sample. Sci Rep. 2021;11:1–13.
32. Mathys CD, Lomakina EI, Daunizeau J, Iglesias S, Brodersen KH, Friston KJ, et al. Uncertainty in perception and the Hierarchical Gaussian Filter. Front Hum Neurosci. 2014;8. https://www.frontiersin.org/articles/10.3389/fnhum.2014.00825
33. PAR | MCCB | MATRICS Consensus Cognitive Battery. (n.d.). Retrieved September 27, 2022, from https://www.parinc.com/Products/Pkey/225
34. Schizophrenia. (n.d.). Retrieved September 27, 2022, from https://www.cambridgecognition.com/cantab/test-batteries/schizophrenia/

35. Neuropsychological and Psychological Testing—Medical Clinical Policy Bulletins | Aetna. (n.d.). Retrieved September 27, 2022. http://www.aetna.com/cpb/medical/data/100_199/0158.html

36. Kafadar E, Fisher VL, Quagan B, Hammer A, Jaeger H, Mourgues C, et al. Conditioned Hallucinations and Prior Overweighting Are State-Sensitive Markers of Hallucination Susceptibility. Biol Psychiatry. 2022. https://doi.org/10.1016/j.biopsych.2022.05.007

37. Grujic N, Brus J, Burdakov D, Polania R. Rational inattention in mice. Sci Adv. 2022;8:eabj8935. https://doi.org/10.1126/sciadv.abj8935.

38. Reed EJ, Uddenberg S, Suthaharan P, Mathys CD, Taylor JR, Groman SM. et al. Paranoia as a deficit in non-social belief updating. ELife. 2020;9:e56345. https://doi.org/10.7554/eLife.56345.

39. Bansal S, Bae G-Y, Robinson BM, Hahn B, Waltz J, Erickson M. et al. Association between failures in perceptual updating and the severity of psychosis in schizophrenia. JAMA Psychiatry. 2022;79:169–77. https://doi.org/10.1001/jamapsychiatry.2021.3482.

40. Dawson DR, Marcotte TD. Special issue on ecological validity and cognitive assessment. Neuropsychol Rehab. 2017;27:599–602. https://doi.org/10.1080/09602011.2017.1313379.

41. Chaytor N, Schmitter-Edgecombe M. The ecological validity of neuropsychological tests: a review of the literature on everyday cognitive skills. Neuropsychol Rev. 2004;13:181–97. https://doi.org/10.1023/B:NERV.0000009483.91468.fb.

42. Spooner DM, Pachana NA. Ecological validity in neuropsychological assessment: a case for greater consideration in research with neurologically intact populations. Arch Clin Neuropsychol 2006;21:327–37.

43. Haghani M, Bliemer MCJ, Rose JM, Oppewal H, Lancsar E. Hypothetical bias in stated choice experiments: Part II. Conceptualisation of external validity, sources and explanations of bias and effectiveness of mitigation methods. J Choice Model. 2021;41:100322. https://doi.org/10.1016/j.jocm.2021.100322.

44. Schmuckler MA. What is ecological validity? A dimensional analysis. Infancy. 2001;2:419–36. https://doi.org/10.1207/S15327078IN0204_02.

45. Cook J, Whittington D, Canh DG, Johnson FR, Nyamete A. Reliability of stated preferences for cholera and typhoid vaccines with time to think in Hue, Vietnam. Economic Inq. 2007;45:100–14. https://doi.org/10.1111/j.1465-7295.2006.00038.x.

46. Martland N, Martland R, Cullen AE, Bhattacharyya S. Are adult stressful life events associated with psychotic relapse? A systematic review of 23 studies. Psychol Med. 2020;50:2302–16. https://doi.org/10.1017/S0033291720003554.

47. Wang L-H, Chen B, Hwang G-J, Guan J-Q, Wang Y-Q. Effects of digital game-based STEM education on students' learning achievement: a meta-analysis. Int J STEM Educ. 2022;9:26. https://doi.org/10.1186/s40594-022-00344-0.

48. Huys QJM, Browning M, Paulus MP, Frank MJ. Advances in the computational understanding of mental illness. Neuropsychopharmacology. 2021;46:3–19. https://doi.org/10.1038/s41386-020-0746-4.

49. Smith R, Khalsa SS, Paulus MP. An active inference approach to dissecting reasons for nonadherence to antidepressants. Biol Psychiatry: Cogn Neurosci Neuroimaging. 2021;6:919–34. https://doi.org/10.1016/j.bpsc.2019.11.012.

50. Food and Drug Administration. Assessing the Credibility of Computational Modeling and Simulation in Medical Device Submissions. Draft Guidance for Industry and Food and Drug Administration Staff. Silver Spring, MD, USA: Food and Drug Administration; 2021.

51. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI (2020). An overview of clinical decision support systems: Benefits, risks, and strategies for success. Npj Digital Med. 2020;31. https://doi.org/10.1038/s41746-020-0221-y

52. Golden G, Popescu C, Israel S, Perlman K, Armstrong C, Fratila R, et al. Applying Artificial Intelligence to Clinical Decision Support in Mental Health: What Have We Learned? 2023. https://doi.org/10.48550/arXiv.2303.03511

AUTHOR CONTRIBUTIONS
Initial outline generated by ARP, who also edited the final manuscript. First draft generated by DB. Edits and intellectual contributions provided by VF, AS, CM, and RS.

COMPETING INTERESTS
The authors declare no competing interests.

ADDITIONAL INFORMATION
Correspondence and requests for materials should be addressed to Albert R. Powers.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.