MULTIVARIATE BEHAVIORAL RESEARCH
2023, VOL. 58, NO. 4, 743–761

From Metaphor to Computation: Constructing the Potential Landscape for Multivariate Psychological Formal Models

Jingmeng Cui[a,b], Anna Lichtwarck-Aschoff[a,b], Merlijn Olthof[b], Tiejun Li[c], and Fred Hasselman[b]

[a]Faculty of Behavioural and Social Sciences, University of Groningen; [b]Behavioural Science Institute, Radboud University; [c]LMAM and School of Mathematical Sciences, Peking University

ABSTRACT

For psychological formal models, the stability of different phases is an important property for understanding individual differences and change processes. Many researchers use landscapes as a metaphor to illustrate the concept of stability, but so far there is no method to quantify the stability of a system's phases. We here propose a method to construct the potential landscape for multivariate psychological models. This method is based on the generalized potential function defined by Wang et al. (2008) and Monte Carlo simulation. Based on potential landscapes we define three different types of stability for psychological phases: absolute stability, relative stability, and geometric stability. The panic disorder model by Robinaugh et al. (2019) is used as an example, to demonstrate how the method can be used to quantify the stability of states and phases, illustrate the influence of model parameters, and guide model modifications. An R package, simlandr, was developed to provide an implementation of the method.

KEYWORDS

Complex dynamical systems; clinical psychology; formal theory; computational modeling; potential landscape

Introduction

The past decades have seen a rapid growth of models and theories in the field of psychology, and more specifically in the field of psychopathology. These models, however, are not without critiques. Many *verbal* theories acknowledge the complex and dynamic nature of mental disorders, but they are not always able to make precise and falsifiable predictions (Borsboom et al., 2021; Robinaugh et al., 2021). Statistical models can provide quantitative estimations, but traditional and even cutting-edge statistical methods are largely based on linear, static, and homogeneous assumptions, and usually fail to draw correct conclusions about the nature of the underlying process (Granic & Hollenstein, 2003; Haslbeck et al., 2021; Olthof, Hasselman, & Lichtwarck-Aschoff, 2020). In recent years, *formal* models are gaining momentum in the field of psychology, aimed to address the abovementioned problems and provide a quantitative foundation for theoretical inferences (e.g., Burger et al., 2020; Cramer et al., 2016; Robinaugh et al., 2019; Schiepek et al., 2014).[1]

In those formal models, the elements of psychological systems and their interactions are described mathematically. Based on these specifications, one can simulate how the model evolves over time, observe the characteristics of the model, and investigate how the model output corresponds to real-life phenomena. In complex systems, higher-order, macroscopic *psychological phases*[2]—distinct patterns of psychological systems—can *emerge* from the *self-organization* of these microscopic elements and their interactions (Goldstein, 1999; Olthof et al., in press). Different psychological phases can sometimes be assigned to differences in mental health: a system can for instance be

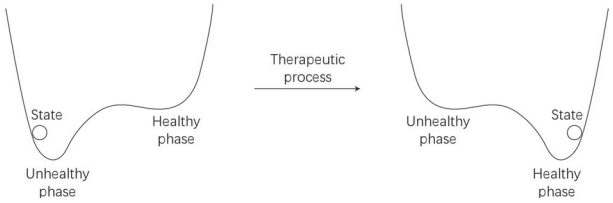

Figure 1. Diagram of concepts in the landscape metaphor of psychopathology.

in an anxious, panicky phase or in a calm and relaxed phase (Robinaugh et al., 2019).

For formal models, the stability of different states and phases is an important quantity that relates to individual differences and change processes of the system. Previous research has often used the landscape metaphor to illustrate this idea: the state of the system is like a ball on the landscape. If the ball is in "a deep valley," the system is stable; if the ball is on "a hill" or in "a shallow valley," the system is unstable, and it tends to "fall down" to a more stable place. The "valleys" or "attractors" correspond to possible phases of the system (Lamothe et al., 2019; Olthof, Hasselman, Oude Maatman, et al., 2020; Wichers et al., 2019; see Figure 1 for a diagram). If the valley of an unhealthy phase of some individuals is deeper, their mental systems are more likely to be trapped there, and they will be more vulnerable to mental disorders. Also, the process of treating mental disorders can be seen as changing the landscape of the system in a way that the stability of the unhealthy phase is decreased while at the same time strengthening the stability of healthy phases (Hayes et al., 2015).

While the metaphorical use of the landscape is certainly a good way of illustration, the concept of stability does not yet have a formal, quantitative representation, which hinders further investigation. This calls for a new line of methods: quantitatively computing the potential landscape from formal dynamic models.

Formal models and case simulations

As its name suggests, a psychological formal model of psychological phenomena mathematically defines how *variables* evolve over time and how they interact with each other. Often the evolution of such systems can be described by a set of *(stochastic) differential equations*. These equations specify the forms and strengths of these interactions among variables and the magnitude of noise in mathematical form. With this precise description, one can determine how the system evolves over time from a starting point and gain knowledge about the theory-implied behavior of the system (Robinaugh et al., 2021).

A well-known psychological formal model is the panic disorder model by Robinaugh et al. (2019; also see Borsboom et al., 2021, and Haslbeck et al., 2021, for discussions on this model). This model is well constructed and contains many typical features that are common in psychological models (e.g., nonlinear relationships, feedback loops, the dependence of system behavior on its unique history, adaption to the environment, and a rather large number of variables). Therefore, we use this model as an example to explain our ideas. Here we briefly introduce the model specification and the main variables and parameters. The relationships of the most important variables that we use in the current paper are shown in Figure 2, and the full description of the model can be found in Robinaugh et al. (2019). The variables and parameters that will be investigated in the current paper are marked in bold font on their first occurrence.

In this model, a panic disorder is considered to emerge from mutually interacting system variables. The core variables of this system are physical arousal (***A***, the level of arousal-related sensations, e.g., heart rate) and perceived threat (***PT***, the cognitive perception that the situation is threatening). The changing rate of physical arousal (A), $\mathrm{d}A/\mathrm{d}t$, is influenced by its own value, perceived threat (PT), and homeostatic feedback (H, the strength of the homeostatic processes that counteract the unsustainably elevated physical arousal),

$$\frac{\mathrm{d}A}{\mathrm{d}t} = r_A(s_{PT,A}PT - A - s_{H,A}H), \tag{1}$$

which represents that physical arousal (A) tends to decrease when itself and homeostatic feedback (H) is high and tends to increase when perceived threat (PT) is high. The parameters r_A, $s_{PT,A}$, and $s_{H,A}$ represent the strength of these influences. The changing rate of homeostatic feedback (H), $\mathrm{d}H/\mathrm{d}t$, is influenced by its own value and physical arousal (A),

$$\frac{\mathrm{d}H}{\mathrm{d}t} = r_H\left(\frac{A^{p_{A,H}}}{A^{p_{A,H}} + h_{A,H}^{p_{A,H}}} - H\right), \tag{2}$$

which represents that homeostatic feedback (H) tends to decrease when itself is high and tends to increase when physical arousal (A) is high. The parameters r_H, $p_{A,H}$, and $\boldsymbol{h_{A,H}}$ represent the strength of these influences. The changing rate of perceived threat (PT), $\mathrm{d}PT/\mathrm{d}t$, is influenced by its own value, physical arousal (A), and escape behavior (E),

$$\frac{\mathrm{d}PT}{\mathrm{d}t} = r_{PT}\left(\frac{A^{p_{A,PT}}}{A^{p_{A,PT}} + h_{A,PT}^{p_{A,PT}}} - PT - s_{E,PT}E\right), \tag{3}$$

which represents that perceived threat (PT) tends to decrease when itself and escape behavior (***E***) is high

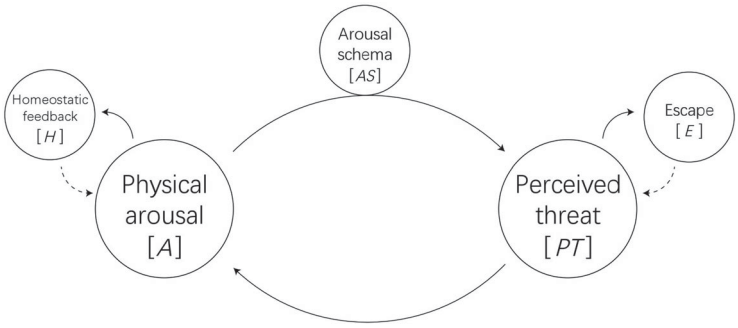

Figure 2. Simplified causal diagram of the panic disorder model (adapted from Robinaugh et al., 2019). Each circle represents a variable of the model. The solid lines represent positive influences on the changing rate of the target variable, and the dashed lines represent negative influences on the changing rate of the target variable. The circle on the path from arousal to perceived threat represents the moderating effect of arousal schema on this relationship.

and tends to increase when physical arousal (A) is high. r_{PT}, $p_{A,PT}$, $h_{A,PT}$, $s_{E,PT}$ are parameters representing the strength of these influences. Here, the influence of A on the changing rate of PT is not linear, but in the form of an S-shaped sigmoid function. Robinaugh et al. (2021) showed that this sigmoid function is necessary for the formal model in order to generate theory-implied behavior as observed in real life. The changing rate of escape behavior (E), $\mathrm{d}E/\mathrm{d}t$, is influenced by its own value and perceived threat,

$$\frac{\mathrm{d}E}{\mathrm{d}t} = r_E\left(\frac{PT^{p_{PT,E}}}{PT^{p_{PT,E}} + h_{PT,E}^{p_{PT,E}}} - E\right), \tag{4}$$

which represents that escape behavior (E) tends to decrease when it is high itself and tends to increase when perceived threat (PT) is high.

People with a panic disorder over-interpret their physical arousal as an indication of danger. When this over-interpretation, termed as *arousal schema* (***AS***) in the model, is high, an increase in physical arousal can result in a larger increase in perceived threat. This effect of AS is represented as its influence on the parameter $h_{A,PT}$,

$$h_{A,PT} = 1 - \frac{AS}{AS + h_{AS,APT}} - s_{sit,APT}c. \tag{5}$$

When AS is higher, $h_{A,PT}$ is lower, which makes the influence of A more dominant in Equation 3. Therefore, it is easier for a relatively high level of physical arousal (A) to lead to a large increase in perceived threat (PT). The increase of perceived threat (PT) can, in turn, amplify physical arousal (A). This generates a vicious circle between the two and finally lead to a panic attack. A collective variable ***fear*** is defined as the geometric mean of A and PT,

$$fear = \sqrt{A \times PT}, \tag{6}$$

and is used to represent the general symptom severity. Panic attacks, therefore, manifest as a sudden increase in the level of *fear*. If panic attacks happen relatively often, the person can be said to have a panic disorder. Arousal schema (AS) is also influenced by a learning mechanism. Its changing rate, $\mathrm{d}AS/\mathrm{d}t$, depends on its own value and the previous history of perceived threat (PT), *fear*, and escape behavior (E),

$$\frac{\mathrm{d}AS}{\mathrm{d}t} = \begin{cases} 0, & if\ \ max\left(fear_{t-\Omega}, \ldots, fear_t\right) < cr_{fear,AS} \\ r_{AS,a}(\max(PT_{t-\Omega}, \ldots, PT_t) - AS), & if\ \ max\left(fear_{t-\Omega}, \ldots, fear_t\right) \geq cr_{fear,AS}\ \ and\ \ max(E_{t-\Omega}, \ldots, E_t) > cr_{E,AS} \\ -r_{AS,e}AS, & if\ \ max\left(fear_{t-\Omega}, \ldots, fear_t\right) \geq cr_{fear,AS}\ \ and\ \ max(E_{t-\Omega}, \ldots, E_t) \leq cr_{E,AS} \end{cases}, \tag{7}$$

which represents three different learning conditions. When the maximum value of *fear* in the previous Ω time points is lower than a critical threshold $cr_{fear,AS}$, no learning processes happened; when the maximum value of *fear* in the previous Ω time points is higher than the critical threshold $cr_{fear,AS}$, the direction of the learning process depends on whether the individual's escape behavior (E) in the previous Ω time points is higher than another threshold $cr_{E,AS}$. If the previous escape behavior (E) is high, the individual does not know how threatening the actual situation is, so he or

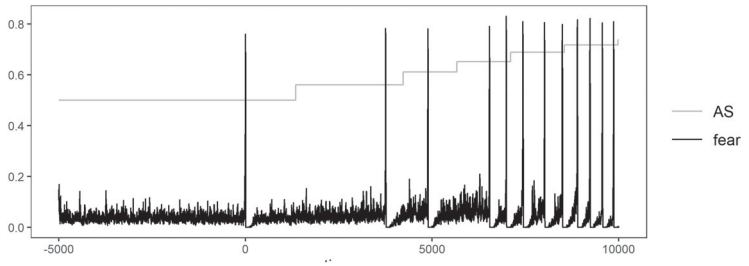

Figure 3. The simulation results using the panic disorder model by Robinaugh et al. (2019). The first panic attack of the system appears probabilistically. To make simulation results comparable, *time* was set as zero at the first panic attack (the time when the peak value of *fear* is reached).

she will learn to update the arousal schema (AS) according to the highest perceived threat (PT) during this time period with an acquiring rate parameter $r_{AS,a}$; if the previous escape behavior (E) is low, the individual will find that the actual situation is not so threatening, so his or her arousal schema (AS) will decrease with an extinguishing rate parameter $r_{AS,e}$.

The model is intended to produce two qualitatively different phases of the (patient) system: a healthy phase and a panic phase. The simulation results of this model (with the parameter values in Robinaugh et al., 2019; Figure 3) indeed show clear sudden increases in *fear*, which represent panic attacks. In these panic attacks, the system moves from the healthy phase to the panic phase, and quickly transitions back to the healthy phase. When the arousal schema (AS) becomes higher, the panic attacks become more frequent, which qualitatively shows the stability of the panic phase increasing, and the stability of the healthy phase diminishing. These simulation results align well with the theoretical foundation of the model.

In this approach, the performance of the model is evaluated by case simulation result: the output of directly simulating the dynamic model. Case simulation is an important way of model evaluation and deduction because it shows how the system evolves if the model correctly represents the system. A mismatch of the simulation result and real-life observations indicates something must be wrong in the model. Case simulation, however, cannot directly provide information about the stability of states: what it shows is how the state changes over time (i.e., state as a function of time), not the stability of different states (i.e., stability as a function of states). Therefore, we need to find a way to define and calculate the stability of the states in a psychological system.

Potential landscape for a dynamical system

Defining or representing stability is not a totally new subject. In physics, the quantity that is used to represent the stability of a state is the *potential function*. Take the gravitational potential energy as an example: for a given object, if it is at a higher position, its gravitational potential energy is higher. This means that the object is more unstable, and it is more likely to "fall down" to more stable states in lower places. Mathematically, if a potential function can be (strictly) defined for a system, its velocity (i.e., how the system state changes over time) should be proportional to the gradient of the potential function. Intuitively, this means that the system always tends to move to the place with lower potential energy, just like a ball on a hill tends to fall along the most convenient way into the valley.

For a unidimensional deterministic system, the potential function can be easily obtained from taking the integral of the dynamic function. This method has already been used for representing the stability of states in unidimensional psychological systems (e.g., Dablander et al., 2022; Robinaugh et al., 2019). However, most psychological models contain many variables, thus are multidimensional. Analyzing the stability for multidimensional systems is more challenging because the multivariate dynamic functions are often not integrable. One way to understand this issue is by looking at the Penrose impossible stairs (Rodríguez-Sánchez et al., 2020; see Figure 4): it is possible that a system keeps whirling around, but then it cannot be represented as always going downstairs in a real 3D space. Therefore, the potential landscape cannot be directly obtained. Mathematicians have developed several generalized methods to construct generalized potential functions (P. Zhou & Li, 2016). These methods relax the requirements of integrability, but the resulting functions can also be used to represent the stability of system states. Among different ways of generalization, we found the potential function by Wang et al. (2008; also see Li & Wang, 2013, and Li & Ye, 2019, for examples of usage in complex biochemical systems) the most suitable for psychological formal models because it can be estimated with

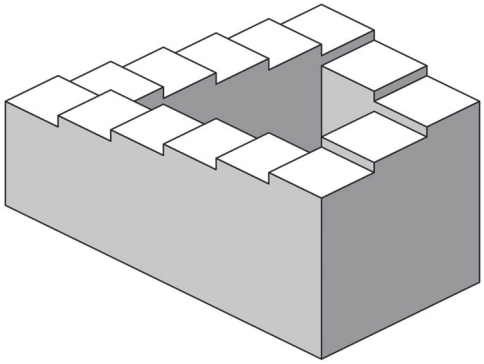

Figure 4. The Penrose impossible stairs (Sakurambo, 2005).

the Monte Carlo method, thus does not have a strict requirement concerning the properties of the dynamic functions. Other generalizations often require the dynamic functions to be continuous, derivable, and independent of the history, properties that psychological models often do not meet (e.g., Equation 7 of the panic disorder model is dependent on the model history and thus not derivable). Wang's definition of potential function is based on the steady-state distribution of the system (denoted as P_{SS}), which refers to the distribution of states that holds constant over time. If we have a single system evolving over time according to a set of (stochastic) differential equations, its state is likely to change every now and then. However, if we have an infinite bunch of systems with the steady-state distribution and let them all evolve together, although each system's state still changes, their distribution can be invariant.[3] Following a generalization of Boltzmann distribution, the potential (U) of a state X is then given by

$$U(\boldsymbol{X}) = -\ln\ P_{SS}(\boldsymbol{X}) \tag{8}$$

which means that the potential is equal to the logarithm of the steady-state distribution. If the probability density for the steady-state distribution is lower, then its potential is higher.

This potential function is also related to some other properties expected from strictly defined potential functions. Here we also explain the usefulness of generalized potential functions from the force decomposition perspective (Zhou & Li, 2016). As mentioned above, the main issue in constructing potential landscapes for multidimensional systems is that they often show whirling behaviors that cannot be represented with a potential landscape. These whirling behaviors are not the whole picture because the total forces in the system also contain gradient parts that represent the general tendency for the system to move to some specific regions. The idea of force decomposition means to decompose those complex forces of a system into a curl part and a gradient part. The stability information of the system is mainly contained in the gradient part, and the gradient part is integrable. Therefore, the generalized potential landscape can be obtained by integrating the gradient part of the forces. The curl part does not contain direct information about the stability because if the system just has the tendency to oscillate between two states without a preference, it is not meaningful to say that one state is more stable than the other. Wang's landscape is originally defined from the steady-state distribution, but it can also be proved that it is equivalent to a possible way of force decomposition (Zhou & Li, 2016). Therefore, the potential landscape constructed with Wang's method can be seen as a representation of the gradient part of the system dynamics.

From either perspective, the potential landscape of the system shows the tendency that the system resides or leaves a specific state. If a state has a lower density in the steady-state distribution, the system is less likely to be around this state in the long run; and if there are a bunch of systems starting with a uniform distribution in the state space, it is more likely that the systems starting around this state will move to other states. If a system is in a higher position on the gradient part of its dynamics, it means that the system tends to fall down to a lower position if not affected by the curl forces and random noises. Therefore, the potential landscape can efficiently represent the stability of psychological systems on the state level.

Having the potential for specific states, we can now describe the stability of the phases. How to clearly define psychological phases is a complex issue on its own. One may propose that a collection of qualitatively similar states constitutes a phase (e.g., the mental states when a patient with depression has a high overall symptom severity), in which case the phase may be seen as a *point attractor*. In contrast, it can also be the case that certain kind of trajectories constitutes a phase (e.g., a mental trajectory in which a patient with cyclothymic disorder switches between states with high depressive symptoms and states with

[3]A concrete illustration: imagine an infinite number of copies of an athlete running on a 400-meter track, and imagine they are independent (so they do not run into each other). Even if they are copies of the same athletes, due to random noises, they will gradually distance themselves from each other. After an infinitely long time, these people will be very evenly distributed in the track. Now even if each person is still running, the population distribution on the track will not change anymore. If each running person is a stochastic dynamic system, then this population distribution is the steady-state distribution for the system.

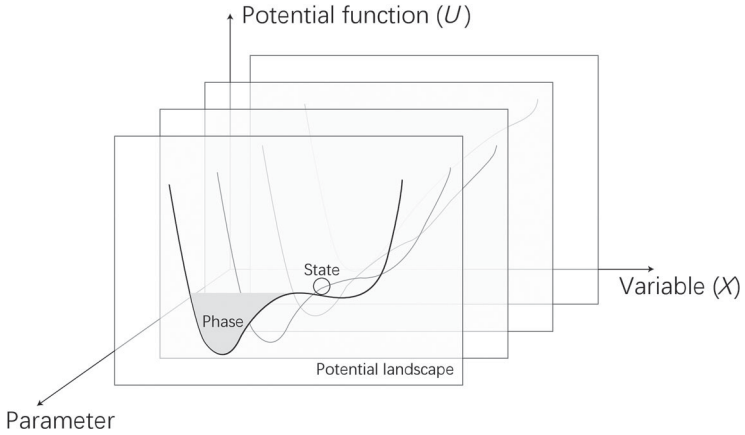

Figure 5. The relationships among the potential function, parameters, variables, state, and phase in the framework of our method.

elevated moods), in which case the phase may be seen as an *oscillating attractor* (Barton, 1994). Sometimes those two definitions can also be interchangeable. For example, if we use the mood variability over a period instead of the valence and intensity of mood as the key variable to describe the cyclothymic disorder, then the patient's mood variability is always high, and may be described as a point attractor. Fully investigating this issue is beyond the scope of the current paper. For clarity and simplicity, we will only look into the phases consisting of similar states and bounded by barriers in the potential landscape (i.e., point attractor-like phases). In other words, we use the term *state* for a *point* in the state space (e.g., the state with $A = 0.1$ and $PT = 0.1$) and the term *phase* for a larger *region* that contains many states within it (e.g., the healthy phase). This distinction, although not commonly made in psychological literature, is important for the method we are introducing. The relationships among the potential function, parameters, variables, state, and phase we used in the current paper are shown in Figure 5.

Based on the potential function of states (defined in Equation 8), we can go one step further to define the stability of phases, which is often more important for psychological systems. From the potential landscape perspective, the stability of phases can be characterized in three ways (see Figure 6). The first one is the local minimum of the potential function within a given phase. Although the phase contains a collection of many states, the local minimum is the most stable state in the phase, hence can provide a quantitative representation. We refer to the potential of the local minimum as the *absolute stability* of the phase. Second, the potential difference between the local minimum and the barrier of the phase is directly related to the difficulty for the system to move out of that phase. We refer to it as the *relative stability*[4] of the phase. The difference between the two is that the former represents if given infinite time, how probable it is for the system to be in a given phase (e.g., the healthy phase in the panic disorder model); the latter represents, if putting the system in a given phase, how difficult it is for the system to escape that phase. Finally, the shape of the potential landscape within and around a certain phase represents how probable the system vibrates within or leaves a phase in a certain direction. We refer to this as the *geometric stability*. Although the geometric stability, as the shape of the landscape, is described qualitatively in the current work, we should note that this qualitative information is based on the quantitative information of the stability of the states within and around a phase of the system.

Besides quantifying the stability, potential landscapes can also be used for other purposes. One advantage of the potential landscape over the case simulation is that it can summarize the stability information concisely and be directly compared across different parameter settings, therefore enables systematic investigation of the influence of various parameters on the model. For example, Robinaugh et al. (2019) only provided one set of values for all 22 parameters in the model, and these parameter values were chosen based on their ability to produce reasonable output. The plausible range of each parameter, however, was not investigated. This is a common practice for performing case simulations for formal models but leaves the robustness of a specific parameter setting questionable. Those parameters also are related to trait-like psychological properties of the system, which are of theoretical interest. It is understandable that most modelers do not show the simulation results with all

[4]This is related to the term "resilience" used in some papers (e.g., Dablander et al., 2022).

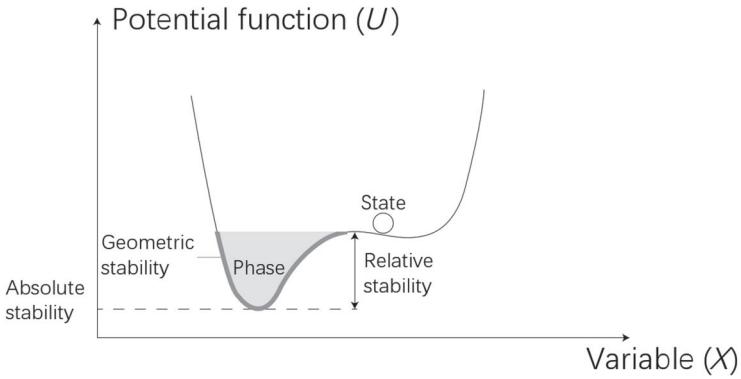

Figure 6. The relationships among the absolute stability, relative stability, and geometric stability of a psychological phase.

different parameter settings because it does not show the influence of parameters in an informative way. With the help of potential landscapes, these problems can be addressed more clearly.

The potential landscape can also provide guidance on model modification. It is nearly impossible to have a model that successfully explains every real-life phenomenon. Modification is often needed to continuously improve the model. Case simulations do not always provide enough guidance on how to modify the model. For example, there is a clinical phenomenon that the panic attack model could not explain: some people only have some (non-clinical) panic attacks but do not develop a panic disorder (Robinaugh et al., 2019). In the original model, however, if the panic attack happens once, the system would always develop into a panic disorder. The authors provided a way of mending as an example of model improvement, namely adding another parameter called "escape schema" (S_E). When S_E is higher, the parameter $h_{PT,E}$ in Equation 4 is higher. This parameter represents the extent to which the individual believes escaping could help to cope with the perceived threat. After adding this parameter, individuals with a low S_E would have panic attacks but would not develop a panic disorder. The rationale behind this solution is mainly based on a known theoretical mechanism that is translated into a model parameter, but not based on the model output. In other words, the simulation results only provide information about *whether* there is a problem, but not *how* the problem arises. The potential landscape method in contrast can provide insight into the problem—at least problems can be systematically analyzed from the perspective of stability.

Aim of the current research

The current research aims to provide a method to compute the potential landscape for psychological formal models and examine its usefulness in understanding psychological systems. First, we develop a set of tools to compute the potential landscape and related stability indicators from psychological dynamic models. Then, we illustrate how the procedure works by using the panic disorder model by Robinaugh et al. (2019). We will show how to (1) analyze the stability of states and phases from the potential landscape perspective, (2) systematically investigate the influence of various parameters from the potential landscape, and (3) use the potential landscape to guide model modification.

Methods

The method section is divided into three parts. In the first part, we show the adjustments we made to make the panic disorder model more suitable for landscape construction. In the second part, we explain the preparatory analyses we did that ensure the validity of the landscape results. In the third part, we introduce the method that we used in the main analysis. The simulations and analyses in this study were performed in R 4.1.2 (R Core Team, 2021). The replicable R scripts used for this study, as well as the animation or interactive version of the landscapes, can be found at the OSF repository of this project (https://osf.io/ke3xb/). An R package, *simlandr*,[5] was developed to organize the methods we used and to facilitate future applications. We try to involve minimal yet sufficient mathematics in the current article. For readers seeking more rigorous technical details of this method, we refer them to *S3. Practical information on programming issues* in the Supplementary Materials and Cui et al. (2021).

Model adjustments

In order to construct the landscape for the system, some modifications are needed. First, we adjusted

[5]The package is available at https://cran.r-project.org/package=simlandr.

some ways of computational implementations to make the simulation more effective. The new implementation produces, in principle,[6] the same outputs as the implementation by Robinaugh et al. (2019). We refer to this model as the *original model*.

Then, as we use long-term simulations to estimate the steady-state distribution of the model, we need the model to have global ergodicity. In other words, the model should travel through its entire realistic phase space in a sufficiently long time. Complex systems often display local ergodicity breaking, which occurs when the system gets trapped in a local minimum, or, attractor state. To ensure global ergodicity, we want the model to have large enough noise terms so that the system can escape local minima. The original model does not meet this requirement because there is only one noise term on the changing rate of physical arousal (A). This term is not enough to ensure the ergodicity of the whole system. Therefore, we added several Gaussian noise terms on all model variables to replace the single noise term in the original model. We also calibrated the standard deviation of the noise term to make sure the panic phase still exists. Second, in the original model, there is a short-long term difference: some variables (e.g., arousal schema, AS) are updated every "day," while other variables are updated per "minute." However, in a Monte Carlo simulation, the purpose is to estimate the steady-state distribution, not to represent the actual time scale of real-life changes. Keeping the short-long term difference in the model for Monte Carlo simulation is not meaningful and will reduce the effective sample size for those slowly updated variables. Therefore, we deleted the short-long term difference by updating all the variables in the same frequency. We refer to this model as the *simplified model*.

In the original model, AS is an important variable that influences the stability of the system, and it changes much more slowly compared to other variables. In the simplified model, we added noise to it and made it change faster. These modifications make it difficult to see how the value of AS changes the stability of the system. Therefore, we also made a model that is based on the simplified model but holding AS constant. In this model, AS is a parameter instead of a variable, and it is useful for investigating the influence of AS. We refer to it as the *constant* AS *model*.

Finally, we also extracted the deterministic part of the constant AS model (i.e., all noise terms were deleted). This is because a model of deterministic ordinary differential equations (ODEs) enables mathematical analyses of the stable points of the model. We refer to this model as the *deterministic model*.

Preparatory analyses

The number and stability of equilibrium points

Using the deterministic model, it is possible to analytically tell how many equilibrium states there are in the system. If the system is in its equilibrium points, all the time derivatives of the models should be zero. Moreover, if the eigenvalues of the Jacobian matrix of the system all have negative real parts, the equilibrium point is stable; if some of the eigenvalues have positive real parts, the equilibrium point is unstable (see Sayama, 2015, for explanations of this method). Therefore, the stabilities of the equilibrium points are determined by the *dominant eigenvalue* (λ_d), which means the eigenvalue with the largest real part. To know how many equilibrium points there are in the system, we hold every time derivative zero except for dPT/dt, and then calculate how it changes over PT. The result is shown in Figure S1 of the Supplementary Materials. Based on these results, we can find that when AS is low, there is only one zero point at $PT = 0$. However, as AS increases, two new equilibrium points appear.

The real part of these λ_ds is shown in Table S1 (Supplementary material). In the one-equilibrium-point cases, that point is stable; in the three-equilibrium-point cases, the first and the third points are stable, but the second one is unstable. The first and the third point here correspond to the healthy and the panic phase, respectively. The third equilibrium point only emerges when AS is sufficiently large. The second (unstable) equilibrium point corresponds to the saddle point of the system. Ideally, the system can be in equilibrium there, but with a small disturbance, the system will move to either of those two stable phases. Based on the results of stability analysis, we can confirm the heuristic that the landscape we construct should have one or two phases, depending on parameter values.

Checking convergence and determining the simulation length

Monte Carlo estimation of the steady-state distribution is only valid if the simulation *converges*, which means

[6]The simulation function is implemented using Rcpp (Eddelbuettel & François, 2011) instead of the implementation in R by Robinaugh et al. (2019), and random numbers are generated using RcppZiggurat (Eddelbuettel, 2020) in this implementation. Some differences in pseudo-random number generation may lead to very minor differences in the model output. These differences, however, do not influence the results meaningfully.

that there are enough data points sampled that the joint distribution of the variables will not change even when the simulation length is extended. We checked this by comparing the distribution of key variables in the initial, middle, and final stages of the simulation. With a simulation length of 10^7 timesteps, the distributions in different stages are sufficiently stable (Supplementary materials, Figure S2). Therefore, we use 10^7 as the simulation length for constructing landscapes.

Main analysis

The stability of states and phases

For constructing the potential landscape function for each possible state, we first estimate the steady-state distribution (P_{SS}) of the model with Monte Carlo simulation. The raw potential landscape function is defined in a high dimensional space, where the dimension equals the number of variables in the model. To make this function understandable, we need to perform a dimensionality reduction. This was done with a simple but widely used approach, which is obtaining the marginal distributions (e.g., Li & Wang, 2013; Zhang et al., 2020). This method is capable of visualizing up to three selected variables for the model. Kernel smooth methods were used to calculate smooth distribution density with a reduced set of variables each time, and Equation 8 was used to calculate the potential landscape.

For the stability of phases, we first calculated the absolute stability by finding the local minima of the potential function within each phase. Then, we looked for the *minimum energy path* (MEP) and the *saddle point* between the two phases. The minimal energy path is the path that the system would be most likely to travel from one local minimum to another if the system was purely gradient. The point with the highest potential in the MEP is the saddle point. It can be proved that this path should first go along the steepest ascending path from the starting point and then go along the steepest descending path to the end point (E & Vanden-Eijnden, 2010). Its geometric form, from which a Dijkstra algorithm (Dijkstra, 1959) can be derived, was used to find the MEP (Heymann & Vanden-Eijnden, 2008). After that, the relative stability defined by the barrier height was calculated as the potential difference between the saddle point and the local minima. The geometric stability of the phases is described qualitatively.

The effect of parameters

To investigate the influence of parameters on the stability of the phases, multiple simulations with different parameter values were performed. For a single parameter, its value was sampled evenly within a parameter space. The range of this parameter space is roughly centered around the original parameter values used by Robinaugh et al. (2019) with a plausible width that is large enough to clearly show the influence of the parameter on the potential landscape. The potential landscape was constructed separately for each parameter value, and the barrier height was computed respectively. Later, the barrier heights were compared across parameter values to show the influence of parameters on the difficulty for the system to escape a certain phase and transition into another phase. For the joint influence of two parameters, a sample grid was made for the combination of parameter values, and the potential landscape was calculated for each condition.

Model modification

Based on the information provided by the potential landscape, we propose the following general strategy for model modification. In the first step, the problem of the model is identified. It is usually some differences in stability between model outcome and real-life phenomena. In the second step, the reason for this inconsistency is analyzed from the potential landscape perspective. It can be that the stabilities of different phases are not suitable, the barrier height between phases is too high or too low, or the landscape has more or fewer phases than it should have. After that, the model is adjusted accordingly. For example, adding or removing time derivative terms can tilt the landscape and stabilize the states in a certain direction. Finally, both case simulation outputs and the potential landscape of the modified model are checked to test if the problem has been solved. Using this strategy, we analyzed a problem of the panic disorder model, provided a way of modification, and evaluated the modified model (Figure 6).

Results

Stability of states and phases

Among all the model variables, physical arousal (*A*), perceived threat (*PT*), and *fear* are the core variables representing the symptoms of panic disorder. A higher value of these variables represents higher symptom severity. Besides that, arousal schema (*AS*) represents the key control variable for the disorder. Therefore, we first constructed potential landscapes for these variables with the simplified model. In the potential landscape of *A* and *PT* (Figure 7a), we can find two local minima. The position of the first one is

at $A = 0.01$, $PT = 0.00$ ($U = -4.47$), and the second one is at $A = 0.63$, $PT = 0.73$ ($U = 3.01$). The saddle point is at $A = 0.38$, $PT = 0.41$ ($U = 3.99$). The barrier heights are $\Delta U = 8.46$ and $\Delta U = 0.98$ for the two phases, respectively. The first phase has a lower symptom severity, hence corresponds to the healthy phase of the system. The symptom severity of the second one is higher, hence corresponds to the panic phase of the system. The potential of the local minimum within the healthy phase is lower than that of the panic phase, and the barrier height of the healthy phase is higher than that of the panic phase, indicating that the healthy phase has a higher absolute and relative stability. Both phases show a regular circle-like shape, which means that the system state tends to vibrate around the local minimum symmetrically in both *A* and *PT* directions. There is a single pathway connecting them, which is the path that the system is likely to take when transitioning from one to the other. For example, if the system is going from the healthy phase to the panic phase, *A* and *PT* will increase together until the system reaches the region of the panic phase.

The landscape of *AS* and *fear* is shown in Figure 7b. Similarly, the phase with a lower *fear* level is the healthy phase, and the phase with a higher *fear* level is the panic phase. According to the landscape, the panic phase only appears when *AS* is high enough. The local minimum with lower *fear* is at $AS = 0.77$, $fear = 0.01$ ($U = -3.08$), and the local minimum with higher *fear* is at $AS = 0.78$, $fear = 0.68$ ($U = 2.83$). The saddle point is at $AS = 0.78$, $fear = 0.41$ ($U = 3.54$). The barrier heights are $\Delta U = 6.62$ and $\Delta U = 0.71$ for the two phases, respectively. Again, both absolute and relative stability indices support that the healthy phase is more stable than the panic phase. While the panic phase shows a circle-like shape, the healthy phase shows a slender shape, indicating that the potential landscape is rather flat along the *AS* direction. This shows that when *AS* is lower than the local minimum point (0.78), the higher and lower *AS* ranges do not differ much in stability, which means there is little resistance for *AS* to increase or decrease within the healthy phase. However, when *AS* is higher than the local minimum point (0.78), there is a strong tendency for *AS* to decrease. This indicates that there are some mechanisms in the system preventing *AS* from rising too much.

Besides the most central variables, the panic disorder model also has many other elements. Here, we calculated the landscape of escape (*E*) and *fear* to show the stability-related properties of *E*, which is an important behavioral mechanism in panic disorder. This potential landscape is shown in Figure 7c. As the previous potential landscapes, this potential landscape also shows two phases, with the healthy one being more stable than the panic phase (local minimum in the healthy phase: $E = 0.00$, $fear = 0.00$, $U = -4.97$; local minimum in the panic phase: $E = 0.95$, $fear = 0.66$, $U = 3.10$; saddle point: $E = 0.34$, $fear = 0.55$, $U = 5.65$; barrier heights: $\Delta U = 10.63$ and $\Delta U = 2.55$). *E* is in a lower range when the system is in the healthy phase and in a higher range when the system is in the panic phase, which indicates that the panic phase is related to a higher tendency of escaping.

While the potential landscape (Figure 7c) around two local minima shows a regular shape, the landscape between them shows a unique phenomenon: between the healthy phase and the panic phase, there are two pathways instead of one. Between these two paths, there is also a small "hill" inside the "valley." This indicates that there are two possible paths of transitioning from one phase to the other. The path taken by the system from the healthy phase to the panic phase is different from the path taken by the system from the panic phase back to the healthy phase. For systems where two variables have asymmetric relationships (e.g., the Lotka–Volterra predator-prey model), this type of behavior is not uncommon. It indicates that maybe the process of a panic attack is a one-way street: if a panic attack has started, it may not be possible to stop it before the *fear* level reaches its peak. Further, because the path taken by the system from the healthy phase to the panic phase is different to the path from the panic phase to the healthy phase, maybe it is possible to tell the transition direction from the state of the system.[7]

Influence of parameters

As mentioned earlier, *AS* is an important variable in the model that controls the stability of the phases. In the constant *AS* model, *AS* is a parameter instead of a variable. Therefore, we can first investigate *AS* again using the method for parameters. We constructed a series of landscapes from the constant *AS* model (Figure 8a). Comparing those landscapes, we can find that when *AS* is 0.3 or lower, there is only one phase in the system, namely the healthy phase. The panic phase only appears when *AS* is 0.7 or higher. For the critical condition that $AS = 0.5$, the system can go to

[7]The landscape for *AS*, *E*, and *fear* also confirms the results above. In this landscape plot, three variables are represented with *x*-, *y*-, and *z*-axis, and the potential value is represented by color. See the OSF repository (https://osf.io/ke3xb/) for the visualization of this landscape.

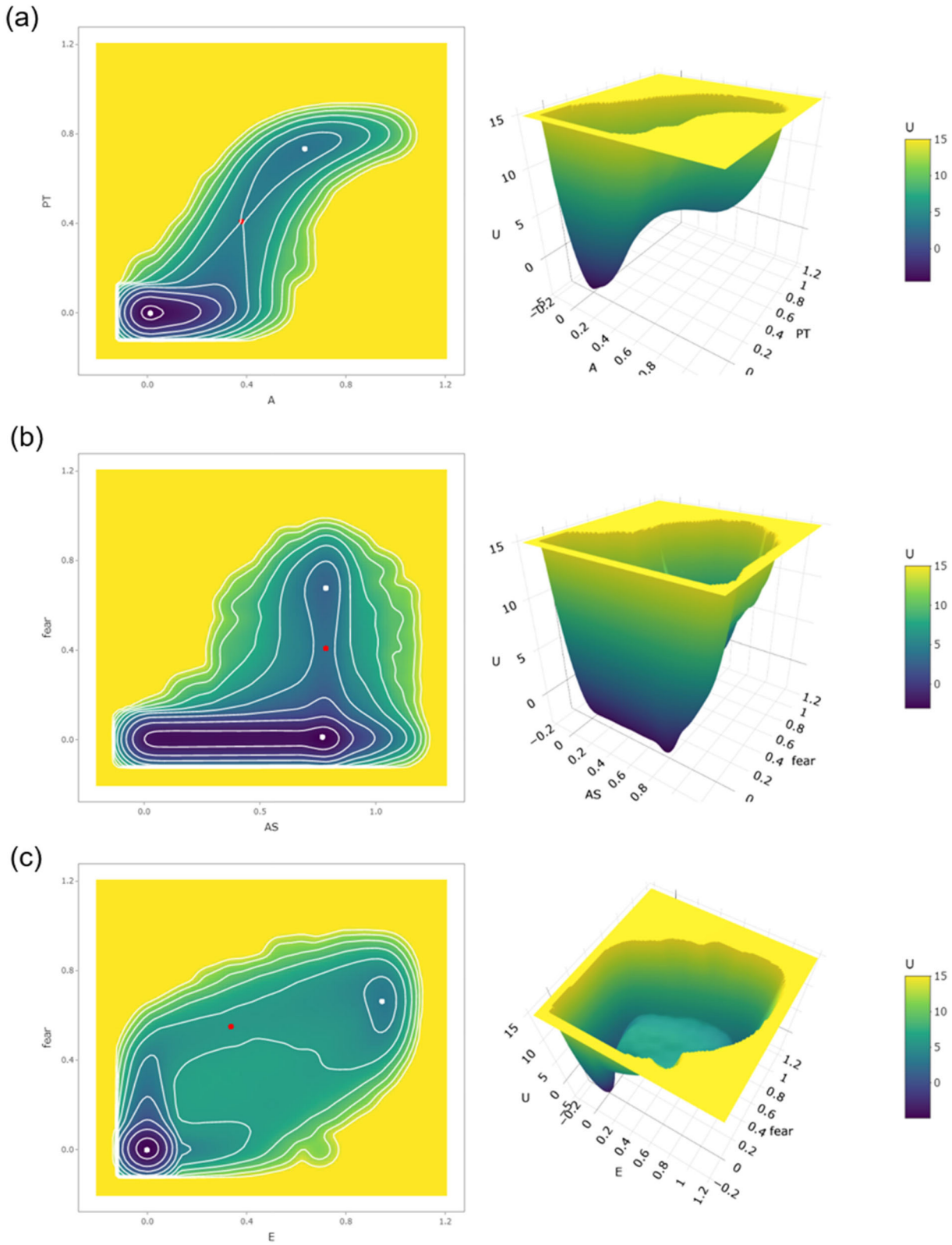

Figure 7. Bivariate potential landscapes of (a) *A* and *PT*, (b) *AS* and *fear*, and (c) *E* and *fear*. All three landscapes were constructed using the simplified model. The first plot in each row is the 2D heatmap with contours, and the second plot is the 3D surface plot. The white and red dots on the 2D heatmaps represent the local minima of the phases and the saddle points, respectively.

the region of panic phase, but there is not a local minimum in that region (i.e., the potential landscape increases monotonically in the direction toward higher *A* and *E*). We call it a *quasi-stable phase*. Because there is no local minimum for a quasi-stable phase, its potential value and barrier height cannot be calculated.

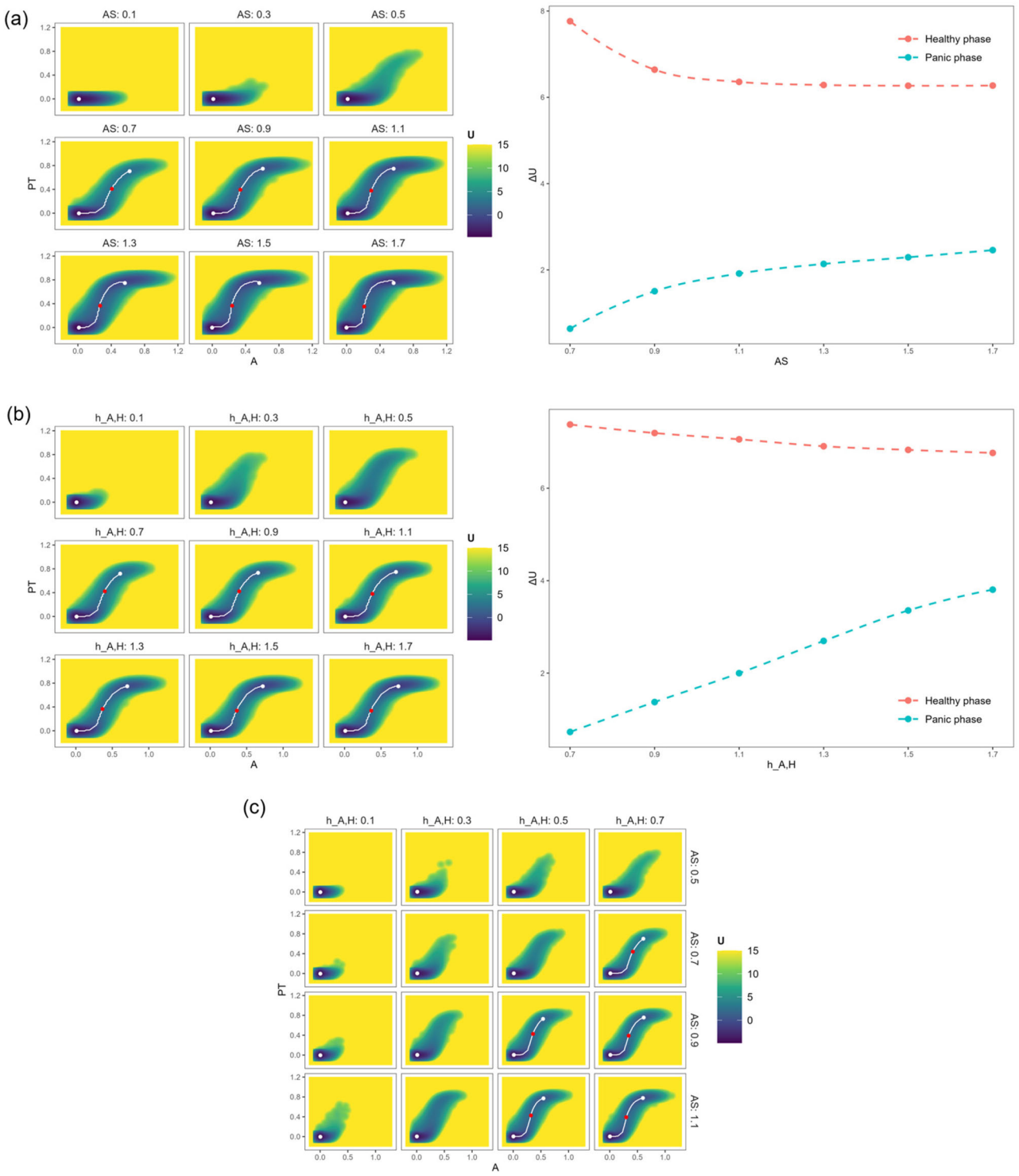

Figure 8. Bivariate potential landscapes of *A* and *PT* with different (a) *AS*, (b) $h_{A,H}$, and (c) both *AS* and $h_{A,H}$. The landscapes were constructed using the constant *AS* model. The first plot in each row is the potential landscapes represented in 2D heatmaps. The white dots represent the local minima of the phases in the landscape, the white lines connecting two white dots represent the minimum energy path between two local minima, and the red dots represent the saddle points in the paths. The second plot in (a) and (b) shows the barrier heights (ΔU) of both phases in the potential landscapes for different parameter values. For cases where there is only one phase in the system, it is not possible for the system to transition to an alternative phase, hence the barrier height cannot be defined. Therefore, barrier heights were not calculated for those cases.

The barrier heights for the two phases are also shown in Figure 8a. As *AS* becomes larger, the healthy phase gradually becomes more unstable, while the panic phase becomes more stable. Nevertheless, the healthy phase is always more stable than the panic phase. These results again confirm the findings above.

While *AS* is related to the relationship between *A* and *PT*, which is at the core of the system, there are also some other variables whose effects are less obvious. Here we investigate another parameter, $h_{A,H}$, as an example. This parameter represents the relationship between physical arousal (*A*) and homeostatic feedback of arousal (*H*): the physical arousal at a later point in time tends to reduce more if it is higher in the previous time point. *A* has a positive influence on *H* and *H* has a negative influence on *A*. Therefore, the state of *A* is "stored" in *H* and will be influenced by *H* later. When $h_{A,H}$ is higher, the influence of *A* on *H* is weaker.

The landscapes and barrier heights are shown in Figure 8b. Comparing the landscapes with different $h_{A,H}$ values (Figure 8b) and those with different *AS* values (Figure 8a), we can find that both parameters have a similar role in controlling the landscape of *A* and *PT*. When arousal has weaker homeostatic feedback of arousal (i.e., when the delayed negative feedback gets weaker), the panic phase gradually appears and stabilizes. If changing both parameters together, their roles are similar and independent: increasing either will stabilize the healthy phase of the system (see Figure 8c). This suggests that the same phenomenon of the system (e.g., panic disorder) can stem from different underlying mechanisms.

Guidance for model modification

Potential landscapes can also guide model modification, as we show here with the panic disorder model as an example. As mentioned by Robinaugh et al. (2019), the problem of the original model is that the system should have non-clinal panic attacks, but the model failed to generate them. From the landscape with *AS* and *fear* (Figure 7b), we can see that the panic phase only appears when *AS* is high enough. However, the landscape in the *AS* direction is rather flat, which means that *AS* does not have a strong tendency to increase or decrease within its plausible range. This is related to the problem mentioned above because whenever *AS* happens to be in the high range, it does not go back. Hence, an increase in *AS* can easily lead to a panic disorder.

To solve this problem, what we should do is tilt the landscape to make the low-*AS* region more stable. The way to tilt the landscape is straightforward. As the gradient of the landscape corresponds to the changing rate of the variables, we can simply add a negative term to the dynamic function to tilt it toward zero. Here we added a small negative term to the time derivative of *AS*:

$$\frac{dAS}{dt} = g(X) + \begin{cases} -r_{\text{extinction}}(AS - AS_{\text{baseline}}), & AS > AS_{\text{baseline}} \\ 0, & AS \le AS_{\text{baseline}} \end{cases} \quad (9)$$

where g(X) represents the terms in the original functions, and the term to the right of the brace is the added term: when *AS* is larger than a given baseline value (AS_{baseline}), it declines exponentially. This setting can add a small tendency toward lower *AS* around its critical range while not letting it decline to zero. To distinguish this way of modification and the way in the original paper by Robinaugh et al. (2019, which used a higher S_E to increase the parameter $h_{PT,E}$ as a constant; see Equation 4), we refer to the modification in Equation 9 as the AS *extinction modification* and the modification by Robinaugh et al. (2019) as the *escape schema modification.*

Both modifications show a good effect on solving the problem in the original model. As shown in Figure 9, for both modifications, the system can have a limited number of panic attacks without later going into a full-blown panic disorder. After the first panic attack, the *AS* value of the original model increases and stays in a higher value until the next panic attack, which further increases *AS* (Figure 3), whereas the *AS* value decreases after the first panic attack for both modified models. The difference in the direction of change for *AS* shows the effects of both modification mechanisms. The potential landscapes of both modifications are shown in Figure 10. Comparing them with the landscape of the original model (Figure 7b), the landscapes of those modified models show the expected tendency that the lower-*AS* region of the healthy phase is more stable.

However, there are several important differences between the two modifications. For the escape schema modification, the *fear* value reaches a higher peak level and the *AS* value directly decreases after the panic attack; in the *AS* extinction model, the peak *fear* value is similar to the case in the original model, and the *AS* value first increases before it declines back to a lower value (Figure 9). Why is this the case? The landscapes of the models can provide further information. The general shape of the landscape of the *AS* extinction modification (Figure 10a) is similar to the original model (Figure 7b), but only the stability of the lower *AS* range of the healthy phase is increased. The shape of the landscape of the escape schema modification (Figure 10b), however, is much different from the original model, especially around the panic phase. There is a small island in the middle between

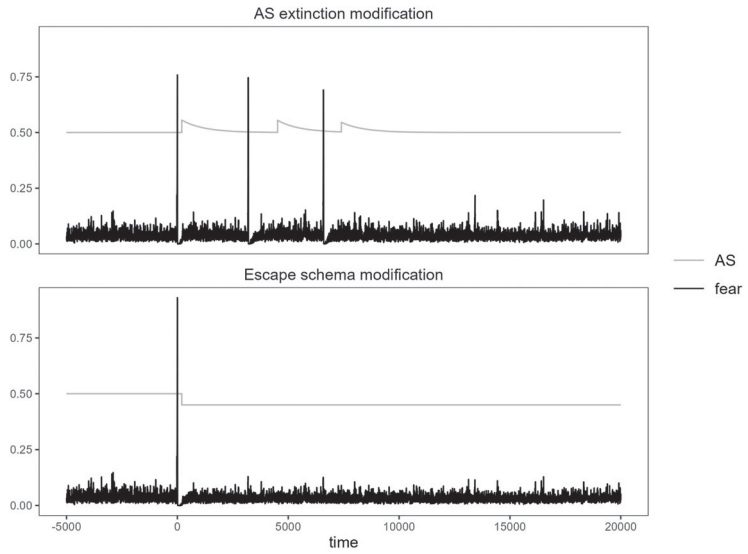

Figure 9. Model simulation results of *fear* and *AS* for the two modifications. The first panic attack of the system appears probabilistically. To make simulation results comparable, *time* was set as zero at the first panic attack (the time when the peak value of *fear* is reached). In the *AS* extinction modification, $r_{extinction} = 0.001$, $AS_{baseline} = 0.5$; in the escape schema modification, $S_E = 0$.

the two phases. As explained earlier (for the landscape of *E* and *fear* in Figure 7c), this indicates that the pathway the system takes from the healthy phase to the panic phase is different than the other way around, so that the system does not recover in the same way. In this model, this means that the system goes to the panic phase when *AS* is high, and then the *AS* level declines *during* the same panic attack. When the system goes back to the healthy phase, the *AS* value is already at a lower level. In the original model and the *AS* extinction modification, *AS* decreases at a slower time scale, so that the system recovers to the healthy phase with a similarly high *AS* value. These differences enable further theoretical and empirical examinations of these two modifications.

Discussion

We here introduced a new method to construct potential landscapes for multivariate psychological formal models. Based on the steady-state distribution, the stability of any state of the system can be quantified with the potential function. We illustrated the method with the panic disorder model by Robinaugh et al. (2019). After several adjustments to the original model, we constructed potential landscapes for the system and analyzed the absolute, relative, and geometric stability of the healthy phase and the panic phase. Then, we examined the influence of two model parameters on the potential landscape of the system. Finally, based on the information from the potential landscape, we came up with a new way of model modification and compared the simulation output and the potential landscape of it with the modification suggested by Robinaugh et al. (2019).

The results of the potential landscape showed that there are one or two phases in the system, depending on the parameter settings. The healthy phase is always present, while the panic phase only appears under certain conditions (e.g., with high *AS*) and is always less stable than the healthy phase. Increasing *AS* can thus stabilize the panic phase and destabilize the healthy phase. These results are well aligned with the conclusions from the case simulations in the original paper. It is important to note that we do not think that the potential landscape method can or should replace case simulations. On the contrary, we claim that both methods provide important information about the nature of psychological formal models. The advantage of our method is that it can present the concept of stability in a clear, explicit way. By filtering out the time-related information in the model output, the stability-related information that does not change over time emerges clearly on the potential landscapes. Instead of relying on the heuristics from observation, the stability of states is now specified as positions in a potential landscape. Specifically, the stability of psychological phases can be described accurately based on three aspects: absolute stability, relative stability, and geometric stability. This makes the stability information more apparent for researchers, hence facilitating understanding and communication of the models.

Moreover, the potential landscape method also enables easy investigation of variables and parameters. A common problem for psychological formal models is that the number of variables and parameters used for

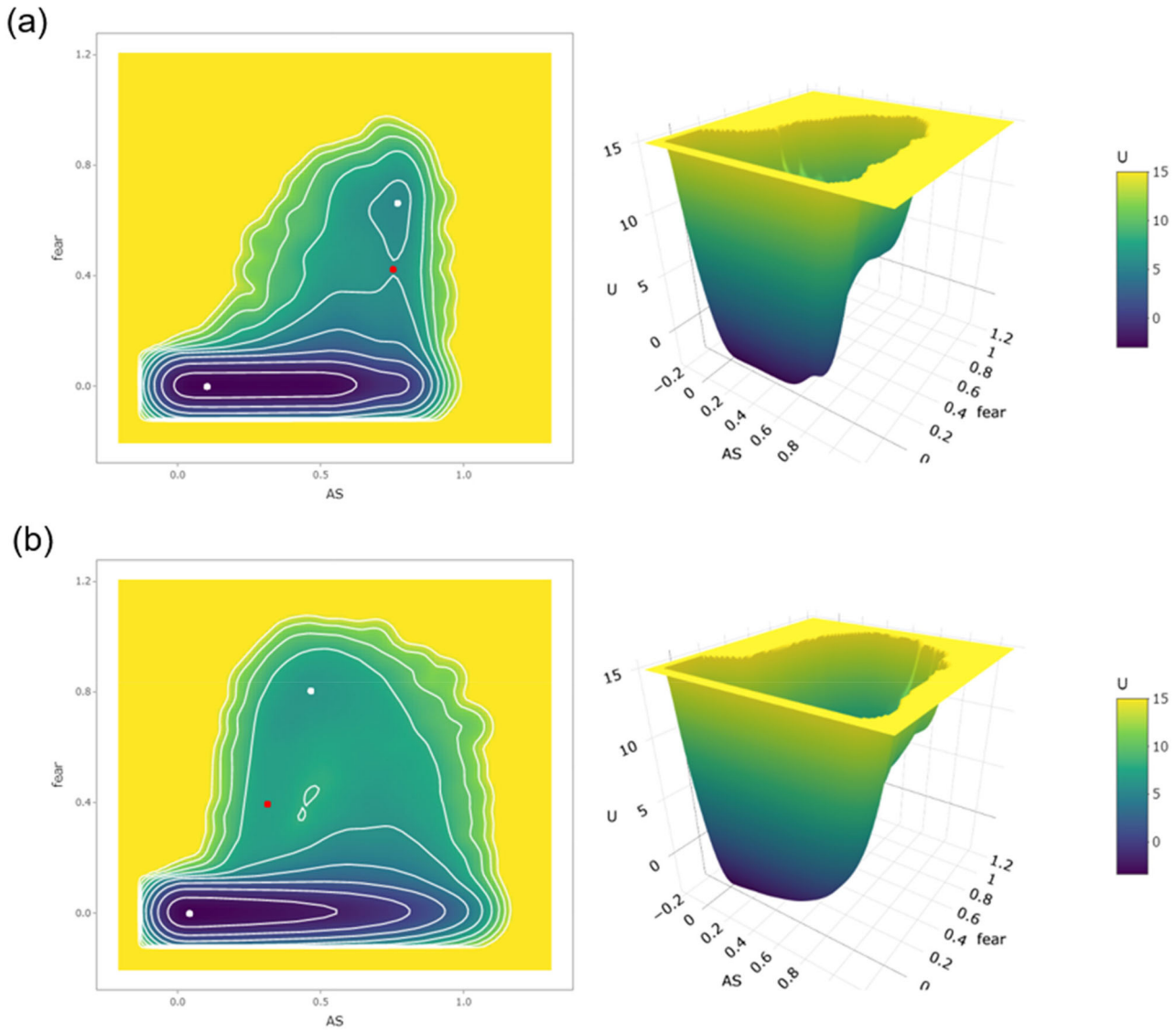

Figure 10. Bivariate potential landscapes of *AS* and *fear* for (a) the *AS* extinction modification and (b) the escape schema modification. The first plot in each row is the 2D heatmap with contours, and the second plot is the 3D surface plot. The white and red dots on the 2D heatmaps represent the local minima of the phases and the saddle points, respectively.

constructing the model is much higher than the critical variables that are examined to evaluate the model. In this paper, we showed how to construct the potential landscape for *A*, *PT*, *AS*, *E*, and *fear* and how the potential landscape changes with *AS* and $h_{A,H}$. With the tools we provided in the simlandr package, the same landscape construction method can be easily extended to all other variables and parameters in the model. Our method allows researchers to evaluate each of their variables and parameters systematically and see if the resulting potential landscape is consistent with theory and empirical findings, which parameter range produces the expected behavior, and how the parameters influence the stability. This not only makes model evaluation more effective and comprehensive but also helps to clarify the scope and boundaries of a model or theory.

Based on the information from the potential landscape, we provided a new modification of the original model. The advantage of using a stability measure to guide model modification is that, in some cases, the discrepancy between model output and real-life phenomena is more closely related to the stability of states or phases rather than the simulated trajectories. In the case of the panic disorder model, for example, the problem of non-clinical panic attacks can be directly attributed to the geometric stability of the healthy phase (i.e., its landscape on the *AS* direction is too flat). Also, because of the close link between the potential landscape and the dynamic functions, changing the landscape of the system is usually not difficult, making it straightforward to solve the identified problems. In this work, we proposed a way of modification

for the panic disorder model based on its potential landscape. There is also a possible theoretical explanation for this modification, namely *fear extinction*. After the association (arousal schema, *AS*) between physical arousal (*A*) and perceived threat (*PT*) is learned, even when there are no new fear-inducing events, the strength of this association still decreases (Mattera et al., 2020; Milad & Quirk, 2012). It is important to emphasize that the decision to add a negative value to the *AS* time derivative in the *AS* extinction modification was inspired by the potential landscape of the model, not based on a theoretical analysis. Nevertheless, it later helps to point out a direction in which researchers could search for relevant theories.

Limitations and possible pitfalls

Despite the advantages of the method discussed above, we also want to point out several limitations and possible pitfalls of our method. First, the potential landscape is calculated from a generalized potential function. This means that it does not contain all the dynamic information of the model. In other words, psychological systems are not totally the same as a ball on a landscape. Some additional non-gradient forces also influence the system, which can drive it in a different direction instead of the direct path toward the most stable state. For example, in the landscape of *E* and *fear*, there are two paths connecting two phases of the system. We suppose these paths are different in directions (i.e., the system state goes through one path to the healthy phase and the other path to the panic phase). However, the choice of which path to take is influenced by some non-gradient forces, which are not shown in the potential landscape. In psychology, non-gradient forces may be related to emotional or behavioral inertia, which means that some emotional or behavioral variables in the system may be more resistant to change (e.g., *E* and *fear* in the panic disorder model; Alós-Ferrer et al., 2016; Kuppens et al., 2010). These non-gradient forces are canceled out when calculating the stability of states. Nevertheless, researchers should be aware of their existence and investigate their influences if the dynamic properties of the system are also a concern.

Second, to ensure the ergodicity in the Monte Carlo simulation, we added some noise terms to the original model. We suggest that researchers do the same if their simulation is difficult to converge. However, we want to point out that for some systems these noise terms can affect the stability of phases. For example, Van den Broeck et al. (1994) showed that for a specific kind of dynamic system, an ordered phase only exists if the noise is in a certain range. The conditions under which noise terms can affect the stability of phases are not yet clear. Further investigations on this issue are needed.

Third, although psychological models usually have a large number of variables, the potential landscape can only be visualized in a lower (up to three) dimensional space. This means one phase or transition path on the landscape can actually be several ones in the high-dimensional space. For example, there is only one pathway between the healthy phase and the panic phase in the landscape of *A* and *PT*, but two pathways in the landscape of *fear* and *E*. Therefore, we suggest researchers look into different combinations of variables and investigate how the phases and paths in those low dimension landscapes correspond.

Fourth, the potential landscape can only show the stability of individual states. For a system that has periodic or chaotic phases (e.g., Schiepek et al., 2017), the potential landscape may not be able to clearly represent its stability. In these cases, some preprocessing of the original variables may be needed. For example, if two phases of a system have similar mean values but differ in their variation, then the moving-window standard deviation can be used to construct the landscape instead of the original variable values. The suitable way of transformation depends on the exact system of interest.

Future directions

Incorporating the potential landscape method opens up new avenues for psychological (modeling) research. Here we want to point out several possible directions. First, there is still much room to further develop the potential landscape method for psychological models. Apart from the generalized potential function used in our work, there are also several other potential functions that have different definitions and calculation methods (see Zhou & Li, 2016, for a review). These methods may have a higher requirement for the form of dynamic functions, but they also have strengths in representing other aspects of stability (e.g., the quasi-potential landscape by Zhou et al., 2012, emphasizes more on the transition path between states). We encourage future research to test the use of those methods for psychological models.

Second, we encourage modelers to further explore the usage of this method and apply this method to different kinds of models. In this paper, we showed several applications of the potential landscape

methods, namely representing stability, investigating the influence of parameters, and guiding modifications, and the formal model we used is an emotion-cognition-behavior model of a mental disorder. However, we are confident that the usage of this method is not limited to this range. For example, it may be possible to draw a *phase diagram* to show the parameter ranges where different phases exist and use the potential landscape as a way to choose parameter values systematically. Some cognitive models also have multistability (e.g., Kogo et al., 2011), which can possibly be analyzed with this method. We look forward to future researchers exploring those possibilities.

Third, potential landscapes can also be helpful in clinical practice. Understanding individual differences of psychological phases and clinical change processes is critical to establish more effective therapies. From the potential landscape perspective, the difference in people's parameter values makes them differ in their vulnerability to mental disorders, and successful therapies are related to changes in parameter values. This idea can be better understood if we can construct landscapes with different parameters to represent different individuals or different therapeutic stages. We illustrated how systems with different AS and $h_{A,H}$ values differ in the stability of the healthy phase and the panic phase. These AS and $h_{A,H}$ values may correspond to different individuals and/or stages. Apart from panic disorder, this approach can also be readily applied for conceptualizing other mental disorders. Constructing landscapes from empirical data, however, requires future methodological development. Previous research found that recovering the psychological dynamics from experience sampling data is difficult, while the distribution of psychological variables is more assessable (Haslbeck & Ryan, 2021). The generalized potential landscape is defined from the steady-state distribution of the system, which can be possibly estimated from the observed variable distribution. Therefore, maybe fitting the potential landscape directly from empirical data is a possible alternative to fitting the dynamic functions.

Finally, we want to note that, in the abstract sense, the form of psychological formal models does not differ much from many dynamic models in biology, chemistry, and other natural science fields. For example, we coincidentally found out that the panic disorder model also has an equivalent form represented by chemical reactions (see *S4. Equivalent chemical representations* in the Supplementary Materials). While formalizing psychological theory is a new trend in psychology, the same effort has been undertaken in many fields, with various analytical methods readily available. With the current work as an example, we would encourage further interdisciplinary cooperation in the psychological modeling field. We hope the insights from expertise in all fields of science can somehow come together, helping us to better understand human psychology, a complex yet fascinating subject.

Conclusion

The stability of states and phases is an important property for psychological formal models, yet not concretely addressed with the common case simulation method. By incorporating the generalized potential function by Wang et al. (2008) and Monte Carlo simulation, we developed a method to construct the potential landscape for multivariate psychological dynamic models. This method can contribute to a better understanding of the stability concept, the influence of model parameters, and the way to modify a model. We hope this method can help researchers to better evaluate and develop their models and ultimately help to guide clinical practice in the future.

Article information

Conflict of interest disclosures: Each author signed a form for disclosure of potential conflicts of interest. No authors reported any financial or other conflicts of interest in relation to the work described.

Ethical principles: The authors affirm having followed professional ethical guidelines in preparing this work. These guidelines include obtaining informed consent from human participants, maintaining ethical treatment and respect for the rights of human or animal participants, and ensuring the privacy of participants and their data, such as ensuring that individual participants cannot be identified in reported results or from publicly available original or archival data.

Funding: This work was supported by the Behavioural Science Institute, Radboud University. TL was supported by the NSFC under Grant No. 11825102 and the Beijing Academy of Artificial Intelligence (BAAI). ALA was supported by a NWO VIDI grant, Grant No. VI.Vidi.191.178.

Role of the funders/sponsors: None of the funders or sponsors of this research had any role in the design and conduct of the study; collection, management, analysis, and interpretation of data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

References

Alós-Ferrer, C., Hügelschäfer, S., & Li, J. (2016). Inertia and decision making. *Frontiers in Psychology*, *7*, Article 169. https://doi.org/10.3389/fpsyg.2016.00169

Barton, S. (1994). Chaos, self-organization, and psychology. *The American Psychologist*, *49*(1), 5–14. https://doi.org/10.1037/0003-066X.49.1.5

Borsboom, D., van der Maas, H. L. J., Dalege, J., Kievit, R. A., & Haig, B. D. (2021). Theory construction methodology: A practical framework for building theories in psychology. *Perspectives on Psychological Science*, *16*(4), 756–766. https://doi.org/10.1177/1745691620969647

Burger, J., van der Veen, D. C., Robinaugh, D. J., Quax, R., Riese, H., Schoevers, R. A., & Epskamp, S. (2020). Bridging the gap between complexity science and clinical practice by formalizing idiographic theories: A computational model of functional analysis. *BMC Medicine*, *18*, Article 99. https://doi.org/10.1186/s12916-020-01558-1

Cramer, A. O. J., Borkulo, C. D. v., Giltay, E. J., Maas, H. L. J. v d., Kendler, K. S., Scheffer, M., & Borsboom, D. (2016). Major depression as a complex dynamic system. *PLoS One*, *11*(12), Article e0167490. https://doi.org/10.1371/journal.pone.0167490

Cui, J., Olthof, M., Lichtwarck-Aschoff, A., Li, T., & Hasselman, F. (2021). simlandr: Simulation-based landscape construction for dynamical systems. PsyArXiv. https://doi.org/10.31234/osf.io/pzva3

Dablander, F., Pichler, A., Cika, A., & Bacilieri, A. (2022). Anticipating critical transitions in psychological systems using early warning signals: Theoretical and practical considerations. *Psychological Methods*. Advance online publication. https://doi.org/10.1037/met0000450

Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*, *1*, 269–271. https://doi.org/10.1007/BF01386390

E, W., & Vanden-Eijnden, E. (2010). Transition-path theory and path-finding algorithms for the study of rare events. *Annual Review of Physical Chemistry*, *61*(1), 391–420. https://doi.org/10.1146/annurev.physchem.040808.090412

Eddelbuettel, D. (2020). *RcppZiggurat: "Rcpp" integration of different 'Ziggurat' normal RNG Implementations*. https://CRAN.R-project.org/package=RcppZiggurat

Eddelbuettel, D., & François, R. (2011). Rcpp: Seamless R and C++ integration. *Journal of Statistical Software*, *40*(8), 1–18. https://doi.org/10.18637/jss.v040.i08

Goldstein, J. (1999). Emergence as a construct: History and issues. *Emergence*, *1*(1), 49–72. https://doi.org/10.1207/s15327000em0101_4

Granic, I., & Hollenstein, T. (2003). Dynamic systems methods for models of developmental psychopathology. *Development and Psychopathology*, *15*(3), 641–669. https://doi.org/10.1017/s0954579403000324

Haslbeck, J., & Ryan, O. (2021). Recovering within-person dynamics from psychological time series. *Multivariate Behavioral Research*. Advance online publication. https://doi.org/10.1080/00273171.2021.1896353

Haslbeck, J., Ryan, O., Robinaugh, D., Waldorp, L., & Borsboom, D. (2021). Modeling psychopathology: From data models to formal theories. *Psychological Methods*. Advance online publication. https://doi.org/10.1037/met0000303

Hayes, A. M., Yasinski, C., Ben Barnes, J., & Bockting, C. L. H. (2015). Network destabilization and transition in depression: New methods for studying the dynamics of therapeutic change. *Clinical Psychology Review*, *41*, 27–39. https://doi.org/10.1016/j.cpr.2015.06.007

Heymann, M., & Vanden-Eijnden, E. (2008). Pathways of maximum likelihood for rare events in nonequilibrium systems: Application to nucleation in the presence of shear. *Physical Review Letters*, *100*(14), Article 140601. https://doi.org/10.1103/PhysRevLett.100.140601

Kogo, N., Galli, A., & Wagemans, J. (2011). Switching dynamics of border ownership: A stochastic model for bi-stable perception. *Vision Research*, *51*(18), 2085–2098. https://doi.org/10.1016/j.visres.2011.08.010

Kuppens, P., Allen, N. B., & Sheeber, L. B. (2010). Emotional inertia and psychological maladjustment. *Psychological Science*, *21*(7), 984–991. https://doi.org/10.1177/0956797610372634

Lamothe, K. A., Somers, K. M., & Jackson, D. A. (2019). Linking the ball-and-cup analogy and ordination trajectories to describe ecosystem stability, resistance, and resilience. *Ecosphere*, *10*(3), Article e02629. https://doi.org/10.1002/ecs2.2629

Li, C., & Wang, J. (2013). Quantifying cell fate decisions for differentiation and reprogramming of a human stem cell network: Landscape and biological paths. *PLoS Computational Biology*, *9*(8), Article e1003165. https://doi.org/10.1371/journal.pcbi.1003165

Li, C., & Ye, L. (2019). Landscape and flux govern cellular mode-hopping between oscillations. *The Journal of Chemical Physics*, *151*(17), Article 175101. https://doi.org/10.1063/1.5125046

Mattera, A., Pagani, M., & Baldassarre, G. (2020). A computational model integrating multiple phenomena on cued fear conditioning, extinction, and reinstatement. *Frontiers*

in Systems Neuroscience, *14*, Article 569108. https://doi.org/10.3389/fnsys.2020.569108

Milad, M. R., & Quirk, G. J. (2012). Fear extinction as a model for translational neuroscience: Ten years of progress. *Annual Review of Psychology*, *63*, 129–151. https://doi.org/10.1146/annurev.psych.121208.131631

Olthof, M., Hasselman, F., & Lichtwarck-Aschoff, A. (2020). Complexity in psychological self-ratings: Implications for research and practice. *BMC Medicine*, *18*, Article 317. https://doi.org/10.1186/s12916-020-01727-2

Olthof, M., Hasselman, F., Oude Maatman, F., Bosman, A. M. T., & Lichtwarck-Aschoff, A. (in press). Complexity theory of psychopathology. *Journal of Psychopathology and Clinical Science*. https://doi.org/10.31234/osf.io/f68ej

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Robinaugh, D. J., Haslbeck, J., Ryan, O., Fried, E. I., & Waldorp, L. J. (2021). Invisible hands and fine calipers: A call to use formal theory as a toolkit for theory construction. *Perspectives on Psychological Science*, *16*(4), 725–743. https://doi.org/10.1177/1745691620974697

Robinaugh, D. J., Haslbeck, J., Waldorp, L., Kossakowski, J. J., Fried, E. I., Millner, A., McNally, R. J., van Nes, E. H., Scheffer, M., Kendler, K. S., & Borsboom, D. (2019). Advancing the network theory of mental disorders: A computational model of panic disorder, PsyArXiv. https://doi.org/10.31234/osf.io/km37w

Rodríguez-Sánchez, P., Nes, E. H. v., & Scheffer, M. (2020). Climbing Escher's stairs: A way to approximate stability landscapes in multidimensional systems. *PLoS Computational Biology*, *16*(4), Article e1007788. https://doi.org/10.1371/journal.pcbi.1007788

Sakurambo. (2005). *An illusion of an endless staircase using false perspective* [Digital painting]. Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Impossible_staircase.svg

Sayama, H. (2015). Continuous-time models II: Analysis. In *Introduction to the modeling and analysis of complex systems* (pp. 111–129).

Schiepek, G., Tominschek, I., & Heinzel, S. (2014). Self-organization in psychotherapy: Testing the synergetic model of change processes. *Frontiers in Psychology*, *5*, Article 1089. https://doi.org/10.3389/fpsyg.2014.01089

Schiepek, G., Viol, K., Aichhorn, W., Hütt, M.-T., Sungler, K., Pincus, D., & Schöller, H. J. (2017). Psychotherapy is chaotic—(Not only) in a computational world. *Frontiers in Psychology*, *8*, Article 379. https://doi.org/10.3389/fpsyg.2017.00379

Van den Broeck, C., Parrondo, J. M. R., & Toral, R. (1994). Noise-induced nonequilibrium phase transition. *Physical Review Letters*, *73*(25), 3395–3398. https://doi.org/10.1103/PhysRevLett.73.3395

Wang, J., Xu, L., & Wang, E. (2008). Potential landscape and flux framework of nonequilibrium networks: Robustness, dissipation, and coherence of biochemical oscillations. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(34), 12271–12276. https://doi.org/10.1073/pnas.0800579105

Wichers, M., Schreuder, M. J., Goekoop, R., & Groen, R. N. (2019). Can we predict the direction of sudden shifts in symptoms? Transdiagnostic implications from a complex systems perspective on psychopathology. *Psychological Medicine*, *49*(3), 380–387. https://doi.org/10.1017/S0033291718002064

Zhang, X., Chong, K. H., & Zheng, J. (2020). A Monte Carlo method for in silico modeling and visualization of Waddington's epigenetic landscape with intermediate details. *Biosystems*, *198*, Article 104275. https://doi.org/10.1016/j.biosystems.2020.104275

Zhou, J. X., Aliyu, M. D. S., Aurell, E., & Huang, S. (2012). Quasi-potential landscape in complex multi-stable systems. *Journal of the Royal Society, Interface*, *9*(77), 3539–3553. https://doi.org/10.1098/rsif.2012.0434

Zhou, P., & Li, T. (2016). Construction of the landscape for multi-stable systems: Potential landscape, quasi-potential, A-type integral and beyond. *The Journal of Chemical Physics*, *144*, Article 094109. https://doi.org/10.1063/1.4943096