

RESEARCH ARTICLE

A hierarchical stochastic model for bistable perception

Stefan Albert¹, Katharina Schmack², Philipp Sterzer², Gaby Schneider^{1*}

1 Institute of Mathematics, Goethe University, Frankfurt (Main), Germany, **2** Department of Psychiatry and Psychotherapy, Charité Universitätsmedizin Berlin, Germany

* schneider@math.uni-frankfurt.de



Abstract

Viewing of ambiguous stimuli can lead to bistable perception alternating between the possible percepts. During continuous presentation of ambiguous stimuli, percept changes occur as single events, whereas during intermittent presentation of ambiguous stimuli, percept changes occur at more or less regular intervals either as single events or bursts. Response patterns can be highly variable and have been reported to show systematic differences between patients with schizophrenia and healthy controls. Existing models of bistable perception often use detailed assumptions and large parameter sets which make parameter estimation challenging. Here we propose a parsimonious stochastic model that provides a link between empirical data analysis of the observed response patterns and detailed models of underlying neuronal processes. Firstly, we use a Hidden Markov Model (HMM) for the times between percept changes, which assumes one single state in continuous presentation and a stable and an unstable state in intermittent presentation. The HMM captures the observed differences between patients with schizophrenia and healthy controls, but remains descriptive. Therefore, we secondly propose a hierarchical Brownian model (HBM), which produces similar response patterns but also provides a relation to potential underlying mechanisms. The main idea is that neuronal activity is described as an activity difference between two competing neuronal populations reflected in Brownian motions with drift. This differential activity generates switching between the two conflicting percepts and between stable and unstable states with similar mechanisms on different neuronal levels. With only a small number of parameters, the HBM can be fitted closely to a high variety of response patterns and captures group differences between healthy controls and patients with schizophrenia. At the same time, it provides a link to mechanistic models of bistable perception, linking the group differences to potential underlying mechanisms.

OPEN ACCESS

Citation: Albert S, Schmack K, Sterzer P, Schneider G (2017) A hierarchical stochastic model for bistable perception. *PLoS Comput Biol* 13(11): e1005856. <https://doi.org/10.1371/journal.pcbi.1005856>

Editor: Wolfgang Einhäuser, Technische Universität Chemnitz, GERMANY

Received: July 25, 2017

Accepted: October 29, 2017

Published: November 20, 2017

Copyright: © 2017 Albert et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by the German Federal Ministry of Education and Research (www.bmbf.de) within the framework of the e:Med research and funding concept (grant number 01ZX1404B to SA and GS, 01ZX1404A to KS), by the Priority Program 1665 of the Deutsche Forschungsgemeinschaft (grant number SCHN 1370/02-1 to GS, www.dfg.de) and by a Research Grant from Deutsche Forschungsgemeinschaft

Author summary

Patients suffering from schizophrenia show specific cognitive deficits. One way to study these cognitive phenomena works with the presentation of ambiguous stimuli. During viewing of an ambiguous stimulus, perception alters spontaneously between different percepts. Percept changes occur as single events during continuous presentation, whereas

(grant number STE 1430/7-1 to PS). KS is participant in the Charité Clinical Scientist Program funded by the Charité Universitätsmedizin Berlin and the Berlin Institute of Health. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

during intermittent presentation, percept changes occur at regular intervals either as single events or bursts. Here we investigate perceptual responses to continuous and intermittent stimulation in healthy control subjects and patients with schizophrenia. Interestingly, the response patterns can be highly variable but show systematic group differences. We propose a model that connects these perceptual responses to underlying neuronal processes. The model mainly describes the activity difference between competing neuronal populations on different neuronal levels. In a hierarchical manner, the differential activity generates switching between the conflicting percepts as well as between states of higher and lower perceptual stability. By fitting the model directly to empirical responses, a high variety of patterns can be reproduced, and group differences between healthy controls and patients with schizophrenia can be captured. This helps to link the observed group differences to potential neuronal mechanisms, suggesting that patients with schizophrenia tend to spend more time in neuronal states of lower perceptual stability.

Introduction

The phenomenon of bistable perception has fascinated researchers for a long time [1, 2, 3]. Recently, the description of response patterns to bistable stimuli such as the Necker Cube, Rubin's vase or rotating spheres with switching rotation direction gained increasing interest in computational neuroscience [4, 5, 6, 7, 8]. By modeling dynamic changes of perception during viewing of one and the same stimulus, one aims at providing potential explanations for neuronal mechanisms underlying perception and perceptual changes and to identify related brain areas as well as potential dysfunctions, e.g. in schizophrenia [9, 10].

Interestingly, the response patterns to continuously shown bistable stimuli often share common properties [7, 11]. Typically, the distribution of intervals of constant perception, termed dominance times, is unimodal and right-skewed, and extremely short dominance times, i.e., rapidly fluctuating percepts, are rare [12, 13, 14]. The dominance times under continuous stimulation are therefore often modeled as Gamma distributed [15, 16, 17, 18, 19, 20]. The mean of dominance times can be highly variable across subjects [14, 20], whereas the coefficient of variation (CV) is often comparable [21].

In comparison to a continuous presentation, intermittent presentation of a bistable stimulus, i.e., by repetitive interruption of stimulation for short time periods, has been observed to stabilize the percept if the interruption period is long enough, typically longer than 0.7 seconds [15, 18, 22, 23, 24, 25]. In this case, dominance times get longer and can also show a certain degree of periodicity [14]. In addition, such stable phases with long dominance times during intermittent presentation can also interchange with unstable phases of rapid percept changes. Fig 1 shows examples of response patterns to continuous and intermittent presentation of a bistable stimulus from the dataset reported in [10].

Modeling studies with elaborated mathematical models have been proposed that can explain a number of properties of bistable perception like the distribution of dominance times under continuous stimulation [8, 17, 19, 20, 21, 26] or cyclic behavior and the impact of the duration of the stimulus presentation on the dominance times in intermittent stimulation [14, 18]. One key ingredient of these models of bistable perception is typically a competition between neuronal populations that correspond to the different percepts [14, 17, 18, 20, 27]. In order to account for stabilized perception in intermittent viewing, the use of multiple time-scales for memory traces of past perception has been proposed by [14] and [18].

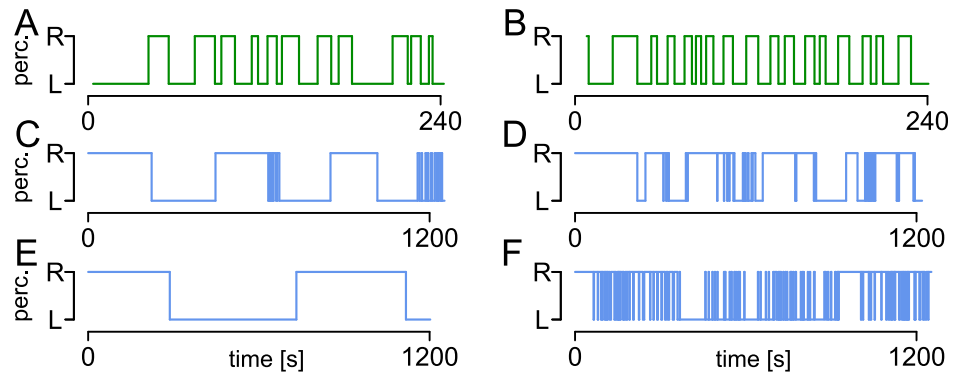


Fig 1. Examples of response patterns to a bistable stimulus. Response patterns to continuous (green, A, B) and intermittent (blue, C-F) presentation from the dataset reported in [10]. While the distribution of dominance times tends to be unimodal in the continuous case, stable and unstable phases interchange in intermittent stimulation. In addition, response patterns can be highly variable across subjects.

<https://doi.org/10.1371/journal.pcbi.1005856.g001>

Many such models require a high number of parameters in order to describe the variety of response patterns. As a consequence, they can often hardly be fitted to experimental data, especially in the typical cases when only a few dozen dominance times are observed. In addition, the majority of models focus either on continuous or on intermittent viewing. Interesting models that are applicable to both cases have been proposed by [14, 17, 18].

The relevance of a joint description of continuous and intermittent viewing is illustrated here on a dataset including responses of patients with schizophrenia and of healthy controls to continuous and intermittent presentation of a rotating sphere with ambiguous rotation direction reported earlier in [9, 10]. In [10], an enhanced alternation rate for the group of patients with schizophrenia during intermittent presentation was reported. Interestingly, when we analyzed previously unpublished data recorded in the same participants during continuous presentation, the opposite could be observed [Fig 2; the data was collected during an initial training run for which the experimental procedures but not the results are described in 10]. Due to the differences in patterns and time scales between continuous and intermittent

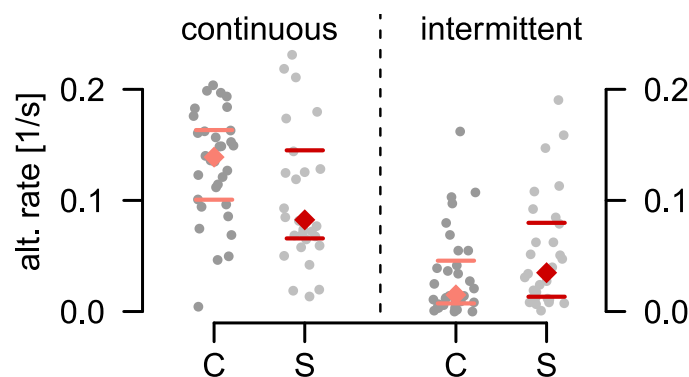


Fig 2. Alternation rates in control subjects and subjects with schizophrenia. During continuous presentation healthy controls (C) showed higher alternation rates compared to patients with schizophrenia (S) (left), while the opposite could be observed for intermittent presentation (right). Each grey dot indicates the perceptual alternation rate from one individual participant, colored diamonds and lines indicate group medians with 25%/75% quantiles. Two-sided Wilcoxon tests yielded $p < .1$ for both continuous and intermittent stimulation.

<https://doi.org/10.1371/journal.pcbi.1005856.g002>

presentation, the potential neuronal mechanisms underlying the transitions between the different response properties remain unclear.

Therefore, we propose here a new model for the description of response patterns to bistable perception that links the observed behavior in continuous and intermittent stimulation to potential underlying neuronal processes. First, the model should be able to describe the high variety of both, continuous and intermittent stimulation within one model framework. Second, we use a minimal number of parameters in order to allow parameter estimation and model fitting to the typically short experimental data. This can then allow the statistical investigation of differences between clinical groups.

Note that strictly speaking, the term ‘dominance time’ refers to slightly different objects in continuous and intermittent viewing. While during continuous presentation, switches occur from a dominant to a suppressed percept (percept-switch), dominance times during intermittent presentation consist of multiple continuous presentation periods, and switches typically occur because of different perceptual choices at the onset of the presentation (percept-choice) [28]. In the present model, the observed sequences of dominance times are treated as conceptually similar. This simplification allows for a parsimonious model description in both continuous and intermittent viewing but may not fully capture the relation between the perceptual processes in the two regimes.

The remainder of the article is organized as follows. First, we use a simple Hidden Markov Model (HMM) that describes the observed perceptual processes with a few parameters. For continuous presentation, one state produces independent and identically distributed dominance times with a two-parametric distribution. For intermittent presentation, switching between stable and unstable phases requires two hidden states with short and long dominance times, respectively. The HMM has the advantage that it allows straightforward model fitting and data description with a minimal number of parameters. However, it remains descriptive and lacks relations to potential underlying mechanisms. Therefore, we link the HMM to a hypothetical underlying stochastic model. This model is termed here Hierarchical Brownian Model (HBM) and intends to describe aggregated underlying neuronal activity, producing the observed behavioral responses.

The HBM is based on two main ideas: First, it assumes that switching between percepts results from two conflicting neuronal populations [cmp., e.g., 18]. In order to minimize the number of parameters, this process is reduced to a simple Brownian motion with drift that fluctuates between two thresholds, where the first passage times indicate state changes [similar to 21]. For continuous presentation, one therefore requires only two parameters, i.e., the drift of the Brownian motion and the threshold. The distribution of the resulting first passage times—i.e., dominance times—is then the same as in the HMM, with a simple relation between the two HBM and the two HMM parameters. Second, in order to describe intermittent presentation in the same model framework, we use a hierarchical model. The idea is to describe the switching between stable and unstable phases that is typical for intermittent presentation by using an analogous threshold crossing mechanism of conflicting neuronal populations. Specifically, we assume a second pair of neuronal populations whose corresponding Brownian motion modulates the drift and threshold of the first population pair and thus causes switching between stable and unstable phases. We give a set of model assumptions under which the HBM parameters are comparable to the HMM parameters, thus allowing both model fitting to experimental data sets and potential relation to underlying mechanisms. The parameter estimation is straightforward using maximum likelihood and the HBM can reproduce both, the unimodal distribution in the continuous presentation and the bimodal distribution of dominance times in the intermittent presentation, including also various different response

patterns. Moreover, it allows the identification of specific differences between the clinical groups in [10] and relates these to the hypothesized underlying processes.

Results

A simple Hidden Markov Model for perceptual responses

As a first step to describe the processes observed in bistable perception, we reduce data analysis to the dominance times, i.e., the times between reported changes of the percept. As described above, the distribution of dominance times tends to be unimodal in the continuous case, while stable and unstable phases interchange in intermittent stimulation.

We denote the dominance times by $d_i, i = 1, 2, \dots, n$. In the unimodal continuous case, we assume the d_i to be the realizations of independent and identically (i.i.d.) distributed random variables $D_i, i = 1, 2, \dots, n$ (Fig 3A), where the Gamma or the Inverse Gaussian (IG) distribution are suitable two-parametric distributions [17, 18, 19, 21]. For comparability with the HBM, we focus on the Inverse Gaussian distribution here. For the intermittent case, we assume a HMM with a stable and an unstable state, which are hidden and produce long and short dominance times, respectively (Fig 3B). This requires two parameters for the switches between states, and two parameters for the distribution of dominance times in each state.

Formally, let $Y := (Y_i)_{i=1, \dots, n}$ describe a Markov chain on $\{S, U\}$, where S and U denote the stable and unstable state, respectively. Let $p_{SU} = 1 - p_{SS}$ and $p_{US} = 1 - p_{UU}$ denote the transition probabilities. The dominance times $(d_i)_{i \in \{1, 2, \dots, n\}}$ are assumed to be Inverse Gaussian distributed and conditionally independent given Y , with mean and standard deviation given by (μ_S, σ_S) for $Y_i = S$ and (μ_U, σ_U) for $Y_i = U$.

Note that independence of dominance times is assumed here in continuous presentation. This assumption enables straightforward parameter estimation and is in agreement with the observation that serial correlations of dominance times are typically not reported [e.g., 29, 30]. However, weak long-term dependencies of dominance times reported under continuous presentation [31] cannot be reproduced in the HMM. As such long-term dependence was not observed in the majority of cases in the present data set, also showing no group differences, we use here the simple assumption of independence. In addition, the two used HMM parameters are sufficient to capture the main group difference in the response properties reflected in the alternation rate.

Parameter estimation and precision of parameter estimates. *Continuous presentation.* For parameter estimation in the continuous case, we simply estimate the parameters of the Inverse Gaussian distribution from the dominance times d_1, \dots, d_n , assuming these are

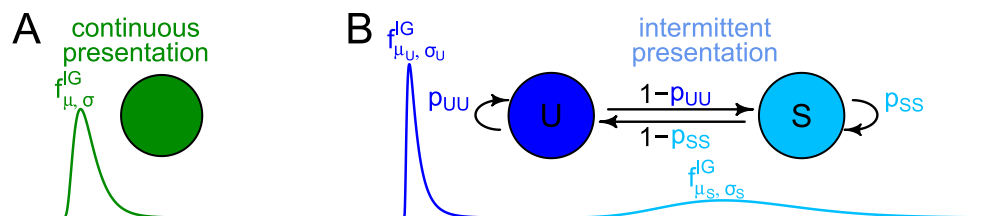


Fig 3. A simple HMM for bistable perception. (A) One state describes a unimodal distribution of dominance times under continuous presentation. (B) Two states (stable, S, and unstable, U) produce long and short dominance times under intermittent presentation.

<https://doi.org/10.1371/journal.pcbi.1005856.g003>

realizations of i.i.d. random variables. Let

$$\bar{\mu}_d := \frac{1}{n} \sum_{i=1}^n d_i$$

denote the sample mean of the observed dominance times. For the Inverse Gaussian (IG) distribution [32] with mean μ and standard deviation σ , the maximum likelihood (ML) estimators are given by

$$\hat{\mu} = \bar{\mu}_d \quad \text{and} \quad \hat{\sigma} = \sqrt{\bar{\mu}_d^3 \frac{1}{n} \sum_{i=1}^n \left(\frac{1}{d_i} - \frac{1}{\bar{\mu}_d} \right)}. \quad (1)$$

The precision of these ML estimators is investigated by parametric bootstrap for the parameter values estimated from the 61 response patterns to continuous presentation in the sample dataset of [10]. For each parameter combination $(\mu_i, \sigma_i)_{i=1, \dots, 61}$ we simulated 1000 response patterns with length $T = 240$ as in the original data. We then compared the estimators $(\hat{\mu}_i, \hat{\sigma}_i)_{i=1, \dots, 61}$ with the parameter values underlying the simulation using the relative error (RE), defined e.g. for μ as $\text{RE}(\mu) := |\hat{\mu} - \mu|/\mu$. The median relative errors for the 61 parameter constellations are shown in Fig 4A. Out of these, 54 (89%) showed estimation errors with median REs less than 0.25 (across the two parameters μ and σ , black). The remaining simulations (gray) showed only few percept changes, $n < 20$, as well as large CVs (σ/μ , Fig 4B).

Intermittent presentation. In order to estimate the parameters in the two-state HMM in intermittent presentation, we use the Baum-Welch-Algorithm [BWA, 33, 34, 35]. See section ‘Intermittent presentation: Baum-Welch-Algorithm’ in the Materials and methods for details on the BWA, the choice of starting values and specifications in the sample data set. For details on HMMs see, e.g., [36, 37].

In order to investigate the estimation precision of this approach, we again apply parametric bootstrap to the 61 parameter combinations estimated from the response patterns to intermittent presentation in the sample data set. Fig 5 shows the median errors obtained in 1000 simulations for every parameter constellation for $T_1 = 1200$ s and $T_2 = 3600$ s. For p_{SS} and p_{UU} the absolute errors (i.e., $\text{AE}(p_{SS}) := |\hat{p}_{SS} - p_{SS}|$) are presented due to the small values of the two parameters. For the time horizon of the data, T_1 (panel A), 50 of the 61 parameter combinations yielded average errors (i.e. mean median errors across all parameters) smaller than 0.25 (black). The remaining cases (gray) showed a large μ_U , i.e., less distinguishable stable and

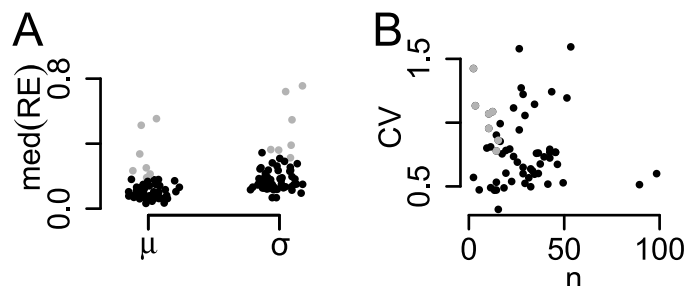


Fig 4. Precision of parameter estimates in the one-state HMM. For each of 61 parameter constellations in continuous presentation, 1000 simulations were performed with sample sizes as in the original data. (A) Median of the relative error (RE) for each parameter. (B) The CV as a function of the number n of simulated dominance times. Black points indicate constellations with mean RE across the parameters smaller than 0.25.

<https://doi.org/10.1371/journal.pcbi.1005856.g004>

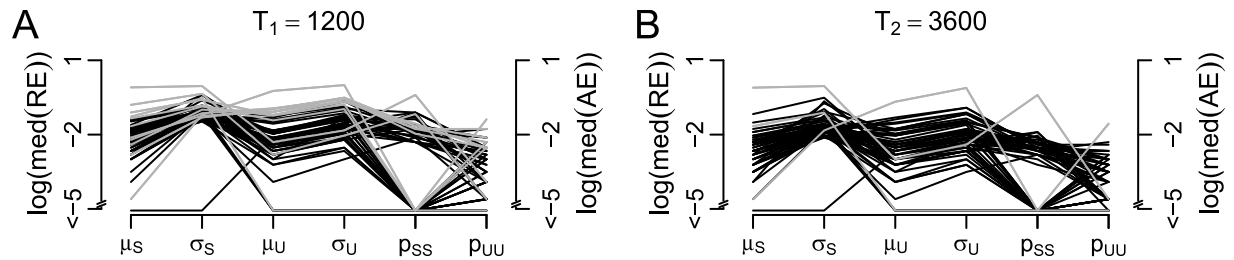


Fig 5. Precision of parameter estimates in the two-state HMM. For each of 61 parameter constellations in intermittent presentation, 1000 simulations were performed. Log(median) of RE (for $\mu_S, \sigma_S, \mu_U, \sigma_U$) or of AE (for ρ_{SS}, ρ_{UU}) for $T_1 = 1200$ (A) and $T_2 = 3600$ (B). Black lines indicate constellations with mean errors across the parameters < 0.25 .

<https://doi.org/10.1371/journal.pcbi.1005856.g005>

unstable distribution, or small sample size $n \approx 10$. For the larger time horizon T_2 (panel B), almost all parameter combinations showed errors smaller than 0.25.

Application of the HMM to the sample data set. Here we use the described methods in order to apply the HMM to the sample dataset presented in [10] consisting of responses to continuous and intermittent stimulation obtained from each of 29 patients with schizophrenia and 32 healthy controls (for details on experimental procedures see [Materials and methods](#)). While response patterns were highly variable across subjects, the CV of dominance times (mean 0.79, SEM 0.04) was comparable to other studies reported in [21]. Serial correlation of adjacent dominance times of the same percept was typically small (mean of Kendall's rank correlation $\bar{\tau} = 0.02$), and statistically significant on the 5% level in less than 7% of the cases, which is about chance level. Concerning long-term dependence [cmp. 31], deviations from the assumption of independent dominance times were not observed in 81% of the cases, and no differences were observed between the experimental groups ($p > .1$, Wilcoxon test, for details on the analysis see [Materials and methods](#)). A correlation between the alternation rates in continuous and intermittent stimulation across subjects was not observed in either group, comparable to the results of [14].

Model fit. By fitting the HMM to response patterns in continuous and intermittent presentation as described above, the typical properties of the observed response patterns can be reproduced in simulations, including unimodal distributions for continuous presentation and changes between stable and unstable stages in intermittent presentation and a high variety of response patterns (Fig 6). For example, subject C shows rather regular stable phases, separated by unstable phases, while subject D shows an irregular response pattern, subject E shows only stable phases, while subject F shows almost only unstable phases. The parameter estimates of these example subjects are given in Tables 1 and 2. Note also that the response patterns of seven of the 61 subjects were described better by only one (stable or unstable) distribution than by the two-state HMM.

In addition to the good correspondence in the response patterns, no strong deviations could be observed from the model assumption of Inverse Gaussian distributed dominance times (Fig 7).

Comparison of repeated trials. The HMM approach also allows studying the reproducibility of response parameters of subjects across multiple sessions. To this end, we used an additional dataset from 105 healthy individuals that contained two separate sessions of continuous presentation [see 9, here we used the first two training runs from Behavioral Experiment 2 as described in this previous work]. These showed highly reproducible response patterns, i.e., a high correlation of parameter estimates of the IG distribution of the same individuals across different sessions (Fig 8).

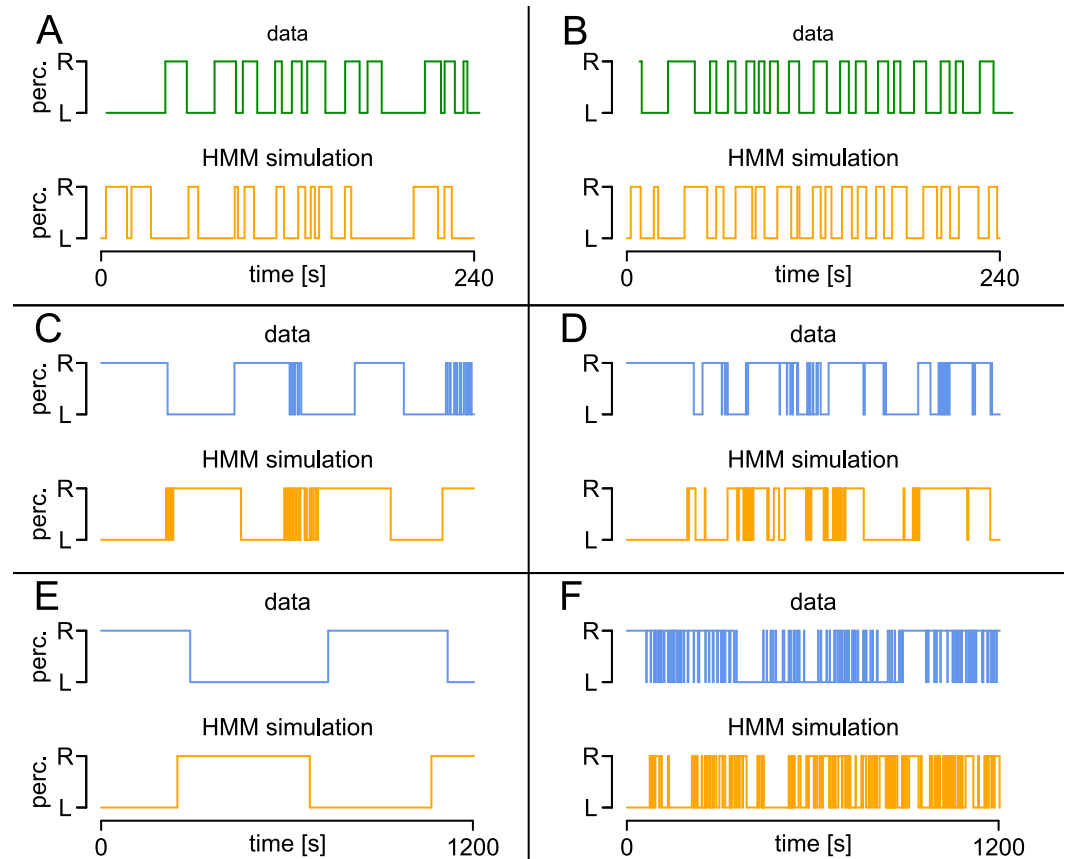


Fig 6. Comparison of empirical response patterns to patterns simulated with the HMM. Examples of response patterns to continuous (green, panels A and B) and intermittent (blue, C-F) stimulation repeated from Fig 1 and corresponding simulations within the HMM (orange). The parameter estimates are given in Tables 1 and 2, respectively.

<https://doi.org/10.1371/journal.pcbi.1005856.g006>

In addition, we also used a likelihood ratio test [38] to investigate the null hypothesis of equality of the parameters of two Inverse Gaussian distributed samples with sample sizes n_1 and n_2 , i.e. $H_0: \mu_1 = \mu_2$ and $\sigma_1 = \sigma_2$. The likelihood ratio derives as

$$Q_n = \prod_{i=1}^2 (n/n_i)^{n_i/2} (S_i/S)^{n_i/2},$$

with $S_i = \sum_{j=1}^{n_i} (d_{ij}^{-1} - \hat{\mu}_i^{-1})$ for $i = 1, 2$, $S_3 = n_1/\hat{\mu}_1 + n_2/\hat{\mu}_2 - n^2(n_1\hat{\mu}_1 + n_2\hat{\mu}_2)^{-1}$, $n = n_1 + n_2$ and $S = S_1 + S_2 + S_3$. Under H_0 the quantity $Q_n^* := -2(1 - 1/6[1/n_1 + 1/n_2] - 1/[12n]) \log Q_n$ is approximately chi-square distributed with two degrees of freedom. Thus, the test rejects the null hypothesis at level 5% if Q_n^* exceeds the 95%-th-quantile of the $\chi^2(2)$ -distribution. In the

Table 1. Estimated HMM parameters of the response patterns to continuous presentation shown in Fig 1A and 1B.

com.	$\hat{\mu}$	$\hat{\sigma}$
A	10.50	8.18
B	6.69	3.58

<https://doi.org/10.1371/journal.pcbi.1005856.t001>

Table 2. Estimated HMM parameters of the response patterns to intermittent presentation shown in Fig 1C–1F.

com.	$\hat{\mu}_s$	$\hat{\sigma}_s$	$\hat{\mu}_u$	$\hat{\sigma}_u$	\hat{p}_{ss}	\hat{p}_{uu}
C	186.45	30.50	5.01	3.06	0.67	0.96
D	67.12	50.81	3.25	2.76	0.33	0.82
E	372.28	68.30	NA	NA	1.00	NA
F	77.13	9.99	5.37	6.26	0.00	0.99

<https://doi.org/10.1371/journal.pcbi.1005856.t002>

sample data set, the likelihood ratio test did not reject the null hypothesis of equal parameter sets in 83 out of 105 subjects (about 79%). For a comparison, we performed 10000 permutations by randomly assigning a first trial of one subject to a second trial of another subject and performing the likelihood ratio tests on the permuted data sets. In the mean, the null hypothesis was not rejected in only about 36% of the randomly assigned pairs, with a maximum percentage across all permutations of 51%.

In summary, the response patterns of the same subject across multiple sessions showed a high degree of reproducibility, with a Pearson correlation coefficient of up to $r = 0.76$ between $\log(\mu_1)$ and $\log(\mu_2)$ (Fig 8). The similarity of response patterns for the same subject across multiple sessions was significantly higher than the similarity of response patterns between subjects.

Group differences. Finally, the HMM provides a relation between the underlying model parameters and the observed group differences reported in the introduction and in [10]. In the continuous case, the decreased alternation rate in the patients with schizophrenia is simply reflected in an increased mean dominance time $\hat{\mu}$ (Fig 9A) in the one-state HMM. For intermittent presentation, we had observed an increased alternation rate in the patients with schizophrenia. The interpretation of this observation was not obvious due to the high variability of response patterns and particularly due to the fluctuation between stable and unstable state. The HMM provides a first explanation of this phenomenon by capturing important response properties in the parameter estimates, which showed the following group differences: In particular, the expected relative time spent in the stable state,

$$\varphi_s := \frac{\mathbb{E}[\text{length of a stable phase}]}{\mathbb{E}[\text{length of a stable phase} + \text{length of an unstable phase}]}$$

was higher in the control group (Fig 9B, detailed formula given in Eq (13) in the Materials and

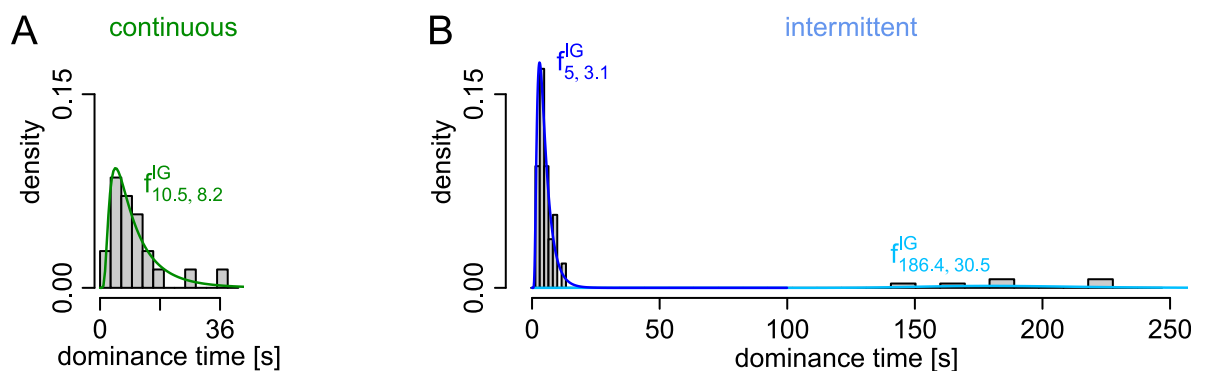


Fig 7. Comparison of distribution of dominance times with theoretical distribution. The theoretical IG distribution in the HMM fitted to the empirical distribution of dominance times for continuous (A) and intermittent (B) presentation shown in Fig 6A and 6C.

<https://doi.org/10.1371/journal.pcbi.1005856.g007>

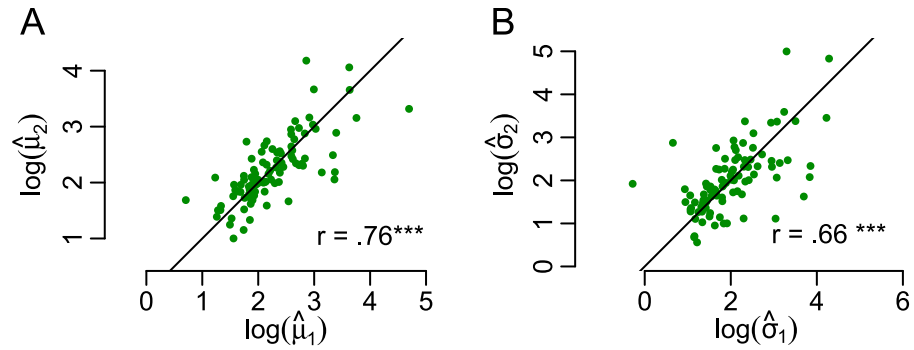


Fig 8. Reproducibility of response patterns. Parameter estimates of $\log(\mu)$ (A) and $\log(\sigma)$ (B) of the IG distribution in two sessions with the same individuals, dataset reported in [9]. The logarithm was applied due to asymmetric distributions of the parameter estimates. Stars indicate highly significant ($p < .0001$) correlation of parameter estimates across different sessions.

<https://doi.org/10.1371/journal.pcbi.1005856.g008>

methods). As the main variable contributing to this difference, we observe that the probability \hat{p}_{ss} to stay in the stable state was higher in healthy controls. In addition, the mean dominance time $\hat{\mu}_U$ in the unstable state was slightly larger in the patients with schizophrenia. The degree of statistical significance was highly similar to the one reported in [10] ($p < .1$ for $\hat{\mu}$, $\hat{\mu}_U$, \hat{p}_{ss} and $\hat{\phi}_s$, two-sided Wilcoxon test). In the following section, we present a model that links these observed differences to potential underlying neuronal mechanisms.

A hierarchical Brownian motion model

As described in the previous section, the HMM captures a high variety of response patterns both in continuous and intermittent viewing, including uni- and bimodal distributions of dominance times with alternations between stable and unstable states and a high variability across subjects. With its small number of parameters, the HMM can be fitted also to short data sections available empirically and therefore also capture differences between experimental groups.

However, the HMM description remains phenomenological and does not provide insight into potential neuronal processes. Also, it cannot provide explanations for potential effects

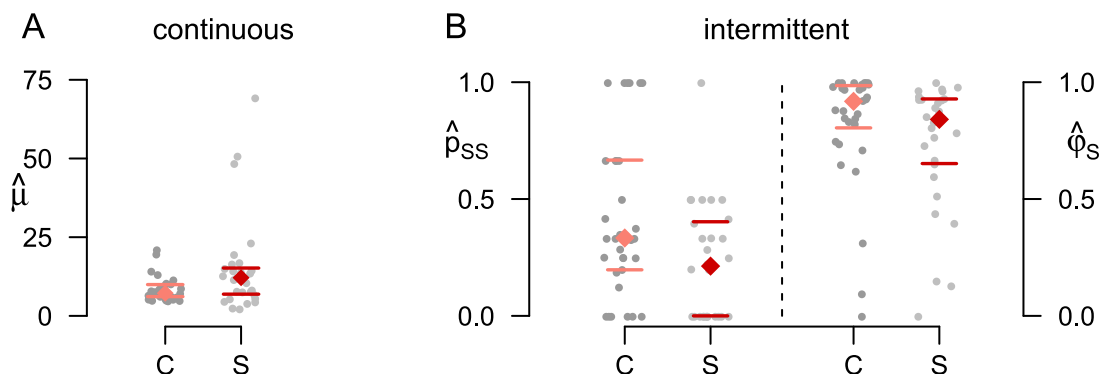


Fig 9. Differences in the HMM parameter estimates between subjects with schizophrenia and control subjects. (A) $\hat{\mu}$ during continuous presentation and \hat{p}_{ss} and $\hat{\phi}_s$ for intermittent presentation (B). Each grey dot indicates one individual participant's data, colored diamonds indicate medians, horizontal bars indicate 25%/75%-quantiles.

<https://doi.org/10.1371/journal.pcbi.1005856.g009>

that different lengths of blank displays could have on the response patterns, as discussed for example by [14, 22, 24, 25]. In addition, the HMM cannot represent the following interesting empirical observation: Before changing from stable to unstable state, the last dominance time tends to be shorter. Therefore, we introduce here a new model, called Hierarchical Brownian Model (HBM), which provides a potential link between the phenomenological description of the response and potential underlying neuronal processes. The HBM assumptions can also provide hypotheses on the effects of different lengths of blank displays and naturally yields shorter dominance times before a state change to the unstable state.

The HBM assumes two competing neuronal populations which indicate perception of right and left rotation, respectively. As has been proposed by various authors [14, 18], we implicitly assume mechanisms of self-excitation, cross-inhibition and adaptation across these neuronal populations, without explicitly modeling them in order to reduce the number of parameters and to allow for model fitting to short trials. In order to obtain a parsimonious model description, we again assume independence of dominance times by neglecting potential mechanisms of week long-term adaptation [31]. For possible model extensions compare section ‘Applicability and model extensions’ in the discussion. We use the simplified assumption that perception arises from the difference in the activity of the two populations, which is modeled here by a Brownian motion with drift [similar to 21] that fluctuates between two thresholds, where the first passage times indicate state changes. This results in two parameters for the case of continuous presentation that are directly linked to the two parametric distribution of dominance times in the HMM. Further, we describe switching between stable and unstable states in intermittent presentation by applying an analogous mechanism, which leads to a hierarchical model. We assume another hierarchical layer of neuronal populations and a corresponding Brownian motion which modulates the drift of the first population pair and thus causes switching between stable and unstable phases.

Continuous presentation. The HBM in continuous presentation (HBMc) simply assumes a Brownian motion with drift $\pm v_0$ between two borders, $\pm b$, where the first hitting times of the borders indicate a percept change and lead to a sign change in the drift. As a potential neurophysiological interpretation, b could be considered the size of the activity difference between the L and R population required for a perception change and thereby be related to the respective population sizes. Roughly speaking, the speed of the drift v_0 could be considered related to the inverse of the connection strengths within and across populations that engage in self excitation and cross inhibition.

Formally, let $b > 0$ be a fixed border, $v_0 > 0$ be a drift and $T > 0$ a time horizon, and let $(W_t)_{t \in [0, T]}$ be a standard Brownian motion. The perception process $P := (P_t)_{t \in [0, T]}$ is then defined by

$$dP_t = S_t v_0 dt + dW_t, \quad \text{where } P_0 = -b,$$

and the process $S_t := S(P_t, t)$ takes the value -1 if P_t last hit b and 1 if P_t last hit $-b$, with $S_0 := 1$. For $t \in (0, T]$ let

$$t^* := t^*(t) := \sup \{x | x < t, |P_x| = b\}, \quad \text{where } t^*(0) = 0$$

be the last time before t that P_t hit either b or $-b$. Then

$$S_t := S(P_t, t) := -\text{sgn}(P_{t^*}).$$

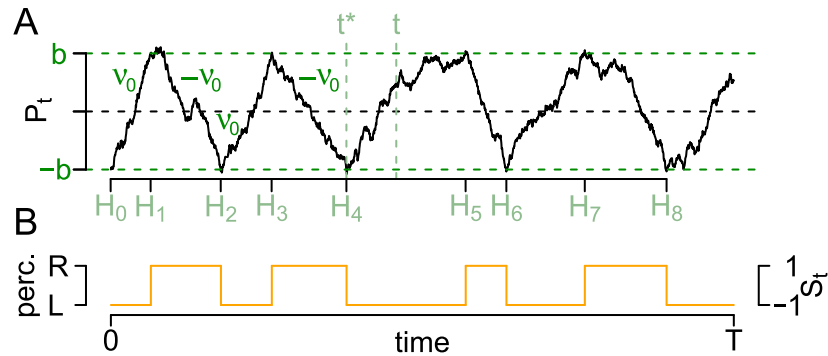


Fig 10. The HBMc. The first passage times (H_i) (green) of a Brownian motion P_t (black, panel A) with drift $\pm v_0$ at borders $\pm b$ indicate the times of the percept changes (orange, panel B). The Brownian motion is assumed to summarize the activity difference of two conflicting neuronal populations with only two parameters.

<https://doi.org/10.1371/journal.pcbi.1005856.g010>

The perception at time $t \geq 0$ takes the value L (left) if $S_t = 1$ and R (right) if $S_t = -1$ and switches at the first-hitting times $(H_i)_i$ of the borders $\pm b$ defined by $H_0 := 0$ and

$$H_i := \inf \{t \leq T | t > H_{i-1}, P_t = S_i b\} \quad i = 1, 2, \dots \quad (2)$$

An example of such a process is shown in Fig 10. Panel A shows the process P , where the sign of the drift changes at each first hitting time of b or $-b$ indicated by the process $(H_i)_i$, which also marks switches in the percept (panel B).

Parameter estimation in the HBMc makes use of the fact that the resulting dominance times

$$d_i := H_i - H_{i-1}, \quad i = 1, 2, \dots \quad (3)$$

are independent and IG distributed, with a known relation between the HBMc parameters b and v_0 and the parameters μ and σ of the IG distribution, given by $\mu = 2b/v_0$ and $\sigma = \sqrt{2b/v_0^3}$ [32]. Note that the CV of the dominance time distribution is therefore given by $1/(2bv_0)$. Thus, an increase in the border b also increases the mean dominance time and decreases the CV. An increase in the drift v_0 decreases the mean dominance time, while again decreasing the CV. The ML estimators \hat{b} and \hat{v}_0 can be derived via

$$\hat{b} = (1/2) \cdot \sqrt{\hat{\mu}^3/\hat{\sigma}^2} \quad \text{and} \quad \hat{v}_0 = \sqrt{\hat{\mu}/\hat{\sigma}^2},$$

where $\hat{\mu}$ and $\hat{\sigma}$ are derived from Eq (1). Explicit expressions are given in the Materials and methods.

The precision of parameter estimates is directly comparable to the HMM for continuous presentation.

Intermittent presentation. In the Hierarchical Brownian Motion model for intermittent presentation (HBMi), we require mechanisms for long dominance times in the stable state as well as for short dominance times in the unstable state. In order to describe the responses to intermittent and continuous presentation in one model framework, we assume the identical perceptual process as in the HBMc during phases of stimulus presentation. The periods of blank display represent the only difference to continuous presentation. In these periods, we assume additional neuronal mechanisms. In particular, we assume that the perceptual process then takes on one of two mean drifts, v_S in the stable state and $v_U \geq v_S$ in the unstable state, with potentially opposite signs of v_0 and v_S for increased stability (Fig 11). Note that the drifts

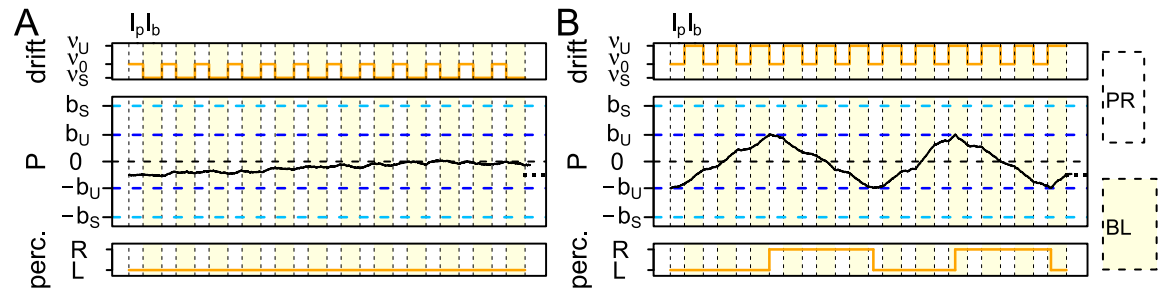


Fig 11. The perception process P in the HBMI during intermittent presentation. During presentation, P has drift v_0 . During blank displays (yellow), P has drift v_S in the stable phase (A), and drift v_U in the unstable phase (B). Typically, we have $v_S \leq v_0$ and $v_U \geq v_0$. The borders are b_S (light blue horizontal line) in the stable phase and b_U (blue horizontal line) in the unstable phase.

<https://doi.org/10.1371/journal.pcbi.1005856.g011>

v_S and v_U are not necessarily constant across the whole period of blank display, but they denote the mean drift of the process, which is sufficient to describe the distribution of dominance times. Interestingly, additional assumptions on the temporal behavior of the drift terms could also allow describing the impact of the lengths of blank displays (cmp. Discussion). Further, in the unstable state the border b_U at which perception and drift direction change is assumed smaller than the border b_S during stable perception. Switches between the stable and unstable state will be caused by a similar mechanism in a so-called background process B described later in this section.

Within a state (S or U), the fluctuation of the perception process between the borders is assumed analogous to the HBMc, except that the borders are dependent on the hidden state and that the drift is v_0 during presentation and v_S or v_U during blank display. Formally, we denote by PR and BL the sets of all periods of stimulus presentation and blank display, respectively. Assuming that we start a trial with a presentation interval and then switch regularly between presentation intervals of length l_p and blank display of length l_b , PR and BL are given by

$$\begin{aligned}
 PR &= \bigcup_{i=1}^{T/l_b} [(i-1)(l_p + l_b); (i-1)(l_p + l_b) + l_p) \\
 BL &= \bigcup_{i=1}^{T/l_b} [(i-1)(l_p + l_b) + l_p; i(l_p + l_b)) \quad (\text{Fig 11}).
 \end{aligned}$$

The perception process $(P_t)_t$ is then given by

$$dP_t = \begin{cases} S_t v_0 dt + dW_t, & \text{if } t \in PR \\ S_t v_{\tilde{Y}_t} dt + dW_t, & \text{if } t \in BL, \end{cases}$$

where $\tilde{Y}_t \in \{S, U\}$ denotes the hidden state and $(W_t)_t$ denotes a standard Brownian motion. As a result, the mean drift per second is given by

$$v_S^* := \frac{l_b \cdot v_S + l_p \cdot v_0}{l_b + l_p} \quad \text{and} \quad v_U^* := \frac{l_b \cdot v_U + l_p \cdot v_0}{l_b + l_p} \quad (4)$$

for states S and U , respectively. Because the periods l_b and l_p are typically short in relation to a dominance time, the behavior of P can be approximated by a Brownian motion with absolute drifts v_S^* and v_U^* , respectively. As in the HBMc, the sign of the drift $S_t := S(P_t, t)$ changes at

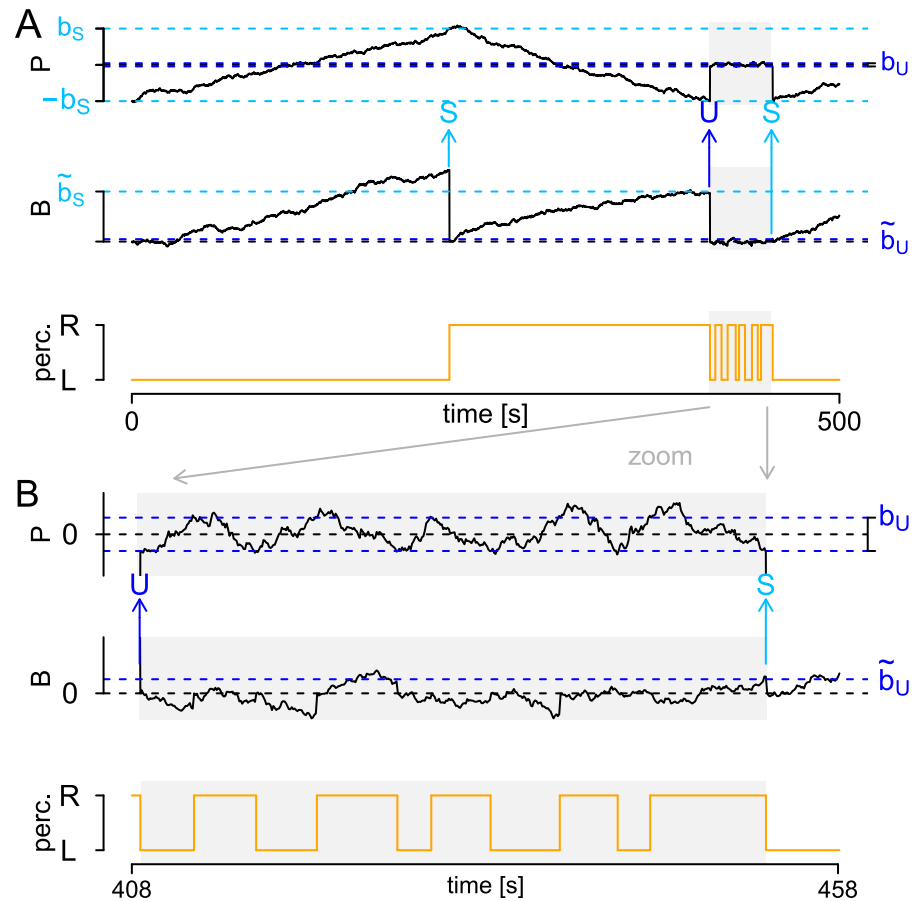


Fig 12. The HBMI. The perception process P , the background process B and the resulting percept. A: A simulation on $[0, 500]$. B: The same realization, zoomed in on the time interval $[408, 458]$. Stable phases indicated by white background, unstable phases indicated by gray background. The beginnings of stable and unstable phases are marked with light blue and blue arrows, respectively.

<https://doi.org/10.1371/journal.pcbi.1005856.g012>

every first hitting time of the respective border, i.e.,

$$S(P_t, t) := -\text{sgn}(P_{t^*}) \text{ where } t^* := t^*(t) := \sup \{x | x < t, |P_x| = b_{\tilde{Y}_x}\} \text{ with } t^*(0) = 0.$$

We initialize $P_0 = -b_{\tilde{Y}_0}$ for the initial state \tilde{Y}_0 which is the stable state with probability $\pi_S := \mathbb{P}(\tilde{Y}_0 = S)$. The perception then takes the value L if $S_t = 1$ and R if $S_t = -1$ and switches at the first-hitting times $(H_i)_i$ of the borders $\pm b_i$ comparable to Eq (2). Note that perception also changes during blank display. The dominance times are therefore again given by $d_i := H_i - H_{i-1}$, $i = 1, 2, \dots$

In order to describe the switching between the two states S and U , we use an analogous upper hierarchical level with another pair of conflicting neuronal populations. Their difference activity is described by a so-called background process $B := (B_t)_t$ (Fig 12A, middle panel). B is also assumed to be a Brownian motion with drift. Its drift is assumed to vanish during

presentation and to take the value $v_B > 0$ during blank display, i.e.,

$$dB_t = \begin{cases} d\tilde{W}_t, & \text{if } t \in \text{PR} \\ v_b dt + d\tilde{W}_t, & \text{if } t \in \text{BL}, \tilde{Y}_t = S \\ -v_b dt + d\tilde{W}_t, & \text{if } t \in \text{BL}, \tilde{Y}_t = U, \end{cases} \quad (5)$$

where $(\tilde{W}_t)_t$ is a Brownian motion independent of $(W_t)_t$. Again, the mean drift across PR and BL intervals is $v_B^* := \frac{v_b v_B}{b_b + v_p}$.

The background process B evokes changes between the stable and the unstable state. Specifically, at the time of a percept change t^* , the question of whether the process stays in the former state (S or U) or switches to the other state depends only on the value of B . Two borders, $b_U < b_S$ determine this switching as follows (see Fig 12). If the former state is S , the process remains stable if and only if $B_{t^*} > \tilde{b}_S$ (first light blue arrow in panel A), while switching to the unstable state if $B_{t^*} \leq \tilde{b}_S$ (blue arrow, panel A). Analogously, if the former state is U , the process switches to S if and only if $B_{t^*} > \tilde{b}_U$ (right light blue arrow, panel A), while staying in U if $B_{t^*} \leq \tilde{b}_U$ (blue arrow, panel B). After the percept change, the background process B is reset to zero and then follows its usual dynamic (Eq (5)), i.e., the sign of its drift changes if and only if the state has changed. Finally, as the perception process P fluctuates between $\pm b_S$ in the stable state and between $\pm b_U$ in the unstable state, the value of P is reset when the state changes, to the value $\text{sgn}(P)b_S$ when changing to the stable state and to $\text{sgn}(P)b_U$ when changing to the unstable state.

Discussion of assumptions and interpretation of parameters. The technical advantage of the HBMI is that the resulting dominance times agree in most parts with the dominance times resulting from the HMM assumptions, which allows model fitting also to short data sections and comparison across clinical groups. In addition, the HBMI also provides a relation to potential underlying neuronal processes, as discussed in the following and illustrated in Fig 13.

Both HBMI-processes P and B are assumed Brownian motions with drift which may be interpreted as the activity difference between neuronal populations. Implicitly, this assumes mechanisms of self-excitation, cross-inhibition and adaptation across these neuronal populations, as proposed by various authors [14, 18]. Without explicitly modeling such mechanisms in order to reduce the number of parameters and allow model fitting, the parameter sets are reduced to the mean drifts v and the borders b . Analogously to the HBMC, the speed of the drifts could be considered related to the inverse of the connection strengths within and across populations that engage in self excitation and cross inhibition. The border b , in analogy to the HBMC, could be considered related to the size of the respective populations under consideration. The use of different borders allows fitting of highly various response patterns and can be motivated as follows.

In the HBMI, the perception process P has two borders, $b_S > b_U$ for the stable and the unstable state. This suggests different population sizes of neurons involved in the stable and unstable state. Typically $b_S > b > b_U$, suggesting that in the stable state, the activity of the dominant population is increased by joining additional neurons to the population, for example by positive feedback mediated by population S . Vice versa, in the unstable state, only a minimal population is involved in the respective percept, leading to fast changes. Thus, one could assume that the dominant percept population size is decreased by the population U (red arrows). The active population sizes are indicated by different circle sizes in the first line of Fig 13 and are assumed modulated by the background populations S and U .

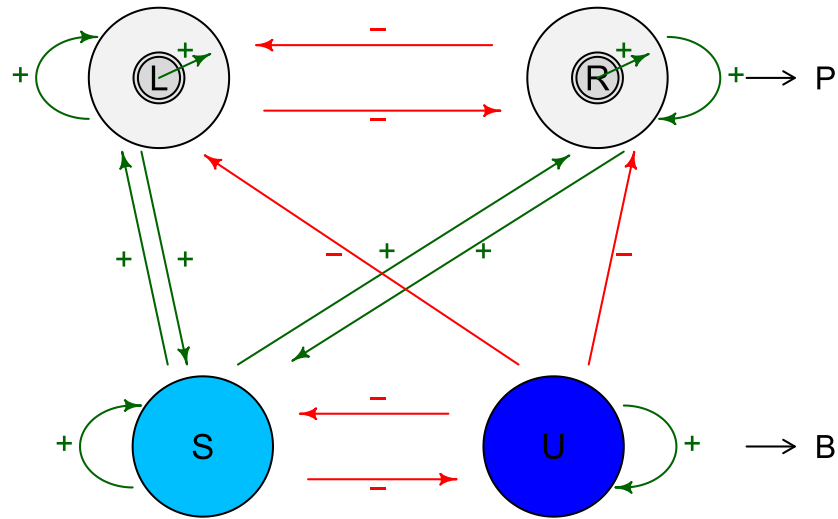


Fig 13. Motivation of HBMI model assumptions. For explanation and details see text.

<https://doi.org/10.1371/journal.pcbi.1005856.g013>

The background process B models the activity difference between S and U and is also associated with two borders, \tilde{b}_S and \tilde{b}_U . The assumption regarding resetting of B at percept change is technically necessary to generate independent dominance times and to thus allow straightforward model fitting (cmp. section ‘Parameter estimation’). In the picture of Fig 13 it can be motivated as follows. Population S is capable of offering positive feedback to the currently active population, L or R , which results in an increased population size as described above. S is also activated by the active population. Therefore, a percept change causes a resetting to zero. However, if S had shown high previous activation (above \tilde{b}_S), the activity of S can increase rapidly again, causing another stable dominance time. In contrast, in case of weak previous activation (below \tilde{b}_S), the unstable population U is taking over, marking the transition to an unstable state. With opposite signs, i.e., negative drift and a small new border \tilde{b}_U , the process proceeds analogously. Similar to the mean drift terms v_S and v_U , the drift v_B is not necessarily constant but describes the mean drift of B during the period of blank display.

In addition to the potential neurophysiological interpretations of the model parameters, we give here a relation of the parameters to the response patterns. Interestingly, the seven HBMI parameters allow the reproduction of highly variable response patterns as are also observed in the empirical data sets (e.g., Fig 1). The following quantities, which are easily derived from the parameters, offer a straightforward pattern interpretation.

First, the parameter sets (b_S, v_S^*) and (b_U, v_U^*) can be interpreted analogously to the parameters (b, v_0) in continuous stimulation. That means, an increase in the border (b_S or b_U) increases the mean dominance time and decreases the CV in the respective state. An increase in the drift (v_S^* or v_U^*) decreases the mean dominance time, while also decreasing the CV. Recall that the CVs of dominance times during stable and unstable states are given by

$$CV_S^* := 1/\sqrt{2b_S v_S^*} \quad \text{and} \quad CV_U^* := 1/\sqrt{2b_U v_U^*}, \quad \text{respectively.}$$

Fig 14 illustrates examples with small CV_S^* (panels A-D) and large CV_S^* (panels E-H).

Second, the parameters \tilde{b}_S and v_B^* can be interpreted best when compared to b_S and v_S^* as follows. Consider the expected fraction of \tilde{b}_S reached by the background process at the end of a

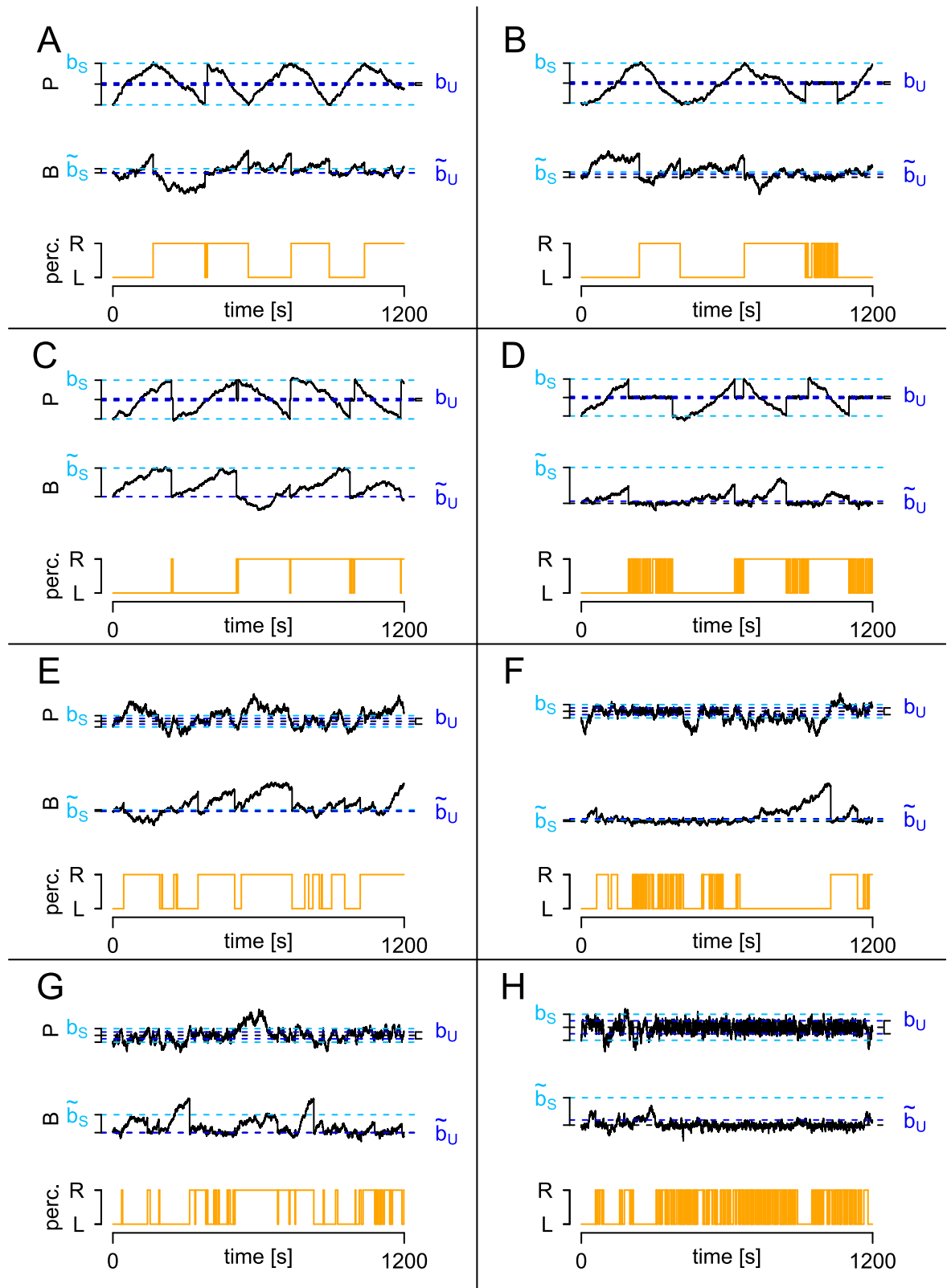


Fig 14. Impact of HBMI parameter values on the response patterns. Examples of simulated response patterns are shown for different values of the three quantities CV_s^* , $2b_s v_B^* / \tilde{b}_s v_s^*$ and \tilde{b}_U . The quantities for panels A-H were $CV_s^* = \{0.2, 0.2, 0.2, 0.2, 1, 1, 1, 1\}$, $2b_s v_B^* / \tilde{b}_s v_s^* = \{4, 4, 0.4, 0.4, 4, 4, 0.4, 0.4\}$, $\tilde{b}_U = \{0, 3, 0, 3, 0, 3, 0, 3\}$.

<https://doi.org/10.1371/journal.pcbi.1005856.g014>

stable dominance time,

$$\frac{\text{Expected duration of a stable dominance time}}{\text{Expected duration until } B \text{ reaches } \tilde{b}_s} = \frac{2b_s/v_s^*}{\tilde{b}_s/v_B^*} = \frac{2b_s v_B^*}{\tilde{b}_s v_s^*},$$

which is related to the transition probability from stable to unstable state. In case of a small background border $\tilde{b}_s < b_s$ and small v_s^* , the probability of B crossing \tilde{b}_s until percept change is high, such that the process remains stable. Fig 14A, 14B, 14E and 14F show such parameter combinations. An analogous term can be derived in comparison to the parameters \tilde{b}_U and v_U^* .

Third, the parameter \tilde{b}_U is related to the number of dominance times in the unstable state before changing to the stable state. Recall that the drift of B is negative during unstable periods. Therefore, a large value of \tilde{b}_U implies a low probability to reach \tilde{b}_U until the percept change. This implies a high expected number of dominance times in the unstable state, or a low transition probability from the unstable to the stable state. Fig 14B, 14D, 14F and 14H show examples with large values of \tilde{b}_U .

Relation of the HBMc to the two state HMM. The relation of the HBMc to the one state HMM is simple as it represents only a reparameterization. Both the one state HMM and the HBMc yield independent and IG distributed dominance times. For the intermittent case, the relation of the HBMc to the two state HMM is not as straightforward. The two models are highly similar in the sense that they use two parameters to describe long and short dominance times, respectively (e.g., (μ_S, σ_S) and (b_s, v_s^*) for the stable state). In the HMM, the dominance times are IG distributed, given the state with the respective parameters. In the HBMc, the dominance times are approximately IG distributed, where the minor deviation from the IG distribution originates from the minor deviation of P from a Brownian motion with drift v_s^* (or v_U^*), instead of exactly assuming drift v_0 during stimulation and v_s (or v_U) during blank display. However, the marginal distribution of P at multiples of such intervals $l_p + l_b$ is identical to the marginal distribution of a Brownian motion with drift v_s^* (or v_U^*) at these time points, and the differences can only be observed in the meantime. Because dominance times usually span multiple trials of duration $l_p + l_b$, the approximation is very close. As another similarity, both models use additional parameters $((p_{SS}, p_{UU})$ and $(\tilde{b}_s, \tilde{b}_U, v_B)$) to describe the transition probabilities between the stable and the unstable state.

The main difference between the HBMc and the two state HMM concerns the dynamic of the state transitions between stable and unstable state. In the HMM, transition probabilities are given by $(1 - p_{SS})$ and $(1 - p_{UU})$ and are independent of the duration of the previous dominance time. In contrast, in the HBMc, a transition from stable to unstable state requires that B has not reached \tilde{b}_s at the end of the respective dominance time. Therefore, the transition probability $\tilde{p}_{SU}(d_i)$ depends on the duration d_i of the i -th dominance time, where shorter dominance times yield higher transition probabilities. Note that the position of B at the end of a dominance time d_i is given by an increment of a Brownian motion with drift v_B^* in the fixed time interval d_i . Therefore the position is normally distributed with mean $d_i \cdot v_B^*$ and variance d_i and the probability to remain in the stable state (which is the probability that the background process exceeds \tilde{b}_s) is given by

$$\tilde{p}_{SS}(d_i) = \mathbb{P}(Y_{i+1} = S | Y_i = S, d_i) \approx 1 - \Phi_{v_B^* d_i, d_i}(\tilde{b}_s), \tag{6}$$

where $\Phi_{\mu, \sigma^2}(\cdot)$ denotes the distribution function of the normal distribution with mean μ and variance σ^2 and Y_i is the hidden state of the i -th dominance time. Analogously, for the transition from unstable to stable state, the transition probability $\tilde{p}_{US}(d_i)$ tends to decrease with

longer dominance times, and we find

$$\tilde{p}_{UU}(d_i) = \mathbb{P}(Y_{i+1} = U | Y_i = U, d_i) \approx \Phi_{-v_B^* d_i, d_i}(\tilde{b}_U). \tag{7}$$

Note that we use the approximate sign ‘ \approx ’ because the drift of B is not exactly v_B^* throughout, but is assumed to change between v_0 and v_B during stimulation and blank display, respectively, yielding a mean drift of v_B^* . Analogously to the above explanation, differences caused by the approximation can be considered minimal.

In order to obtain quantities comparable to the transition probabilities p_{SS} and p_{UU} in the HMM, we derive the marginal transition probabilities in the HBMI as the expected value of \tilde{p}_{SS} and \tilde{p}_{UU} . As shown by [39], the positions X_S and X_U of B at the end of an independent stable or unstable IG distributed dominance time follow the Normal Inverse Gauss (NIG) distribution. The resulting transition probabilities in the HBMI can then be calculated as

$$p_{SS}^* = \mathbb{P}(X_S > \tilde{b}_S) \quad \text{and} \quad p_{UU}^* = \mathbb{P}(X_U \leq \tilde{b}_U). \tag{8}$$

where X_S is NIG-distributed with parameters $(0, \sqrt{v_S^{*2} + v_B^{*2}}, v_B^*, 2b_S)$ and X_U is NIG-distributed with parameters $(0, \sqrt{v_U^{*2} + v_B^{*2}}, -v_B^*, 2b_U)$.

One should note that due to the difference in transition probabilities of the two models, the parameters (b_S, v_S^*) are not direct reparameterizations of (μ_S, σ_S) (and similarly for the unstable state). Furthermore, the dependence of the transition probability on the length of the previous dominance time is one important new aspect of the HBMI not described in the HMM, which will also be used in the section ‘Application of the HBM to the sample data set’ for comparison of models and empirical observations.

Parameter estimation. In order to estimate the HBMI parameter set

$\Theta = (b_S, v_S^*, b_U, v_U^*, \tilde{b}_S, \tilde{b}_U, v_B^*)$, we use ML estimation. The likelihood L is given by

$$L(d_1, \dots, d_n | \Theta) \approx \alpha_S(n) + \alpha_U(n), \tag{9}$$

where the forward variable $\alpha_j(i)$ denotes the probability of being in state j at time i while observing (d_1, \dots, d_i) . The approximation is again due to the averaging of drifts during blank display and stimulus presentation. The forward variables can be derived recursively [34, 40] by

$$\alpha_j(i) = \begin{cases} \pi_j f_j(d_i) & \text{if } i = 1 \\ f_j(d_i) \sum_{j \in \{S,U\}} \alpha_j(i-1) \tilde{p}_{ij}(d_{i-1}) & \text{if } i > 1 \end{cases} \tag{10}$$

with π_j as starting distribution and f_S and f_U denoting the densities of $\text{IG}(2b_S/v_S^*, \sqrt{2b_S/v_S^{*3}})$ and $\text{IG}(2b_U/v_U^*, \sqrt{2b_U/v_U^{*3}})$ distributed random variables, respectively. Details on the maximization algorithm can be found in the [Materials and methods](#). After estimation of Θ , the estimates of \hat{v}_S and \hat{v}_U can be obtained using [Eq \(4\)](#), and analogously for \hat{v}_B .

Precision of parameter estimation. The variability of the parameter estimates in the HBMI is studied analogously to the HMM, using the RE and the AE of the parameter estimates obtained in 1000 simulations of the 61 parameter combinations estimated from the empirical data set. For the border parameters $b_S, b_U, \tilde{b}_S, \tilde{b}_U$, we use the RE, while for the typically small drift parameters v_S^*, v_U^*, v_B^* , the AE is used. Again, one set of simulations uses the time horizon of the empirical data, $T_1 = 1200\text{s}$ ([Fig 15A](#)), and a second simulation was performed using $T_2 = 3600\text{s}$ ([Fig 15B](#)). According to the simulation results, the precision of parameter estimation in the given parameter range is not always satisfactory for the given parameterization $(b_S, b_U, \tilde{b}_S, \tilde{b}_U, v_S^*, v_U^*, v_B^*)$, yielding average errors (i.e., mean REs or AEs across all variables)

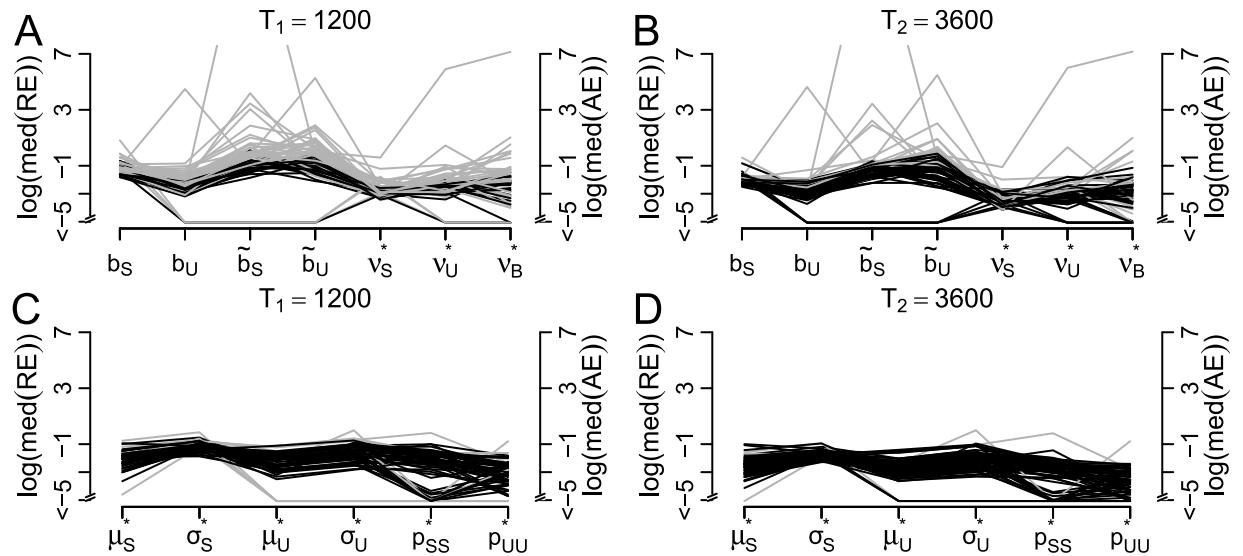


Fig 15. Precision of parameter estimates in the HBMI. For each of the 61 parameter constellations estimated from the sample data set, 1000 simulations were performed with the HBMI. (A) and (B): log(median(REs)) of the border parameters b_s, b_U, b_s^*, b_U^* and log(median(AEs)) of the drift parameters v_s^*, v_U^*, v_B^* with $T = 1200s$ (A) and $T = 3600s$ (B). (C) and (D) log(median(REs)) for the derived model parameters $\mu_s^*, \sigma_s^*, \mu_U^*, \sigma_U^*$ and log(median(AEs)) for the parameters p_{SS}^* and p_{UU}^* for $T = 1200s$ (C) and $T = 3600s$ (D). Parameter combinations with mean errors < 0.25 across all variables plotted in black.

<https://doi.org/10.1371/journal.pcbi.1005856.g015>

smaller than 0.25 in only 24 out of 61 cases for the empirical time horizon T_1 and still in only 44 out of 61 for the tripled sample size of T_2 . This suggests that these raw parameters yield less reliable estimates because different combinations of b and v can yield the same mean stable dominance time. In contrast, the following set of derived parameters that are more easily comparable to the HMM parameters shows better properties. We denote by (μ_s^*, σ_s^*) and (μ_U^*, σ_U^*) the mean and standard deviation of dominance times in the HBMI in the stable and unstable state, respectively. These have analogous interpretations as (μ_s, σ_s) and (μ_U, σ_U) in the HMM and are given by

$$\mu_s^* \approx 2b_s/v_s^* \quad \text{and} \quad \sigma_s^* \approx \sqrt{2b_s/v_s^{*3}}, \quad (11)$$

and analogously for μ_U^* and σ_U^* , where the approximation is again due to the minimal difference between the mean drift v_s^* and the changing drift $v_s + v_0$ during presentation and blank display. In addition, we consider the transition probabilities between the states, p_{SS}^* and p_{UU}^* (cmp. Eq (8)). Fig 15C and 15D show the median REs for $\mu_s^*, \sigma_s^*, \mu_U^*, \sigma_U^*$ and median AEs for p_{SS}^*, p_{UU}^* obtained in the 1000 simulations of length T_1 (panel C) and T_2 (panel D). Concerning this parameterization, 51 parameter combinations yielded average errors smaller than 0.25 across all variables for T_1 , while as many as 58 out of 61 cases showed errors smaller than 0.25 for T_2 . These simulations suggest that the parameterization $(\mu_s^*, \sigma_s^*, \mu_U^*, \sigma_U^*, p_{SS}^*, p_{UU}^*)$ yields more reliable parameter estimates in the HBMI than the original parameterization of borders and drifts.

Application of the HBM to the sample data set. For the analysis of the sample data set, the HBMc and HBMI were fitted to the dataset of [10] containing responses to continuous and intermittent stimulation for each of 29 patients with schizophrenia and 32 control subjects, using the parameter estimation described in the section ‘Parameter estimation’.

Model fit. Because the HBMc represents only a reparameterization of the one-state HMM, results are completely analogous for continuous presentation. Thus, a high variability of

response patterns can be described with the two parametric distribution (see Fig 16A and 16B), including also different means and variances of dominance times. For intermittent presentation, the HBMI and the two-state HMM are similar, but also show a number of differences. As a first similarity to the two-state HMM, the HBMI can also describe and reproduce a high variety of response patterns (Fig 16C–16F). For example, these include highly regular stable states that may or may not be interrupted by short unstable phases (C,E) or response patterns with different degrees of regularity (D, F) and different alternation rates. Note also that the response patterns to intermittent stimulation of six out of the 61 subjects were described better by the one-parametric HBMc (e.g., E).

In addition to the close description and reproduction of the patterns in the empirical data, one interesting additional aspect is captured by the HBMI, which cannot be described in the two-state HMM. As explained in the section ‘Relation of the HBMI to the two state HMM’, the probability of a transition from stable to unstable state decreases with the length of the dominance time in the HBMI. The same is true for the reverse transition. Indeed, the same observation can be made in the empirical data set, while this dependence cannot be captured within the HMM (Fig 17).

Group differences. Due to the high correspondence of the HBMI response patterns with the empirical data and the neurophysiologically related parameterization, the HBMI provides potential additional links to underlying neuronal processes of the observed differences between control subjects and patients with schizophrenia. Here, we consider three aspects related to continuous and intermittent stimulation and to the transition between these two conditions.

First, concerning continuous stimulation, we note that the one-state-HMM and the HBMc yield identically distributed sequences of dominance times. Note that the mean dominance time $\hat{\mu}$ in the HMM (Fig 9A) therefore equals the corresponding value $2\hat{b}/\hat{v}_0$ in the HBMc. As a consequence, the results and interpretation were identical, i.e., the higher alternation rate of the control subjects during continuous presentation (Fig 2) was reflected in a smaller value of $2\hat{b}/\hat{v}_0$. When analyzing the individual parameters b and v_0 , we found no group differences in the drift v_0 , but a tendency for a larger neuronal pool b involved in sensory processing in the group of patients with schizophrenia (Fig 18A, $p < .1$, two-sided Wilcoxon test).

Second, regarding intermittent presentation, we focussed on the derived parameters $(\mu_s^*, \sigma_s^*, \mu_u^*, \sigma_u^*, p_{ss}^*, p_{uu}^*)$, which show correspondence to the HMM and good estimation properties, instead of testing the border and drift parameters individually. Similar to the HMM, we found an increased mean unstable dominance time in the group of patients with schizophrenia as compared to healthy controls. More importantly and also consistent with the HMM, we also found that patients with schizophrenia showed a decreased relative time spent in S , φ_s^* ($p < .1$, two-sided Wilcoxon test, see Eq (17) in the Materials and methods). Again, the probability p_{ss}^* to stay in the stable state seemed to be the main variable contributing to this difference, being significantly reduced in patients with schizophrenia (Fig 18C, $p < .1$, two-sided Wilcoxon test). In order to identify potential neurophysiological mechanisms underlying this group difference, we further investigated the components of p_{ss}^* . Noting that no difference was observed in the mean stable dominance time $\mu_s^* = 2b_s/v_s^*$ and in the relation v_s^*/v_b^* of the drift in stable state and the drift of the background process, one interesting parameter is b_s/\tilde{b}_s . Keeping all other parameters constant, an increase in this quantity means an increase in p_{ss}^* because the crossing of \tilde{b}_s at the end of a stable dominance time, i.e., staying in the stable state, gets more likely. In the empirical data set, we found increased values of b_s/\tilde{b}_s in the control group (Fig 18C, $p < .1$, two-sided Wilcoxon test). In terms of the potential neurophysiological

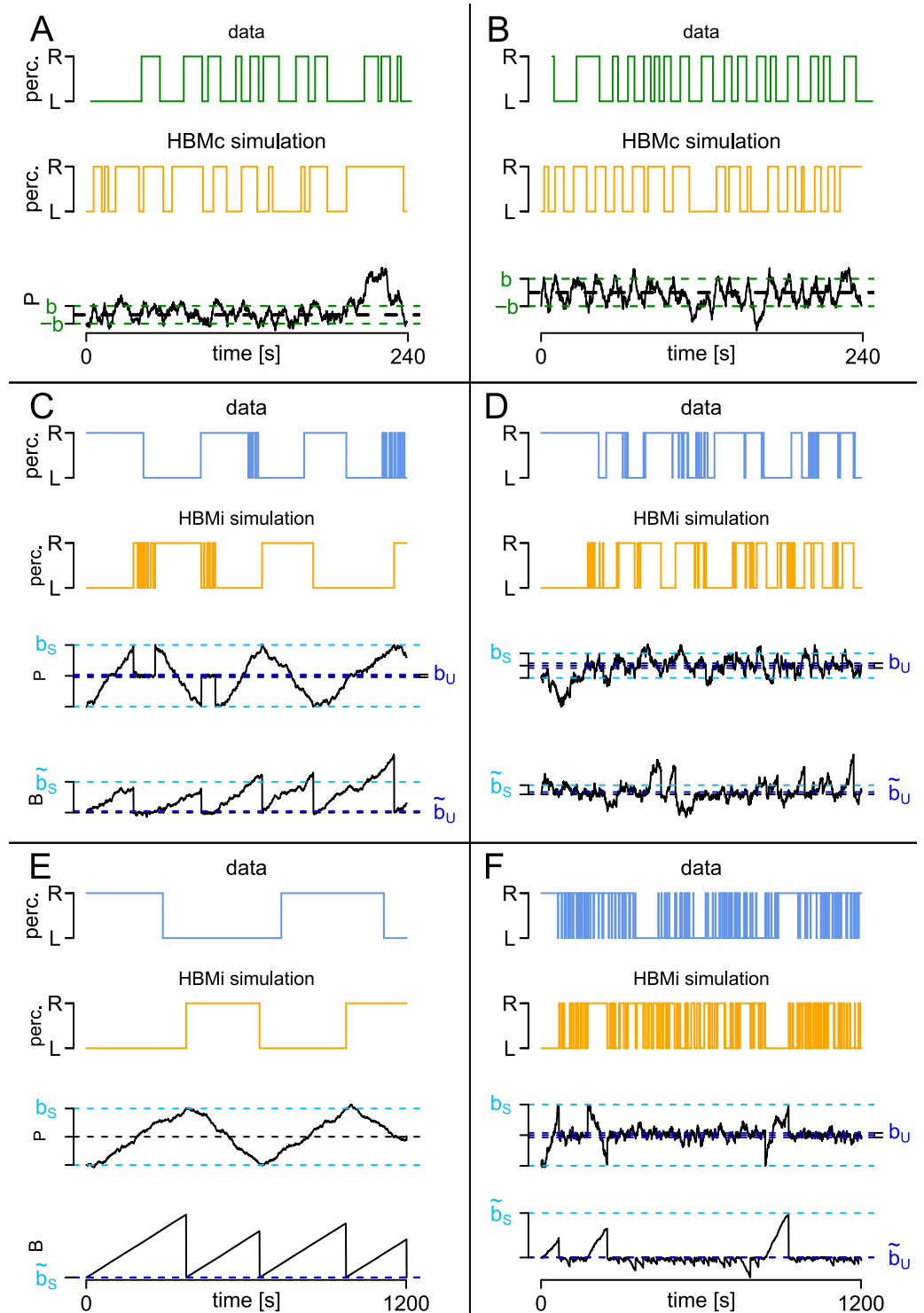


Fig 16. HBM: Examples of simulated response patterns. Example empirical response patterns from Fig 1 and response patterns simulated by the HBM (orange). Responses to continuous and intermittent presentation are plotted in green and blue, respectively. The estimated parameters used for simulation are given in Tables 3 and 4, respectively. (A) and (B) Continuous presentation and HBMc. (C)-(F): Intermittent presentation and HBMi. Below the orange response patterns, one can see the perception process P and (for the HBMi) the background process B corresponding to the respective simulation.

<https://doi.org/10.1371/journal.pcbi.1005856.g016>

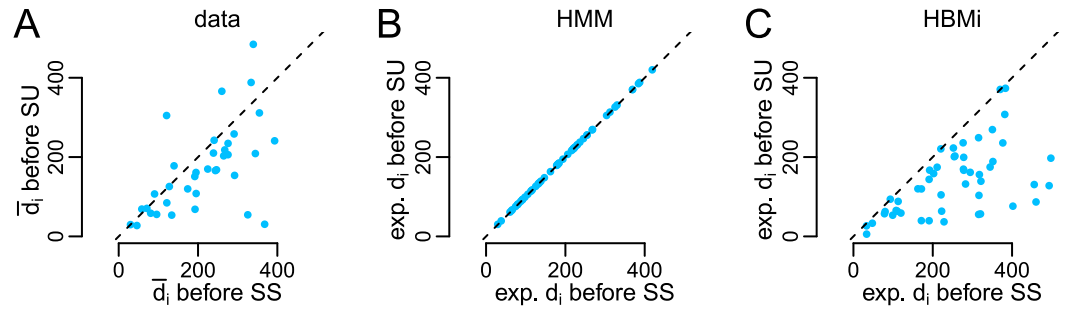


Fig 17. Mean dominance times in stable state as a function of the successive state. In the empirical data set, the mean dominance time before a state transition SU is shorter than the mean dominance time before SS (A). This observation can be reproduced in the HBMi (C), while the expected dominance time in the HMM (B) is equally long for intervals with and without transition, i.e., independent from transition between states. In the empirical data set [10], the state transitions were estimated using the Viterbi paths [34]. Analogous results were obtained using a fixed threshold. For the HBMi, the expected dominance times were derived from the HBMi parameters according to Eq (15) in the Materials and methods. In the HMM, expected dominance times correspond to $\hat{\mu}_S$.

<https://doi.org/10.1371/journal.pcbi.1005856.g017>

interpretation, this would suggest an increased population size involved in stable perception processing as a potential main underlying mechanism.

Third, a similar observation resulted from the derived parameter b_S/b , which describes a transition of population sizes from continuous to intermittent stimulus processing. In the data set, we observed increased values of b_S/b for the control group (Fig 18B) as compared to the group of patients with schizophrenia. Again, applying a potential neurophysiological interpretation, this would suggest that in the control group, the neuronal pool that is additionally recruited during intermittent presentation in stable states could be higher than in the group of patients with schizophrenia. Note that this result could not be obtained in the HMM, which does not explicitly describe a potential transition between continuous and intermittent presentation. However, note that this result should be interpreted carefully due to reduced estimation precision of the parameters b_S/\tilde{b}_S and b_S/b .

In summary, this analysis of the HBM model parameters suggests the following explanation for the observed phenomenon that the alternation rate of the perceived percepts is increased for patients with schizophrenia during intermittent stimulation, while being decreased during continuous stimulation. In general, the relative time spent in the unstable state was increased in patients with schizophrenia. According to the HBM, this was attributed to an increased probability of transition from the stable to the unstable state, which could be potentially related to a decreased recruitment of neurons in the stable state. More specifically, a larger neuronal pool is hypothesized to account for the increased stability, and we accordingly observe a smaller increase in the neuronal pool from continuous to intermittent stable presentation in the patients with schizophrenia, which is suggested by the HBM as a main mechanism for the observed group differences. This analysis of a potential underlying mechanism, explaining the observed group differences also in the transition from continuous to intermittent presentation,

Table 3. Estimated HBMi parameter combinations of the typical response patterns to continuous presentation shown in Fig 1A and 1B.

com.	\hat{b}	\hat{v}_0
A	2.08	0.40
B	2.42	0.72

<https://doi.org/10.1371/journal.pcbi.1005856.t003>

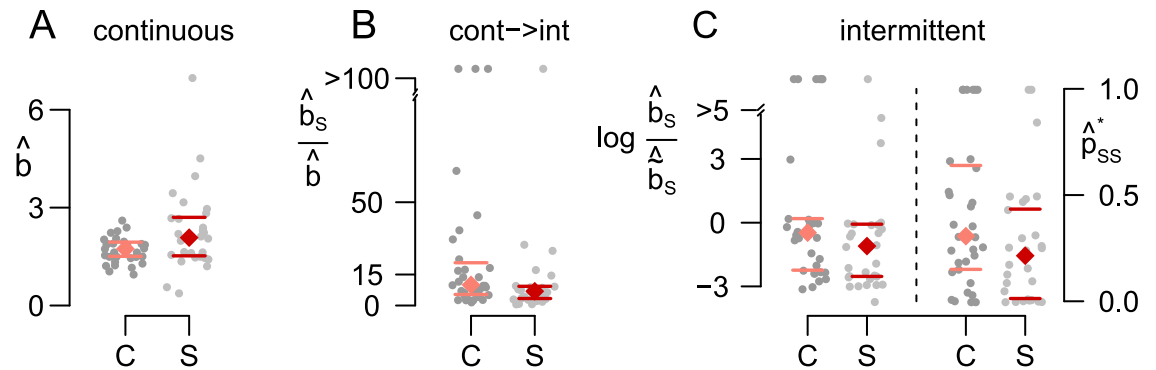


Fig 18. Group differences between patients with schizophrenia and control subjects in the HBM. The HBMc parameter \hat{b} (A) and the relation \hat{b}/\hat{b}_s (B). (C): Differences in \hat{b}_s/\hat{b} and \hat{p}_{ss}^* . The raw data together with the median (colored diamonds) and 25%- and 75%-quantiles are shown. All p -values of a two-sided Wilcoxon test were below .1.

<https://doi.org/10.1371/journal.pcbi.1005856.g018>

Table 4. Estimated HBMi parameter combinations of the typical response patterns to intermittent presentation shown in Fig 1C–1F.

com.	\hat{v}_0	\hat{b}_s	\hat{v}_s	\hat{b}_U	\hat{v}_U	\hat{b}_s	\hat{b}_U	\hat{v}_B
C	0.72	41.74	0.24	1.84	0.74	49.71	2.39	0.54
D	0.27	5.42	0.08	1.06	0.93	4.56	1.13	0.06
E	0.32	52.58	0.25	NA	NA	0.00	NA	17.5
F	0.57	13.5	0.24	1.00	0.24	110.28	0.77	1.90

<https://doi.org/10.1371/journal.pcbi.1005856.t004>

is a particular advantage of the HBM over the HMM, because the HBM provides mechanistic explanations and variables with potential neurophysiological interpretations.

Discussion

Summary and implications

In the present article we have proposed a model framework for the description and analysis of perceptual responses to bistable stimuli. In particular, the first goal was to describe a high number of observed patterns in responses to continuous and intermittent stimulation and their differences between a group of patients with schizophrenia and healthy controls. The variety of patterns includes more or less regular dominance times during continuous stimulation and a switching between long and short dominance times, i.e., stable and unstable states, during intermittent stimulation, with a tendency for periodically occurring percept changes.

We started on a descriptive level, assuming that dominance times were generated by a simple HMM with only one state for continuous presentation and a stable and an unstable state in intermittent presentation. The HMM was sufficiently small to allow model fit to short empirical data sets and could also describe the high variety of empirically observed response patterns in continuous and intermittent presentation. Interestingly, it also revealed a high degree of reproducibility of response patterns of the same subject across different sessions. In addition, it allowed to relate observed group differences in the rate of percept alternations to HMM parameters, suggesting that especially the relative time spent in the stable state was reduced in the patients with schizophrenia.

Our second goal was to relate the observed response patterns and group differences to potential underlying mechanisms and thus, to build a link to models with detailed

neurophysiological assumptions [7, 13, 14, 18, 20] that may not include all types or response patterns and/or may not allow fitting to short data sets. To that end, we proposed a hierarchical model of interacting Brownian motions (HBM). The HBM is based on the common assumption that the sequence of percept changes results from a competition of conflicting neuronal populations [14, 17, 18, 20, 27]. Instead of modeling these in detail, we describe the activity difference by a Brownian motion P with drift v_0 [21] between two borders $\pm b$, where the first hitting times of the borders indicate percept changes. Roughly speaking, the drift v_0 could be considered related to the neuronal interactions within and between the populations, while the border b could be considered related to the population sizes. In order to describe responses to intermittent presentation, this mechanism is adapted in another population pair. These populations exhibit a corresponding background process B that evokes switching between stable and unstable states, similar to switching between the two percepts. In particular, B causes the perception process P to change parameters from small drift v_S and large border b_S in the stable state to fast drift v_U and small border b_U in the unstable state.

The HBM could be fitted nicely to the given empirical data set, reproducing a high variety of response patterns to continuous and intermittent stimulation in healthy subjects and patients with schizophrenia. In particular, the model fit was even improved over the descriptive HMM by reproducing shorter stable dominance times before a change to the unstable state. The HBM also provided more detailed explanations for the observed group difference that patients with schizophrenia showed higher alternation rates during intermittent stimulation, while percept alternation was decreases during continuous presentation. In particular, the HBM contains additional mechanisms of switching between stable and unstable state for intermittent presentation, which is assumed inactive during continuous presentation. The HBM, similar to the HMM, suggests an increased probability of switching to the unstable state for the patients with schizophrenia and thus, a longer relative time spent in the unstable state. The HBM also provides additional potential explanations related to the borders, or assumed population sizes, suggesting a higher increase from continuous (border b) to stable intermittent presentation (border b_S) in the healthy subjects. This is a first finding on the transition from continuous to intermittent presentation, which results from including both continuous and intermittent presentation in one model.

These findings suggested by the HBM, which include a longer relative time spent in the unstable state for the patients with schizophrenia and a smaller population size involved in percept stabilization, are also in agreement with recent findings of [41]. They studied the learning behavior of healthy subjects of whom the degree of delusional ideation [42] had been measured. In compliance with earlier studies [for a review see 43], they reported that subjects with larger delusion proneness made decisions on the basis of less information and were also less resilient against irrelevant information [compare also the literature about jumping to conclusions, e.g. 44, 45]. In the present setting, the ambiguous stimulus represents a constant source of partly contradicting visual information [see also 5, 8]. In that sense, the unstable state could be considered a state in which one is less resilient against this contradicting visual information, which yields a high rate of percept changes. The fact that the patients with schizophrenia spent more time in the unstable state is therefore highly consistent with the findings of [41]. Moreover, this finding is also compatible with current models of schizophrenia in the framework of predictive coding [46] that propose a reduced top-down influence of stored predictions. However, it goes beyond previous work by highlighting the role of a background process that controls the balance between stable and unstable states in perceptual inference. In addition, the population sizes could be considered related to the amount of information taken into consideration to create a percept. Again, consistently with [41], we find, in the stable state, larger estimated population sizes, b_S , of the perceptual populations L and R in healthy controls than in

patients with schizophrenia. Also, these population sizes are typically much larger than the population sizes in the unstable state ($b_S \gg b_U$), which would be consistent with the notion that subjects in the unstable state need less information to change their perception.

Applicability and model extensions

The HBM may also be used to describe dominance times resulting from other experiments with ambiguous visual stimuli studying, e.g., motion-induced-blindness, binocular rivalry, moving plaids, the Necker Cube, orthogonal gratings or the house/face-paradoxon [e.g. 21, 47] or also bistable auditory stimuli [48]. The HBM is, however, not designed for tristable stimuli, and transient stimulus manipulations as used in after-effect studies cannot be captured by the HBM in its current form. In different bistable settings, the HBM cannot be applied directly, but would allow for potential extensions. For example, in its present form, the HBM describes only balanced perception. However, it could be extended with respect to unbalanced bistable displays, e.g., for different eye contrasts during binocular rivalry [49], by choosing different drift parameters for the positive and the negative drift direction during presentation. Similarly, the drift could be chosen to vary as a function of attention [50, 51] or as a function of long-term history (e.g., the cumulative history H proposed in [31]). In studies on mixed perception during binocular rivalry [19], one might use an additional border to define an intermediate range for the perception process in which mixed perception is described.

One should note that the HBMI in its current form is restricted to a duration of blank displays $l_b \leq l_p \cdot v_0/v_S$. For longer blank displays, the mean drift of P during stable states, v_S^* , will be negative, yielding no perception change with high probability. However, it would be possible to extend the model accordingly, assuming a temporal evolution in the drift parameters, given corresponding extended empirical observations. In addition, note that the border of the perception process is assumed to be b during continuous stimulation and b_S (or b_U) during intermittent stimulation. Therefore, an instantaneous change of intermittent to continuous presentation is not yet described. Here, we qualitatively assume that the border jumps very fast from b to b_S with the onset of a blank display, while going back slowly during stimulation. A transition from continuous to intermittent presentation would therefore instantly change the response pattern, while a reverse transition would gradually reverse the change back to the one-state process. Quantitative validation and fitting of this assumption would be interesting, but requires corresponding empirical observations, in which the length of the presentation period l_p is varied. This would also allow investigation of potential relations between the HBMC and HBMI parameters and thus, between the mechanisms assumed to underlie the identified group differences.

Concerning the impact of the duration of the blank display l_b , two aspects should be discussed. First, the HBM can theoretically reproduce a phenomenon reported earlier in [14]. Conditional that one percept has been present for a short while, the probability of a percept change rises with the blank duration l_b . In the HBMI, the same is observed during the unstable state with typically short dominance times: During the unstable state the drift in the blank displays, v_U , is typically larger than the drift v_0 during stimulation. Therefore, longer blank displays speed up P , thereby reducing perceptual stability.

Second, one interesting potential model extension is concerned with the relationship between the length of the blank display and the alternation rate. As reported earlier by [14, 15, 18], the mean dominance time in intermittent presentation has been found to be a function of the relationship between the presentation length l_p (or 'ON'-period) and the length of the blank display l_b (or 'OFF'-period). Particularly, the dependence between l_b and the alternation rate is non-monotonic, as would be implied in the HBMI, but follows an inverted U-shape [22,

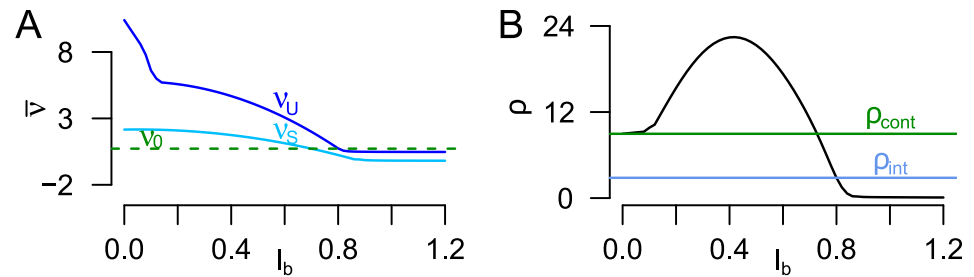


Fig 19. HBMI extension for different lengths of blank display. Parameters derived for subject B and C in Tables 3 and 4, which is the same subject during continuous and intermittent presentation. A) The mean drifts v_S , v_U decrease with the length of the blank display l_b . B) The resulting mean alternation rate ρ under intermittent stimulation (black) is derived using formula (18) (Materials and methods). The empirical mean alternation rates per minute of the subject during intermittent viewing with $l_b = 0.8$ and during continuous viewing are marked by blue and green lines, respectively.

<https://doi.org/10.1371/journal.pcbi.1005856.g019>

[24, 25] with a peak roughly at 0.4 s. Such an inverted U-shape would be possible in a model extension of the HBMI. As discussed in the Results, the drift terms v_S , v_U only represent the mean drift across the period of blank display, which is sufficient and parsimonious in the given data set with fixed length of blank display. However, the model would be fully consistent with the assumption that the drifts change during the ‘OFF’-period, such that the mean drifts $v_S(l_b)$ and $v_U(l_b)$ are functions of the length of the blank display l_b . In Fig 19A these mean drifts v_S , v_U decrease with l_b , where the stronger drifts at the beginning of the blank display could be effects of the recent stimulation. Panel B shows the resulting mean alternation rate, which has an inverted U-shape with a maximum around 0.4 s and shows increased stability under intermittent stimulation for $l_b > 0.7$. This increased stability is caused first by a small drift $v_S < v_0$ in that range. Second, it is also caused by the fact that the time interval l_b in which the background process B has positive drift is longer, leading to an increased probability to reach \tilde{b}_S and thus, to stay in the stable state. Estimation of the functions $v_S(l_b)$ and $v_U(l_b)$ from a suitable data set with variable lengths of blank displays would be an interesting task.

In summary, the proposed HBM intends to provide a link between empirical data analysis and mechanistic modeling. On the one hand, it aims at precisely describing the high variety of response patterns observed in perceptual responses to bistable stimuli. On the other hand, it aims at bridging the gap to detailed mechanistic models of bistable perception, allowing assumed processes to be fitted to short empirical data sets and thus, also the analysis of group differences. Various extension possibilities show a potential of the HBM to investigate related experimental contexts. By including both continuous and intermittent stimulation, the HBM can thus provide interesting new hypotheses on potential neuronal mechanisms of cognitive phenomena.

Materials and methods

Estimation of HMM parameters

Here we describe the estimation procedures of the HMM parameters for continuous presentation and for intermittent presentation. We denote by $d := (d_1, d_2, \dots, d_n)$ the set of dominance times modeled as realizations of random variables $D = (D_1, \dots, D_n)$.

Continuous presentation: ML estimation. The HMM for continuous presentation assumes that all dominance times d_i are independent and Inverse Gaussian distributed with

mean μ and standard deviation σ . The log-likelihood function is then given by

$$\log L(d|\mu, \sigma) = \frac{1}{2} \sum_{i=1}^n \log \left(\frac{\mu^3}{2\pi\sigma^2 d_i^3} \right) - \frac{\mu}{\sigma^2} \frac{(d_i - \mu)^2}{d_i}. \tag{12}$$

Setting the partial derivatives with respect to μ and σ to zero yields the estimates of Eq (12) [32]:

$$\hat{\mu} = \bar{\mu}_d := \frac{1}{n} \sum_{i=1}^n d_i \quad \text{and} \quad \hat{\sigma} = \sqrt{\frac{\bar{\mu}_d^3}{n} \sum_{i=1}^n (1/d_i - 1/\bar{\mu}_d)}.$$

Intermittent presentation: Baum-Welch-Algorithm. The HMM for intermittent presentation uses the parameter set $\vartheta = (\mu_S, \sigma_S, \mu_U, \sigma_U, p_{SS}, p_{UU})$ for the mean and standard deviation in the stable and unstable state, respectively, and the transition probabilities between these two states. This parameter set is estimated with the Baum-Welch-Algorithm [BWA, 33] which is an iteratively working instance of the EM-Algorithm maximizing the model likelihood locally. Here, we explain it briefly [for details see, e.g., 34]. In the first step one applies the so called Forward- and Backward-Algorithm. The forward-variable $\alpha_j(i)$ is defined as the probability of observing the sequence d_1, d_2, \dots, d_i and being in state j at time i , given the model parameters. The backward-variable $\beta_j(i)$ denotes the analogous probability of observing the ending partial sequence $d_{i+1}, d_{i+2}, \dots, d_n$ and being in state j at time i . To avoid underflow we normalize both variables [e.g. 34], resulting in the normalized variables $\tilde{\alpha}_j(i), \tilde{\beta}_j(i)$. The normalized variables can be derived iteratively as follows

$$\alpha_j^*(1) := \pi_j f_{\mu_j, \sigma_j}^{\text{IG}}(d_1), \quad c_i := \alpha_S^*(i) + \alpha_U^*(i), \quad \tilde{\alpha}_j(i) := \alpha_j^*(i)/c_i,$$

$$\alpha_j^*(i) := f_{\mu_j, \sigma_j}^{\text{IG}}(d_i) \sum_{k \in \{S, U\}} \tilde{\alpha}_k(i-1) \tilde{p}_{kj} \quad \text{for } i = 2, \dots, n,$$

$$\tilde{\beta}_j(n) = 1/c_n \quad \text{and} \quad \tilde{\beta}_j(i) = \sum_{k \in \{S, U\}} p_{jk} f_{\mu_k, \sigma_k}^{\text{IG}}(d_{i+1}) \tilde{\beta}_k(i+1)/c_i \quad \text{for } i = n-1, \dots, 1,$$

where π_j denotes the probability that the Markov chain starts in state j , $p_{SU} = 1 - p_{SS}$, $p_{US} = 1 - p_{UU}$ and $f_{\mu, \sigma}^{\text{IG}}(x)$ denotes the density of the IG distribution with expectation μ and standard deviation σ evaluated at x . Note that we suppress the dependence of $\tilde{\alpha}_j(i)$ and $\tilde{\beta}_j(i)$ on the parameter set ϑ for convenience.

The forward and backward variables are used to derive the probability $\gamma_j(i|\vartheta)$ of being in state j at time i , given the whole sequence $d := (d_1, \dots, d_n)$ and the parameters ϑ

$$\gamma_j(i|\vartheta) = \frac{\tilde{\alpha}_j(i) \tilde{\beta}_j(i)}{\tilde{\alpha}_S(i) \tilde{\beta}_S(i) + \tilde{\alpha}_U(i) \tilde{\beta}_U(i)}.$$

Moreover, we need the probability $\xi_{j,k}(i|\vartheta)$ of being in state j at time i and in state k at time $i + 1$, given the data d and the parameters ϑ ,

$$\xi_{j,k}(i|\vartheta) = \frac{\tilde{\alpha}_j(i) p_{jk} \tilde{\beta}_k(i+1) f_{\mu_k, \sigma_k}^{\text{IG}}(d_{i+1})}{\sum_j \sum_k \tilde{\alpha}_j(i) p_{jk} \tilde{\beta}_k(i+1) f_{\mu_k, \sigma_k}^{\text{IG}}(d_{i+1})}.$$

To iteratively derive the parameter estimates, the BWA applies expectation maximization as follows. Let $\vartheta^{(m)}$ denote the parameter estimates after the m -th iteration step, and let \mathfrak{V} denote

the set of all possible state sequences of the hidden Markov chain. Let $Y = (Y_1, \dots, Y_n)$ denote a \mathfrak{Y} -valued random variable and $y = (y_1, \dots, y_n)$ a realization of Y . The BWA then iteratively maximizes the *Q-function* [e.g. 35] over Y ,

$$Q := Q(\mathfrak{g}|\mathfrak{g}^{(m)}) = \sum_{y \in \mathfrak{Y}} \log L(d, y|\mathfrak{g}) \mathbb{P}(Y = y|d, \mathfrak{g}^{(m)}),$$

i.e., the expectation of the complete-data log-likelihood L across all possible paths $y \in \mathfrak{Y}$. The updated parameter set $\mathfrak{g}^{(m+1)}$ is chosen such that it maximizes Q . For a fixed state sequence $y = (y_1, \dots, y_n)$ the log-likelihood of the data is

$$\log L(d, y|\mathfrak{g}) = \log \pi_{y_1} + \log f_{\mu_{y_1}, \sigma_{y_1}}^{\text{IG}}(d_1) + \sum_{i=2}^n (\log(p_{y_{i-1}y_i}) + \log(f_{\mu_{y_i}, \sigma_{y_i}}^{\text{IG}}(d_i))).$$

Insertion into Q yields

$$\begin{aligned} Q(\mathfrak{g}|\mathfrak{g}^{(m)}) &= \sum_{y_1 \in \{S, U\}} \log \pi_{y_1} \mathbb{P}(Y_1 = y_1|d, \mathfrak{g}^{(m)}) \\ &+ \sum_{y_{i-1} \in \{S, U\}} \sum_{y_i \in \{S, U\}} \sum_{i=2}^n \log p_{y_{i-1}y_i} \mathbb{P}(Y_{i-1} = y_{i-1}, Y_i = y_i|d, \mathfrak{g}^{(m)}) \\ &+ \sum_{y_i \in \{S, U\}} \sum_{i=1}^n \frac{1}{2} \mathbb{P}(Y_i = y_i|d, \mathfrak{g}^{(m)}) \left(\log \left(\frac{\mu_{y_i}^3}{2\sigma_{y_i}^2 \pi d_i^3} \right) - \frac{\mu_{y_i}}{\sigma_{y_i}^2} \frac{(d_i - \mu_{y_i})^2}{d_i} \right). \end{aligned}$$

Note that the first line depends only on the initial distribution π , the second line depends on the transition probabilities and the third line depends on the parameters of the Inverse Gaussian distributions. Therefore, iterative parameter estimation separately maximizes these terms. Note further that we can rewrite $\mathbb{P}(Y_i = y_i|d, \mathfrak{g}^{(m)}) = \gamma_{y_i}(i|\mathfrak{g}^{(m)})$ and $\mathbb{P}(Y_{i-1} = y_{i-1}, Y_i = y_i|d, \mathfrak{g}^{(m)}) = \zeta_{y_{i-1}y_i}(i-1|\mathfrak{g}^{(m)})$, which yields the following estimates in the $m + 1$ -st iteration step.

For the entries of the transition matrix, Lagrange maximization of the second line under the constraints $p_{SS} + p_{SU} = p_{UU} + p_{US} = 1$ yields the estimate

$$\hat{p}_{jk}^{(m+1)} = \frac{\sum_{i=1}^{n-1} \zeta_{j,k}(i|\mathfrak{g}^{(m)})}{\sum_{i=1}^{n-1} \gamma_j(i|\mathfrak{g}^{(m)})},$$

in the $(m + 1)$ -th step of the BWA [e.g. 34].

The big bracket in the last line of the Q -function is analogous to the likelihood function of the IG distribution (Eq (12)) given in the section ‘Continuous presentation: ML estimation’ (with the additional indices y_i and the weighting factors $\mathbb{P}(Y_i = y_i|d, \mathfrak{g}^{(m)})$). Therefore, setting the partial derivatives to zero yields for $j \in \{S, U\}$ the analogous estimates

$$\hat{\mu}_j^{(m+1)} = \frac{\sum_{i=1}^n \gamma_j(i|\mathfrak{g}^{(m)}) d_i}{\sum_{i=1}^n \gamma_j(i|\mathfrak{g}^{(m)})}, \quad \hat{\sigma}_j^{(m+1)} = \sqrt{\frac{\hat{\mu}_j^3}{\sum_{i=1}^n \gamma_j(i|\mathfrak{g}^{(m)})} \sum_{i=1}^n \gamma_j(i|\mathfrak{g}^{(m)}) \left(\frac{1}{d_i} - \frac{1}{\hat{\mu}_j} \right)}.$$

which are the updates for parameters of the Inverse Gaussian distributions in the $(m + 1)$ -th step of the BWA.

In order to update the starting distribution, we do not maximize the respective summand of Q but we assume the stationary distribution as the starting distribution π . The stationary

distribution is derived using the eigenvectors of the transpose of the transition matrix [52], which yields

$$\pi_S^{(m+1)} = \frac{p_{UU}^{(m+1)} - 1}{p_{SS}^{(m+1)} + p_{UU}^{(m+1)} - 2}, \quad \text{and} \quad \pi_U^{(m+1)} = 1 - \pi_S^{(m+1)}.$$

These iterative steps are repeated until a desired level of convergence is reached.

Starting values and constraints. As starting values $\mu_S^{(s)}, \sigma_S^{(s)}, \mu_U^{(s)}, \sigma_U^{(s)}, p_{SS}^{(s)}, p_{UU}^{(s)}$ for the Baum-Welch algorithm we chose, in correspondence with the data set, $p_{SS}^{(s)} = p_{UU}^{(s)} = 0.5; \mu_S^{(s)} = 4; \sigma_U^{(s)} = 5$. In order to reduce the probability that the Baum-Welch-Algorithm will be captured in a local extremum, we chose ten equidistant values for $\mu_S^{(s)}$ ranging between 60 and $0.95 \max_i d_i$, and for each value of $\mu_S^{(s)}$ we choose ten equidistant values for $\sigma_S^{(s)}$ between 10 and $1.1 \mu_S^{(s)}$. Out of the resulting one hundred sets of parameter estimates we chose the parameter set with the highest log-likelihood. If the response pattern shows only dominance times larger than 30 seconds we reduce the model to the stable phase. The parameters μ_S and σ_S are derived by ML as described in the section ‘Continuous presentation: ML estimation’, and we set $p_{SS} := 1$. If the dominance time are only smaller than 30 seconds, we only estimate μ_U and σ_U by ML and use $p_{UU} = 1$.

For subjects with relatively clear distinction between long and short dominance times this procedure yields reasonable estimates. For subjects with less clear distinction, we added the following constraints based on the idea that short dominance times should not affect estimation of the stable parameters and long dominance times should not affect estimation of unstable parameters. Note that in continuous presentation where no state exists, about 90% of the dominance times are shorter than 15 seconds, while only about two percent are larger than 30 seconds. Therefore, we first require $\hat{\sigma}_S > 1$ such that not just the largest dominance time is estimated as stable and all others are categorized as unstable (which may increase the likelihood). Second, we do not accept HMMs with $\hat{\mu}_S < 0.98 \hat{\mu}_{15}$ where $\hat{\mu}_{15} := (1/|d \mathbb{1}_{d>15}|) \sum_{i=1}^n d_i \mathbb{1}_{d_i>15}$. This prevents dominance times smaller than 15 seconds to be considered for the estimation of μ_S . Third, we require $\hat{\mu}_S < 1.02 \hat{\mu}_{75}$ where $\hat{\mu}_{75} := (1/|d \mathbb{1}_{d>75}|) \sum_{i=1}^n d_i \mathbb{1}_{d_i>75}$ if any dominance time is larger than 75 seconds and $\hat{\mu}_{75} = 75$ otherwise such that rather stable dominance times longer than 75 seconds are not classified as unstable.

Expected relative time spent in the stable state. Here we derive the formula for the expected time φ_S spent in the stable state investigated in Fig 9. To that end, let $N_j, j \in \{S, U\}$, denote the (random) number of dominance times in state j before a state change. As the number of dominance times before a state change is geometrically distributed with probability $1 - p_{jj}$, we have for the expectation

$$\mathbb{E}[N_j] = \frac{1}{1 - p_{jj}}.$$

Moreover, let $(D_i^S)_{i \geq 1}$ be a sequence of independent $\text{IG}(\mu_S, \sigma_S)$ -distributed random variables and $(D_i^U)_{i \geq 1}$ be a sequence of $\text{IG}(\mu_U, \sigma_U)$ -distributed random variables. We derive φ_S as

$$\varphi_S := \frac{\mathbb{E}[\text{length of a stable phase}]}{\mathbb{E}[\text{length of a stable phase} + \text{length of an unstable phase}]}.$$

The length of a stable phase is a random variable distributed like $\sum_{i=1}^{N_S} D_i^S$, where $D_1^S, \dots, D_{N_S}^S$ are independent from each other and also independent of N_S . Therefore we have

[53]

$$\mathbb{E} \left[\sum_{i=1}^{N_S} D_i^S \right] = \mathbb{E}[N_S] \mathbb{E}[D_i^S] = \frac{\mu_S}{1 - p_{SS}},$$

and analogously for the expected length of an unstable phase. This yields

$$\varphi_S = \frac{\mu_S / (1 - p_{SS})}{\mu_S / (1 - p_{SS}) + \mu_U / (1 - p_{UU})} = \frac{(1 - p_{UU}) \mu_S}{(1 - p_{UU}) \mu_S + (1 - p_{SS}) \mu_U}. \tag{13}$$

Estimation of HBM parameters

Here we describe the estimation procedures of the HBM parameters for continuous presentation and for intermittent presentation.

HBMc. Recall that the HBMc describes the perception process P as a Brownian motion with drift $\pm v_0$, where the sign of the drift changes at the first hitting time of the borders $\pm b$. The two parameters (b, v_0) are estimated using the ML method. Due to transformation invariance of the ML estimates, we can therefore use the ML estimators from the section ‘Continuous presentation: ML estimation’ applying that the resulting dominance times are IG distributed with expectation $\mu = 2b/v_0$ and variance $\sigma^2 = 2b/v_0^3$, or conversely, $b = (1/2) \cdot \sqrt{\mu^3/\sigma^2}$ and $\hat{v}_0 = \sqrt{\mu/\sigma^2}$. This yields

$$\hat{v}_0 = 2n\hat{b} / \sum_{i=1}^n d_i \quad \text{and} \quad \hat{b} = \frac{1}{2} \sqrt{\frac{1}{1/n \sum_{i=1}^n 1/d_i - n / \sum_{i=1}^n d_i}}.$$

HBMi. The parameters of the HBMi are $\Theta = (b_S, v_S^*, b_U, v_U^*, \tilde{b}_S, \tilde{b}_U, v_B^*)$ where (b_S, v_S^*) and (b_U, v_U^*) denote the border and drift parameters of the perception process P in the stable and the unstable state, respectively. $\tilde{b}_S, \tilde{b}_U, v_B^*$ are the border and drift parameters of the background process B which determines the hidden state. Recall that the likelihood of the whole model given the data and the parameter vector Θ is approximately

$$L(d|\Theta) \approx \alpha_S(n) + \alpha_U(n),$$

using the forward variables $\alpha_j(i)$ defined recursively in Eq (10). Note again that we suppress dependence on the parameters Θ for convenience. In practice we need to avoid underflow when calculating $\alpha_j(i)$. To that end the forward variables are normalized such that $\sum_j \alpha_j(i) = 1$ for all time points i , using the following steps [40]. For $j \in \{S, U\}$,

$$\begin{aligned} \alpha_j^*(1) &:= \alpha_j(1), & c_i &:= \alpha_S^*(i) + \alpha_U^*(i), & \tilde{\alpha}_j(i) &:= \alpha_j^*(i)/c_i, \\ \alpha_j^*(i) &:= f_{\mu_j, \sigma_j}^{\text{IG}}(d_i) \sum_{k \in \{S, U\}} \tilde{\alpha}_k(i-1) \tilde{p}_{kj}, & i &> 1. \end{aligned}$$

The likelihood then derives as

$$L(d|\Theta) \approx \prod_{i=1}^n c_i (\tilde{\alpha}_S(n) + \tilde{\alpha}_U(n)) = \prod_{i=1}^n c_i, \quad \text{yielding} \quad \log L \approx \sum_{i=1}^n \log(c_i). \tag{14}$$

Parameter estimation is then obtained by maximizing $\sum_i \log(c_i)$, which is a function of the model parameters Θ . To that end we apply the Newton-type algorithm [54] implemented in

the R-function $\text{nlm}()$. Alternatively the COBYLA-Algorithm [55] for maximization under non-linear constraints can be applied.

Next we discuss the set of starting values $\{b_S^{(s)}, v_S^{*(s)}, b_U^{(s)}, v_U^{*(s)}, \tilde{b}_S^{(s)}, \tilde{b}_U^{(s)}, v_B^{*(s)}\}$ for the optimization algorithm. Let $\mathcal{U} := \{\hat{\mu}_{15}, \hat{\mu}_{30}\}$ and $\mathcal{O} := \{\hat{\sigma}_{15}, \hat{\sigma}_{30}, 1.15\hat{\sigma}_{30}\}$, where $\hat{\mu}_k$ and $\hat{\sigma}_k$ denote the empirical mean and standard deviation of all dominance times larger than k seconds. Then we choose the initial values for b_S and v_S^* from the sets

$$b_S^{(s)} \in \{\sqrt{\mu^3/\sigma^2}/2 | \mu \in \mathcal{U}, \sigma \in \mathcal{O}\}, \quad v_S^{*(s)} \in \{\sqrt{\mu/\sigma^2} | \mu \in \mathcal{U}, \sigma \in \mathcal{O}\}.$$

Depending on $v_S^{*(s)}$ and $b_S^{(s)}$, we choose $b_U^{(s)} \in \{0.01b_S^{(s)}, 0.05b_S^{(s)}, 0.15b_S^{(s)}\}$, $\tilde{b}_S^{(s)} \in \{b_S^{(s)}, 10b_S^{(s)}\}$, $v_U^{*(s)} \in \{1.01v_S^{*(s)}, 3v_S^{*(s)}, 7v_S^{*(s)}\}$, furthermore $v_B^{*(s)} \in \{0.1, 3\}$ and $\tilde{b}_U^{(s)} \in \{-3, 0, 3\}$.

The initial value for $\pi_S = \mathbb{P}(\text{process starts in the stable state})$ is set to $\hat{\pi}_S = 1$ if $d_1 \geq 45$, $\hat{\pi}_S = 0$ if $d_1 \leq 15$, and $\hat{\pi}_S = 1/2$ otherwise. Alternatively and comparable to the HMM the stationary distribution can be used for the initial distribution, thereby reducing the number of parameters by one. As this, however, requires numerical integration and increases the computational effort considerably and the differences between the two approaches were negligible we used the simpler rule.

The maximization algorithm is applied using all combinations of starting values. We then take the set of parameter estimates which yields the highest log-likelihood and fulfills the following constraints

$$\begin{aligned} \text{A) } v_U &\geq v_S; & \text{B) } 0 &\leq b_U \leq b_S; & \text{C) } \tilde{b}_S &\geq 0; & \text{D) } v_B &> 0; \\ \text{E) } \mu_S^* &\approx 2b_S/v_S^* \geq 0.98\hat{\mu}_{15}; & \text{F) } \sigma_S^* &\approx \sqrt{2b_S/v_S^{*3}} < 1.20\hat{\sigma}_{15}. \end{aligned}$$

Constraints A) to D) result from the model assumptions, E) prevents dominance times smaller than 15 seconds to be considered for the estimation of μ_S , analogously to the HMM procedure. Constraint F) prevents too big estimates of standard deviations, which would yield implausible results.

In cases in which only dominance times larger than 30 seconds are observed, we apply the algorithm described for the HBMc, where b_S, v_S^* are estimated like in the section ‘HBMc’ and \tilde{b}_S is set to zero. In cases in which only dominance times up to 30 seconds are observed, we proceed analogously, where b_U, v_U^* are estimated like in the section ‘HBMc’ and set $\tilde{b}_U = 10^{10}$. In either case we set $v_b^* = 10$ and do not estimate the other variables.

HBMi: Dominance time before state changes

Stable dominance time before a change to the unstable state. In the section ‘Relation of the HBMi to the two state HMM’ we discussed that in the HBMi, stable dominance times are shorter when occurring directly before a state change to the unstable state than when followed by another stable dominance time. The same observation was also made in the sample dataset (Fig 17). Therefore, we derive here the expected length μ_S^- of a stable dominance time before a state change and its corresponding expected length μ_S^+ before another stable dominance time.

Let $\Theta = (b_S, v_S^*, b_U, v_U^*, \tilde{b}_S, \tilde{b}_U, v_B^*)$ be the parameter set of an HBMi, and let D_1^S denote an $\text{IG}(2b_S/v_S^*, \sqrt{2b_S/v_S^{*3}})$ -distributed random stable dominance time. Further, let p_{SS}^* denote the probability to stay in the stable state (8), and let Y_i denote the hidden state during the i -th

dominance time. Then it holds

$$\begin{aligned} \mu_s^+ &:= \mathbb{E}[D_1^s | Y_1 = S, Y_2 = S] = \mathbb{E}[D_1^s | Y_1 = S, B_{D_1^s} \geq \tilde{b}_s] \\ &= \frac{\int_0^\infty \mathbb{P}(D_1^s = t, B_{D_1^s} \geq \tilde{b}_s) dt}{\mathbb{P}(B_{D_1^s} \geq \tilde{b}_s)}. \end{aligned}$$

Now we integrate across all possible values t that D_1^s can take, using independence of B and P during D_1^s , IG distribution of D_1^s and normal distribution of $B_{D_1^s}$. Further we note that the probability in the denominator equals p_{SS}^* , which yields

$$\mu_s^+ = (p_{SS}^*)^{-1} \int_0^\infty f_{2b_s/v_s^*, \sqrt{2b_s/v_s^3}}^{IG}(t) (1 - \Phi_{v_s^*, t}(\tilde{b}_s)) dt. \tag{15}$$

Analogously, we obtain

$$\mu_s^- = \mathbb{E}[D_1^s | Y_1 = S, Y_2 = U] = (1 - p_{SS}^*)^{-1} \int_0^\infty f_{2b_s/v_s^*, \sqrt{2b_s/v_s^3}}^{IG}(t) \Phi_{v_s^*, t}(\tilde{b}_s) dt. \tag{16}$$

Expected relative time spent in the stable state. In this section we derive the formula for the expected relative time φ_s^* spent in the stable state in the HBMi analyzed in the section ‘Application of the HBM to the sample data set’. To that end, let $N_j^*, j \in \{S, U\}$, denote the (random) number of dominance times of the HBMi in state j before a state change. As the number of dominance times before a state change is geometrically distributed with probability $1 - p_{jj}^*$ its expectation is given by $\mathbb{E}[N_j^*] = (1 - p_{jj}^*)^{-1}$. (Again, due to taking the mean of different drifts, this derivation is only a close approximation, but we omit approximation signs here for convenience.) Again, we derive φ_s^* as

$$\varphi_s^* := \frac{\mathbb{E}[\text{length of a stable phase}]}{\mathbb{E}[\text{length of a stable phase} + \text{length of an unstable phase}]}.$$

The length of a stable phase is a random variable distributed like $\sum_{i=1}^{N_s^*} D_i^s$, where $D_1^s, \dots, D_{N_s^*}^s$ denote the random stable dominance times. Given N_s^* , we use linearity of expectation to compute the conditional expectation

$$\mathbb{E} \left[\sum_{i=1}^{N_s^*} D_i^s \middle| N_s^* \right] = (N_s^* - 1)\mu_s^+ + \mu_s^-.$$

Now we take expectation over N_s^* to find

$$\mathbb{E} \left[\sum_{i=1}^{N_s^*} D_i^s \right] = \frac{p_{SS}^*}{1 - p_{SS}^*} \mu_s^+ + \mu_s^-.$$

An analogous result holds for the expected length of an unstable phase. This yields

$$\varphi_s^* = \frac{\frac{p_{SS}^*}{1 - p_{SS}^*} \mu_s^+ + \mu_s^-}{\frac{p_{SS}^*}{1 - p_{SS}^*} \mu_s^+ + \mu_s^- + \frac{p_{UU}^*}{1 - p_{UU}^*} \mu_U^+ + \mu_U^-}. \tag{17}$$

HBMi: Derivation of the alternation rate ρ as used in Fig 19B

Here, we derive a formula for the alternation rate ρ used in Fig 19 given the HBMi parameter set $(\mu_s^*, \sigma_s^*, \mu_u^*, \sigma_u^*, p_{ss}^*, p_{uu}^*)$. For each length of blank display l_b and the value of v_s, v_u as shown in Fig 19A the mean drifts per second in the stable and the unstable state v_s^*, v_u^* and the mean drift of the background process v_b^* are derived using Eq (4). Then, we use Eqs (8) and (11) to derive the values $(\mu_s^*, \sigma_s^*, \mu_u^*, \sigma_u^*, p_{ss}^*, p_{uu}^*)$ given the mean drifts per second, and we recall Eq (17) for φ_s^* , and analogously for φ_u^*

We now show that the alternation rate can be described as

$$\rho := \lim_{\Delta \rightarrow \infty} \frac{\mathbb{E}[N^*(\Delta)]}{\Delta} = \frac{\varphi_s^*}{\mu_s^*} + \frac{\varphi_u^*}{\mu_u^*}, \tag{18}$$

where $N^*(\Delta)$ denotes the number of perceptual changes in an interval of length Δ . We split up $N^*(\Delta) = N_s^*(\Delta) + N_u^*(\Delta)$, where $N_s^*(\Delta)$ and $N_u^*(\Delta)$ denote the number of perceptual changes in the respective stable and unstable phases in the interval of length Δ . We then show

$$\frac{\mathbb{E}[N_s^*(\Delta)]}{\Delta} \xrightarrow{\Delta \rightarrow \infty} \frac{\varphi_s^*}{\mu_s^*},$$

and analogously for the unstable state. To that end, let Δ_s be the time spent in the stable state in a time interval of length Δ . By the Elementary Renewal Theorem [e.g. 56] it holds

$$\frac{\mathbb{E}[N_s^*(\Delta)]}{\Delta_s} = \frac{\mathbb{E}[N_s^*(\Delta_s)]}{\Delta_s} \xrightarrow{\Delta \rightarrow \infty} \frac{1}{\mu_s^*}$$

as $\Delta \rightarrow \infty$ naturally implies $\Delta_s \rightarrow \infty$. According to the definition of φ_s^* as the expected relative time spent in the stable state, we get $\Delta_s/\Delta \rightarrow \varphi_s^*$ in probability. This yields the claim

$$\frac{\mathbb{E}[N_s^*(\Delta)]}{\Delta} = \frac{\Delta_s}{\Delta} \cdot \frac{\mathbb{E}[N_s^*(\Delta)]}{\Delta_s} \xrightarrow{\Delta \rightarrow \infty} \frac{\varphi_s^*}{\mu_s^*}.$$

Experimental protocol and data preprocessing

Details on experimental protocol. The sample data set for main analysis was collected as described in [10], the sample data for the analysis of reproducibility was collected as described in [9]. In both cases, participants took part in a behavioral experiment including continuous and intermittent stimulation with the same ambiguous stimulus. This stimulus was a structure-from-motion stimulus that appears as a rotating sphere. The rotation direction of the sphere is ambiguous and equally compatible with leftward and rightward percepts. For continuous stimulation, the sphere was presented throughout four minutes in which participants' perception spontaneously alternated between the two percepts. Participants indicated perceptual alternations via button presses on a keyboard. For intermittent stimulation, the sphere was presented for twenty minutes repeatedly for short intervals of 0.6 s interleaved by blank screens of 0.8 s duration. Here, participants indicated the perceived rotation direction at every 1.4 s at each presentation of the sphere. Please refer to [9, 10] for a detailed description of the data collection. As suggested in [10], missing responses during intermittent stimulation were replaced by their preceding responses because the reported percept typically persisted to the next available response.

Investigation of history dependence in the response patterns. Long-term dependencies in the data set of [10] during continuous presentation are analyzed using the Pearson

correlation coefficient c_H between the dominance times and the cumulative history H as introduced in [31]. The history H is a function of the length and recency of previously dominated percepts. For each of the 57 subjects with at least five dominance times, c_H is estimated as explained in [31]. To assess statistical significance, 1000 data sets are obtained for each subject by permutation of the dominance times to approximate the distribution of c_H under the null hypothesis of independent and identically distributed dominance times. Statistical significance on the 5% level is obtained by comparison of the empirical history c_H to the 95% quantile of the distribution of c_H derived from the permuted data sets.

Supporting information

S1 Data. Raw data including dominance times from the studies [9, 10]. Data include dominance times for intermittent (AlbertetalDataINT2015.RData) and continuous (AlbertetalDataCON2015.RData) stimulation for each of the 61 subjects of [10], and dominance times of two trials of continuous stimulation for each of the 105 subjects of [9] (AlbertetalDataCON2013.RData). A pdf file (AlbertetalDataDescription.pdf) provides a detailed description of the data-sets. The data files and pdf file are packed in the supporting information file S1_Data.zip. (ZIP)

Acknowledgments

We thank Brooks Ferebee and Michael Messer for inspiring discussions.

Author Contributions

Conceptualization: Philipp Sterzer, Gaby Schneider.

Data curation: Katharina Schmack.

Formal analysis: Stefan Albert.

Funding acquisition: Katharina Schmack, Philipp Sterzer, Gaby Schneider.

Investigation: Katharina Schmack.

Methodology: Stefan Albert, Gaby Schneider.

Project administration: Gaby Schneider.

Software: Stefan Albert.

Supervision: Gaby Schneider.

Visualization: Stefan Albert.

Writing – original draft: Stefan Albert, Gaby Schneider.

Writing – review & editing: Katharina Schmack, Philipp Sterzer.

References

1. Necker LA. Observations on some remarkable optical phaenomena seen in Switzerland; and on an optical phaenomenon which occurs on viewing a figure of a crystal or geometical solid. London and Edinburgh Philosophical Magazine and Journal of Science. 1832; 1:329–337.
2. von Helmholtz H. Treatise on Physiological Optics. vol. 3. Birmingham, AL: The Optical Society of America; 1866.
3. Levelt W. On binocular rivalry. Soesterberg, The Netherlands: Institute for Perception; 1965. Available from: <https://pdfs.semanticscholar.org/7d66/4283485e4297924fe0abd03c7c75812c74a7.pdf>.

4. Leopold DA, Logothetis NK. Multistable phenomena: changing views in perception. *Trends Cogn Sci.* 1999; 3(7):254–264. [https://doi.org/10.1016/S1364-6613\(99\)01332-7](https://doi.org/10.1016/S1364-6613(99)01332-7) PMID: 10377540
5. Hohwy J, Roepstorff A, Friston K. Predictive coding explains binocular rivalry: An epistemological review. *Cognition.* 2008; 108:687–701. <https://doi.org/10.1016/j.cognition.2008.05.010> PMID: 18649876
6. Sterzer P, Kleinschmidt A, Rees G. The neural bases of multistable perception. *Trends Cogn Sci.* 2009; 13(7):310–318. <https://doi.org/10.1016/j.tics.2009.04.006> PMID: 19540794
7. Braun J, Mattia M. Attractors and noise: Twin drivers of decisions and multistability. *NeuroImage.* 2010; 52(3):740–751. <https://doi.org/10.1016/j.neuroimage.2009.12.126> PMID: 20083212
8. Weilhhammer V, Stuke H, Hesselmann G, Sterzer P, Schmack K. A predictive coding account of bistable perception—a model-based fMRI study. *PLOS Comput Biol.* 2017; 13(5):e1005536. <https://doi.org/10.1371/journal.pcbi.1005536> PMID: 28505152
9. Schmack K, Gómez-Carrillo de Castro A, Rothkirch M, Sekutowicz M, Rössler H, Haynes JD, et al. Delusions and the Role of Beliefs in Perceptual Inference. *J Neurosci.* 2013; 33(34):13701–13712. <https://doi.org/10.1523/JNEUROSCI.1778-13.2013> PMID: 23966692
10. Schmack K, Schnack A, Priller J, Sterzer P. Perceptual instability in schizophrenia: Probing predictive coding accounts of delusions with ambiguous stimuli. *Schizophr Res Cog.* 2015; 2(2):72–77. <https://doi.org/10.1016/j.scog.2015.03.005>
11. Brascamp JW, Klink PC, Levelt WJ. The 'laws' of binocular rivalry: 50 years of Levelt's propositions. *Vision Res.* 2015; 109:20–37. <https://doi.org/10.1016/j.visres.2015.02.019> PMID: 25749677
12. Levelt WJ. Note on the distribution of dominance times in binocular rivalry. *Br J Psychol.* 1967; 58(1): 143–145. <https://doi.org/10.1111/j.2044-8295.1967.tb01068.x> PMID: 5582864
13. Moreno-Bote R, Rinzel J, Rubin N. Noise-induced alternations in an attractor network model of perceptual bistability. *J Neurophysiol.* 2007; 98(3):1125–1139. <https://doi.org/10.1152/jn.00116.2007> PMID: 17615138
14. Brascamp JW, Pearson J, Blake R, van den Berg AV. Intermittent ambiguous stimuli: Implicit memory causes periodic perceptual alternations. *J Vis.* 2009; 9(3):3, 1–23. <https://doi.org/10.1167/9.3.3> PMID: 19757942
15. Leopold DA, Wilke M, Maier A, Logothetis NK. Stable perception of visually ambiguous patterns. *Nat Neurosci.* 2002; 5(6):605–609. <https://doi.org/10.1038/nn0602-851> PMID: 11992115
16. Murata T, Matsui N, Miyauchi S, Kakita Y, Yanagida T. Discrete stochastic process underlying perceptual rivalry. *NeuroReport.* 2003; 14(10):1347–1352. <https://doi.org/10.1097/00001756-200307180-00013> PMID: 12876471
17. Wilson HR. Minimal physiological conditions for binocular rivalry and rivalry memory. *Vision Res.* 2007; 47(21):2741–2750. <https://doi.org/10.1016/j.visres.2007.07.007> PMID: 17764714
18. Gigante G, Mattia M, Braun J, Del Giudice P. Bistable Perception Modeled as Competing Stochastic Integrations at Two Levels. *PLOS Comput Biol.* 2009; 5(7):e1000430. <https://doi.org/10.1371/journal.pcbi.1000430> PMID: 19593372
19. Gershman SJ, Vul E, Tenenbaum JB. Multistability and Perceptual Inference. *Neural Comput.* 2012; 24(1):1–24. https://doi.org/10.1162/NECO_a_00226 PMID: 22023198
20. Pastukhov A, Garcia-Rodriguez P, Haenicke J, Guillamon A, Deco G, Braun J. Multi-stable perception balances stability and sensitivity. *Front Comput Neurosci.* 2013; 7(17). <https://doi.org/10.3389/fncom.2013.00017> PMID: 23518509
21. Cao R, Pastukhov A, Mattia M, Braun J. Collective Activity of Many Bistable Assemblies Reproduces Characteristic Dynamics of Multistable Perception. *J Neurosci.* 2016; 36(26):6957–6972. <https://doi.org/10.1523/JNEUROSCI.4626-15.2016> PMID: 27358454
22. Orbach J, Ehrlich D, Heath H. Reversibility of the Necker cube. I. An examination of the concept of 'satiation or orientation'. *Percept Mot Skills.* 1963; 17:439–458. <https://doi.org/10.2466/pms.1963.17.2.439> PMID: 14065532
23. Maier A, Wilke M, Logothetis NK, Leopold DA. Perception of temporally interleaved ambiguous patterns. *Curr Biol.* 2003; 13(13):1076–1085. [https://doi.org/10.1016/S0960-9822\(03\)00414-7](https://doi.org/10.1016/S0960-9822(03)00414-7) PMID: 12842006
24. Kornmeier J, Bach M. Early neural activity in Necker-cube reversal: Evidence for low-level processing of a gestalt phenomenon. *Psychophysiology.* 2004; 41(1):1–8. <https://doi.org/10.1046/j.1469-8986.2003.00126.x> PMID: 14692995
25. Pastukhov A, Braun J. Structure-from-motion: dissociating perception, neural persistence, and sensory memory of illusory depth and illusory rotation. *Atten Percept Psychophys.* 2013; 75(2):322–340. <https://doi.org/10.3758/s13414-012-0390-0> PMID: 23150214

26. Cao R, Braun J, Mattia M. Stochastic Accumulation by Cortical Columns May Explain the Scalar Property of Multistable Perception. *Phys Rev Lett*. 2014; 113(9):098103. <https://doi.org/10.1103/PhysRevLett.113.098103> PMID: 25216009
27. Laing CR, Chow CC. A Spiking Neuron Model for Binocular Rivalry. *J Comput Neurosci*. 2002; 12(1): 39–53. <https://doi.org/10.1023/A:1014942129705> PMID: 11932559
28. Noest AJ, van Ee R, Nijs MM, van Wezel RJA. Percept-choice sequences driven by interrupted ambiguous stimuli: A low-level neural model. *J Vis*. 2007; 7(8):10, 1–14. <https://doi.org/10.1167/7.8.10> PMID: 17685817
29. Walker P. Stochastic properties of binocular rivalry alternations. *Percept Psychophys*. 1975; 18(6): 467–473. <https://doi.org/10.3758/BF03204122>
30. Lehky SR. Binocular rivalry is not chaotic. *Proc R Soc Lond B Biol Sci*. 1995; 259:71–76. <https://doi.org/10.1098/rspb.1995.0011>
31. Pastukhov A, Braun J. Cumulative history quantifies the role of neural adaptation in multistable perception. *J Vis*. 2011; 11(10):12, 1–10. <https://doi.org/10.1167/11.10.12> PMID: 21931128
32. Seshadri V. *The Inverse Gaussian Distribution*. Oxford University Press; 1993.
33. Baum LE, Petrie T, Soules G, Weiss N. A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Annals of Mathematical Statistics*. 1970; 41(1):164–171. <https://doi.org/10.1214/aoms/1177697196>
34. Rabiner LR. A tutorial on hidden markov models and selected applications in speech recognition. *IEEE Trans Inf Theory*. 1989; 77(2):257–284.
35. Bilmes JA. A gentle tutorial of the EM algorithm and its application to parameter estimation for gaussian mixture and hidden markov models. International Computer Science Institute; 1998.
36. MacDonald IL, Zucchini W. *Hidden Markov and other models for discrete-valued time series*. vol. 70 of *Monographs on Statistics and Applied Probability*. Chapman&Hall, London; 1997.
37. Ephraim Y, Merhav N. Hidden Markov processes. *IEEE Trans Inf Theory*. 2002; 48(6):1518–1569. <https://doi.org/10.1109/TIT.2002.1003838>
38. Samanta M. On Tests Of Equality Of Two Inverse Gaussian Distributions. *S Afr Stat J*. 1985; 19(2): 83–95.
39. Barndorff-Nielsen OE. Normal Inverse Gaussian Distributions and Stochastic Volatility Modeling. *Scand Stat Theory Appl*. 1997; 24(1):1–13.
40. Turner R. Direct maximization of the likelihood of a hidden Markov model. *Comput Stat Data Anal*. 2008; 52(9):4147–4160. <https://doi.org/10.1016/j.csda.2008.01.029>
41. Stuke H, Stuke H, Weillhammer VA, Schmack K. Psychotic Experiences and Overhasty Inferences Are Related to Maladaptive Learning. *PLOS Comput Biol*. 2017; 13(1):e1005328. <https://doi.org/10.1371/journal.pcbi.1005328> PMID: 28107344
42. Peters E, Joseph S, Garety P. Measurement of delusional ideation in the normal population: introducing the PDI (Peters et al. Delusions Inventory). *Schizophr Bull*. 1999; 25(3):533–576. <https://doi.org/10.1093/oxfordjournals.schbul.a033401>
43. Fletcher PC, Frith CD. Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. *Nat Rev Neurosci*. 2009; 10:48–58. <https://doi.org/10.1038/nrn2536> PMID: 19050712
44. Evans SL, Averbeck BB, Furl N. Jumping to conclusions in schizophrenia. *Neuropsychiatr Dis Treat*. 2015; 11:1615–1624. <https://doi.org/10.2147/NDT.S56870> PMID: 26170674
45. Ross RM, McKay R, Coltheart M, Langdon R. Jumping to Conclusions About the Beads Task? A Meta-analysis of Delusional Ideation and Data-Gathering. *Schizophr Bull*. 2015; 41(5):1183–1191. <https://doi.org/10.1093/schbul/sbu187>. PMID: 25616503
46. Adams RA, Stephan KE, Brown HR, Frith CD, Friston KJ. The computational anatomy of psychosis. *Front Psychiatry*. 2013; 4:47. <https://doi.org/10.3389/fpsy.2013.00047> PMID: 23750138
47. van Ee R. Dynamics of perceptual bi-stability for stereoscopic slant rivalry and a comparison with grating, house-face, and Necker cube rivalry. *Vision Res*. 2005; 45(1):29–40. <https://doi.org/10.1016/j.visres.2004.07.039> PMID: 15571736
48. Rankin J, Sussman E, Rinzel J. Neuromechanistic Model of Auditory Bistability. *PLOS Comput Biol*. 2015; 11(11):e1004555. <https://doi.org/10.1371/journal.pcbi.1004555>. PMID: 26562507
49. Brascamp JW, van Ee R, Noest AJ, Jacobs RAH, van den Berg A. The time course of binocular rivalry reveals a fundamental role of noise. *J Vis*. 2006; 6(11):1244–1256. <https://doi.org/10.1167/6.11.8>. PMID: 17209732
50. Pastukhov A, Braun J. Perceptual reversals need no prompting by attention. *J Vis*. 2007; 7(10):5, 1–17. <https://doi.org/10.1167/7.10.5> PMID: 17997674

51. Dieter KC, Brascamp J, D T, Blake R. Does visual attention drive the dynamics of bistable perception? *Atten Percept Psychophys*. 2016; 78(7):1861–1873. <https://doi.org/10.3758/s13414-016-1143-2> PMID: [27230785](https://pubmed.ncbi.nlm.nih.gov/27230785/)
52. Grimmett G, Stirzaker D. *Probability and random processes*. 3rd ed. Oxford University Press, New York; 2001.
53. Wald A. On cumulative sums of random variables. *The Annals of Mathematical Statistics*. 1944; 15(3): 7283–296. <https://doi.org/10.1214/aoms/1177731235>
54. Dennis JE, Schnabel RB. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Prentice-Hall, Englewood Cliffs, NJ; 1983.
55. Powell MJD. A direct search optimization method that models the objective and constraint functions by linear interpolation. In: Gomez S, Hennart JP, editors. *Advances in Optimization and Numerical Analysis*; 1994. p. 51–67.
56. Ross SM. *Stochastic processes*. Wiley, New York; 1996.