iScience, Volume *24*

Supplemental information

Bistable perception alternates between internal and external modes of sensory processing

Veith Weilnhammer, Meera Chikermane, and Philipp Sterzer

Supplementary Information

"Blasts from the past: Bistable perception alternates between internal and external modes of sensory processing."

Authors:

Veith Weilnhammer[1,2], Meera Chikermane[1], Philipp Sterzer[1,2,3,4]

Affiliations and Contribution:

[1] Department of Psychiatry, Charité-Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, 10117 Berlin, Germany

[2] Berlin Institute of Health, Charité-Universitätsmedizin Berlin and Max Delbrück Center, 10178 Berlin, Germany

[3] Bernstein Center for Computational Neuroscience, Charité-Universitätsmedizin Berlin, 10117 Berlin, Germany

[4] Berlin School of Mind and Brain, Humboldt-Universität zu Berlin, 10099 Berlin, Germany

Corresponding Author:

Veith Weilnhammer, Department of Psychiatry, Charité Campus Mitte, Charitéplatz 1, 10117 Berlin, phone: 0049 (0)30 450 517 317, email: veith-andreas.weilnhammer@charite.de

Word Count: 3415 words (main text without methods)

OSF-Project: https://osf.io/y2cfm/

1 Transparent Methods

1.1 Participants

We recruited a total of 20 participants (9 female; age: 27.45 ± 1.01 years). All participants had (corrected-to-) normal vision, were naive to the purpose of the study and gave informed, written consent prior to the experiment authorized by the Charité ethics committee.

1.2 Apparatus

Stimuli were presented on a 98PDF-CRT-Monitor (60 Hz, 1040 x 1050 pixels, 60 cm viewing distance, 41.38 pixels per degree [°] visual angle) using Psychtoolbox 3 and Matlab R2007b (MathWorks). 3D stimulation was achieved using 3D red-blue filter glasses. The blue filter was placed over the right eye.

1.3 Heterochromatic Flicker Photometry

Subjective differences in luminance can induce 3D effects based on the Pulfrich effect. To preclude that this phenomenon induces biases with regard to direction of rotation in partially ambiguous structure-from-motion stimuli, we conducted a separate pre-test experiment. We used *Heterochromatic Flicker Photometry* to estimate subjective equiluminance between red and blue. We presented red and blue circles (diameter: 6.45° visual angle) alternating at a frequency of 15 Hz. In case of subjective differences in luminance, participants perceived a flicker, which they reduced by adjusting the luminance of the blue stimulus initially presented at a random luminance between 0 and 125% relative to the red stimulus presented at a fixed luminance of 100%. Average equiluminance estimated across 10 such trials determined the monitor- and participant-specific luminance of the red- and blue-channels (average blue luminance: 110.85 ± 4.74%).

1.4 Main experiment

The main experiment assessed how perceptual history was integrated with varying levels of disambiguating sensory information. To induce bistable perception, we generated rotating discontinuous structure-from-motion stimuli by placing a total of 2000 dots (each subtending 0.08° visual angle, overall stimulus size: 14.5° x 14.5°) on the surface of a Lissajous band (see Figure 1A and additional Supplementary Video V1). The Lissajous band was formed by the perpendicular intersection of two sinusoids ($x(t) = sin(A * t)$ and $y(t) = cos(B * t + \delta)$ with $A = 3$, $B = 8$). Within each trial, the stimulus was presented for 2 sec, while δ increased from 0 to 0.5π. The width of the Lissajous band was set to 0.04π ° rotational angle. Fixation intervals between trials were uniformly jittered around 2.5 ± 0.25 sec.

To generate parametric 3D stimuli, we attached a stereo-disparity signal to a fraction of the dots on the Lissajous band. Dots that carried stereo-disparity information were represented on separate monocular channels. To this end, corresponding pairs of red (left eye) and blue (right eye) dots were shifted against each other by 0.01π rotational angle. Dots without stereo-disparity information were presented binocularly. The wavelength of binocular dots (purple) was defined by adding the individual wavelengths of red and blue (see Heterochromatic Flicker Photometry). Throughout the experiment, we varied the signal-to-ambiguity ratio (SAR) by manipulating the fraction of dots that carried stereo-disparity information (see below). The direction of disambiguating sensory information (i.e., whether the front surface of the partially disambiguated sphere moved to the left or to the right) was randomized across trials.

We instructed participants to indicate the perceived direction of rotation of the Lissajous band by pressing the arrow-keys on a standard keyboard (right index finger: rotation of the front surface of the Lissajous to the *left*; right ring finger: rotation to the *right*; middle finger: *unclear* or mixed direction of rotation). *Error* responses were defined for trials at which participants did not respond before the end of stimulus presentation or indicated more than one perceptual response.

Within one run, participants viewed a total of 120 trials. In runs R1-4, we adjusted the SAR dynamically based on a staircase procedure (Figure 1A; see Gekas et al. for a similar approach that manipulated the ambiguity of Gabor stimuli by parametrically varying their orientation (Gekas et al., 2019)). To this end, we defined *checkpoint-trials* that occurred in intervals of 10 trials, starting at the 11th trial of each run. At each checkpoint-trial, we computed the number of stimulus-congruent trials in the block of 10 trials preceding the checkpoint-trial. If more than 8 trials within the preceding block were stimulus-congruent (i.e., perceived in congruence with disambiguating sensory information), we decreased the SAR for the upcoming block by 5%. For 8 stimulus-congruent trials, the SAR remained unchanged in the upcoming block. If we observed less than 8 and more than 5 stimulus-congruent trials, we increased the SAR by 5% in the upcoming block. For less than 6 stimulus-congruent trials, we increased the SAR by 10% in the upcoming block. Run R1 started at an initial SAR of 100%. Runs R2-4 started at the final SAR obtained in the preceding run. During the final runs R5 and R6, we fixed the SAR to the average SAR obtained during runs R1-4.

1.5 Analyses

As dependent variables-of-interest, we computed the proportion of stimulus-congruent trials (i.e., trials perceived in congruence with disambiguating sensory information) and history-congruent trials (i.e., trials perceived in congruence with the immediately preceding percept). At every trial, we recorded the specific perceptual response (left, right, unclear or error) and response time (difference between the button-press indicating the percept and trial onset). *Directed biases* in perception were assessed via the probability of trials perceived as rotating to the right (ranging from 0 to 100%). We computed *absolute biases* by taking the absolute difference between the probability of trials perceived as rotating to the right and chance level at 50%. For summary statistics, we computed the dependent variables within runs R1-6 or for levels of SAR, respectively, and averaged across participants. For dynamic analyses, we computed the dependent variables at each trial within a sliding window of $\pm$ 5 trials. Trials were allocated to the *internal mode* of perceptual processing if the sliding probability of history-congruent percepts was above the sliding probability of stimulus-congruent percepts (vice versa for *external mode*). *Intermediate mode* was designed as a rest category accounting for the fraction of trials where the sliding probabilities of history- and stimulus-congruent percepts were equal (see Supplementary Figure 1C for a representative time course). In 6 participants, we detected runs in which no mode-switch occurred. These runs were excluded when computing the average duration between mode-switches.

Statistical procedures were carried in *R* (summary statistics) and *Matlab* (computational modeling). We conducted group-level pair-wise comparisons using two-sided paired t-tests. Differences from chance-level were evaluated using two-sided one-sample t-tests. We performed correlative analyses using Pearson correlation. We applied the R-method *glm* with a binomial link-function for logistic regression and used the R-packages *lmer* and *afex* for linear mixed effects modeling.

Bayes factors were computed using the R-package *BayesFactor*, using the function *ttestBF* and *lmBF* for linear models. For t-tests, we placed a noninformative Jeffreys prior on the variance of the normal population and a Cauchy prior (rscale = 0.71) on the standardized effect size. Linear models used g-priors (fixed effects: rscale = 0.71; random effects = 1). To obtain Bayes factors for main effects and interactions, we estimated full and reduced models and divided the respective Bayes Factors.

1.5.1 Logistic regression and simulation analyses

In simulation analyses, we asked whether logistic regression reproduced the overall probability of stimulus- and history-congruent percepts. Moreover, we used these simulations to test whether the Markovian assumption of logistic regression (i.e., that the percept at trial t depends exclusively on the current sensory information at trial t and the percept at the immediately preceding trial *t-1*) could explain the observed fluctuations between external and internal modes of perceptual processing. Prior to simulation analysis, we estimated logistic regression models that predicted the perceptual response p at each trial t based the dependent variables h (perceptual history) and d (disambiguating sensory information):

$$p(t) = \beta_h * h(t) + \beta_d * d(t) \tag{1}$$

The dependent variable *perceptual history* (*h(t)*) was defined by the perceptual response *p(t)* (0: leftward rotation; 1: rightward rotation) at the preceding trial:

$$h(t) = p(t-1) \tag{2}$$

The dependent variable *Disambiguating sensory information* (*d(t)*) was defined by a linear transform of the *Direction of disambiguation* (*DIR*, 0: leftward rotation; 1: rightward rotation) and the signal-to-ambiguity ratio (*SAR*, ranging from 0 to 100%) at trial t:

$$d(t) = 0.5 + (DIR(t) - 0.5) * SAR/100 \tag{3}$$

By setting either β_h or β_d to zero, we created reduced logistic regression models that were compared based on Akaike Information Criterion (AIC). As indicated by equation (1), none of the logistic regression models contained an interaction term. For simulation, we used the full logistic regression model, with *SAR* set to the individual threshold SAR used in run R5 and 6. *DIR* was chosen at random for every simulated trial. In analogy to the actual experiment, we simulated 120 trials per run. The total number of simulated runs amounted to 1000 for each participant.

1.5.2 Computational modeling

We constructed all models using the Hierarchical Gaussian Filter toolbox (Mathys et al., 2014) as implemented in the HGF 4.0 toolbox (distributed within the TAPAS toolbox; https://www.tnu.ethz.ch/de/software/tapas). At each trial t, the possible perceptual states y were coded as

$$y(t) = \begin{cases} 1: & \rightarrow \quad (rotation) \\ 0: & \leftarrow \quad (rotation) \end{cases} \tag{4}$$

The input to the model u was provided a linear combination of the direction of disambiguation (DIR) and the signal-to-ambiguity ratio (SAR):

$$u(t) = 0.5 + (DIR(t) - 0.5) * SAR/100 \tag{5}$$

To predict the participants' trial-wise perceptual responses, we combined input u with the prior probability of the perceptual states $\hat{\mu_1}(t)$ into the first-level posterior μ_1.

$$\eta_1(t) = exp(-(u(t) - 1)^2/(2 * \alpha)) \tag{6}$$

$$\eta_0(t) = exp(-(u(t))^2/(2 * \alpha)) \tag{7}$$

$$\mu_1(t) = \frac{\hat{\mu_1}(t) * \eta_1(t)}{\hat{\mu_1}(t) * \eta_1(t) + (1 - \hat{\mu_1}(t)) * \eta_0(t)} \tag{8}$$

In these equations, the influence of disambiguating sensory information on perception scales with the sensitivity parameter α, which was estimated as a free parameter in all models. When α approaches zero, $\mu_1(t)$ is close to the binary values of $u(t)$ (i.e., 0: stimulation with

3D-information for leftward rotation; 1: rightward rotation), signaling high sensitivity to sensory information. Conversely, for α increasing toward infitiy, $\mu_1(t)$ is close to $\hat{\mu_1}(t)$ (see below), reflecting low sensitivity to sensory information.

The influence of perceptual history, in turn, is represented by $\hat{\mu_1}(t)$. The value of $\hat{\mu_1}(t)$ depends on the dynamic accumulation of history effects in μ_2 (i.e, the estimated prior probability of perceptual states represented at the second level of the HFG), which represents the tendency of the first level posterior towards $\mu_1(t) = 1$. For higher values of κ, the prior probability of perceptual states μ_2 has a stronger impact on $\hat{\mu_1}(t)$. The influence of perceptual history on the participants' experience therefore scales with κ:

$$\hat{\mu_1}(t) = s(\kappa * \mu_2(t-1)) \tag{9}$$

Importantly, the models considered in this manuscript differ with respect to the computation of μ_2 (Dimension 1) and κ (Dimension 2).

1.5.2.1 Dimension 1

For models $M_{Learning+/Oscillation-}$ and $M_{Learning+/Oscillation+}$, μ_2 is updated via precision-weighted prediction errors that are generated by the sequence of perceptual states:

$$\mu_2(t) = \hat{\mu_2}(t) + \frac{1}{\pi_2(t)} * \delta_1(t) \tag{10}$$

$$\hat{\mu_2}(t) = \mu_2(t-1) \tag{11}$$

The precision of the second-level representation of perceptual history is governed by $\pi_2(t)$

$$\pi_2(t) = \hat{\pi_2}(t) + \frac{1}{\hat{\pi_1}(t)} \tag{12}$$

The difference between the first level perceptual prediction $\hat{\mu_1}(t)$ and the first-level posterior $\mu_1(t)$ yields the prediction error $\delta_1(t)$:

$$\delta_1(t) = \mu(t) - \hat{\mu}_1(t) \tag{13}$$

$\delta_1(t)$ is combined with the second level precision π_2, yielding the precision-weighted prediction error $\epsilon_2(t)$, which updates the second level prediction $\hat{\mu}_2(t)$:

$$\epsilon_2(t) = \frac{1}{\pi_2} * \delta_1(t) \tag{14}$$

In addition to κ and α, $M_{Learning+/Oscillation+}$ and $M_{Learning+/Oscillation-}$ incorporate a learning rate ω. This free parameter determines how swiftly μ_2 is updated in response to predicition errors, thereby controlling the speed at which the second-level precision $\hat{\pi_2}(t)$ changes over time.

$$\hat{\pi_1}(t) = \frac{1}{\hat{\mu_1}(t) * (1 - \hat{\mu_1}(t))} \tag{15}$$

$$\hat{\pi_2}(t) = \frac{1}{\frac{1}{\pi_2(t)} + \exp(\omega_2)} \tag{16}$$

By contrast, for models $M_{Learning-/Oscillation-}$ and $M_{Learning-/Oscillating+}$, $\mu_2(t)$ is defined by the immediately preceding perceptual state:

$$\mu_2(t) = \begin{cases} 1: & y(t) = 1 \\ -1: & y(t) = 0 \end{cases} \tag{17}$$

Thus, $M_{Learning-/Oscillation-}$ and $M_{Learning-/Oscillating+}$ do not incorporate any second-level accumulation of perceptual history and are thus governed only by the parameters κ and α.

1.5.2.2 Dimension 2

For models $M_{Learning-/Oscillation-}$ and $M_{Learning+/Oscillating-}$, κ is estimated as a stable parameter. By contrast, for models $M_{Learning-/Oscillation+}$ and $M_{Learning+/Oscillating+}$, κ fluctuates dynamically according to the frequency parameter f (in $nb\ trials^{-1}$), the phase parameter p and the amplitude parameter amp

$$\kappa = \frac{(amp * sin(f * t + p) + 1)}{2} \tag{18}$$

1.5.3 Model inversion

We used a free energy minimization approach for model inversion (Friston, 2010), maximizing a lower bound on the log-model evidence for the individual participants' data. Parameters were optimized using quasi-Newton Broyden-Fletcher-Goldfarb-Shanno minimization. Parameters were inverted using the following priors:

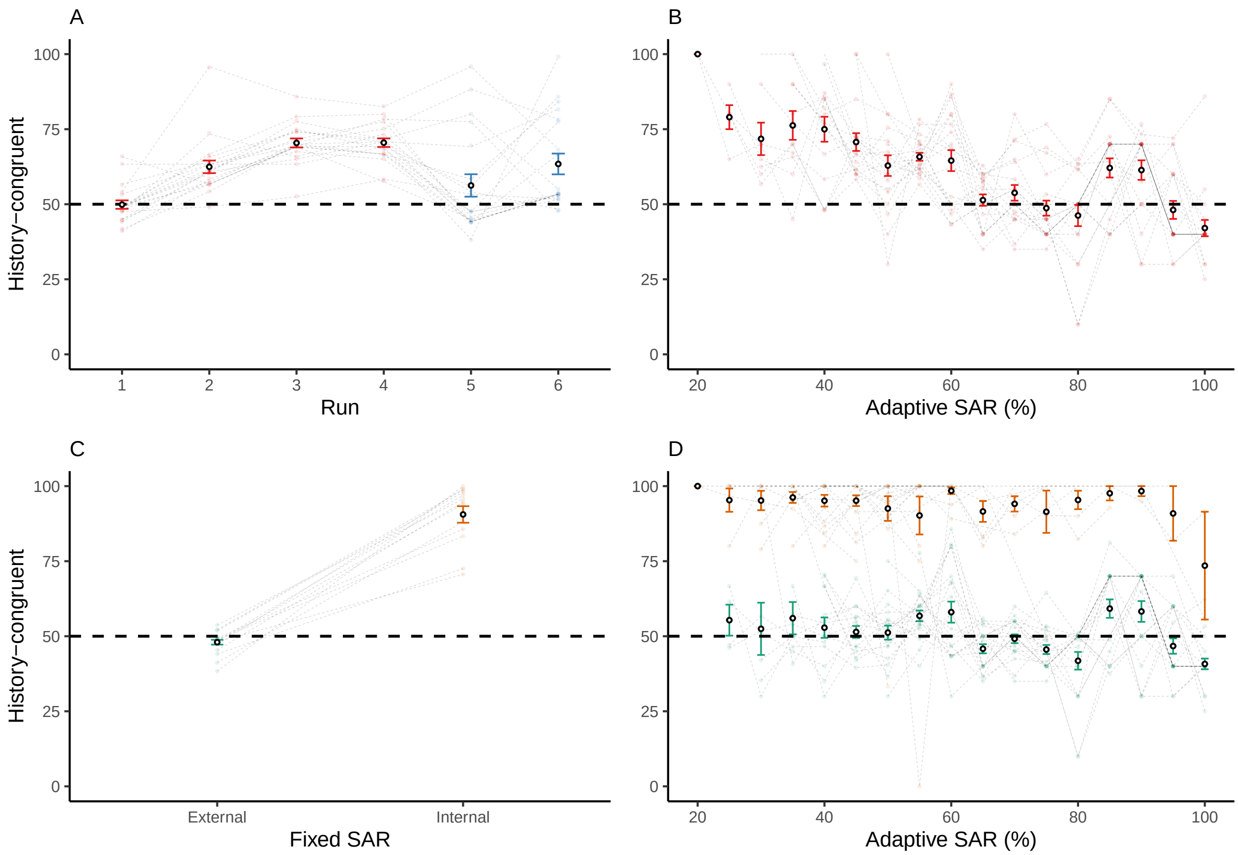
- Dimension 1: κ = prior mean of log(1) and prior variance of 1; α = prior mean of log(0.1) and prior variance of 1; ω = prior mean of 0 and prior variance of 16.
- Dimension 2: α = prior mean of log(0.1) and prior variance of 1; f = prior mean of log(0.1 and prior variance of 0.1; p = prior mean of $\pi/2$ and prior variance of $\pi/2$; amp = prior mean of log(1) and prior variance of 1.

1.5.4 Model-level inference

Models were compared using random-effects Bayesian model selection (Stephan et al., 2009) as implemented in SPM12 (http://www.fil.ion.ucl.ac.uk/spm/software/spm12). We report protected exceedance probabilities for group-level inference and individual exceedance probabilities at the participant-level.

2 Supplementary Figures

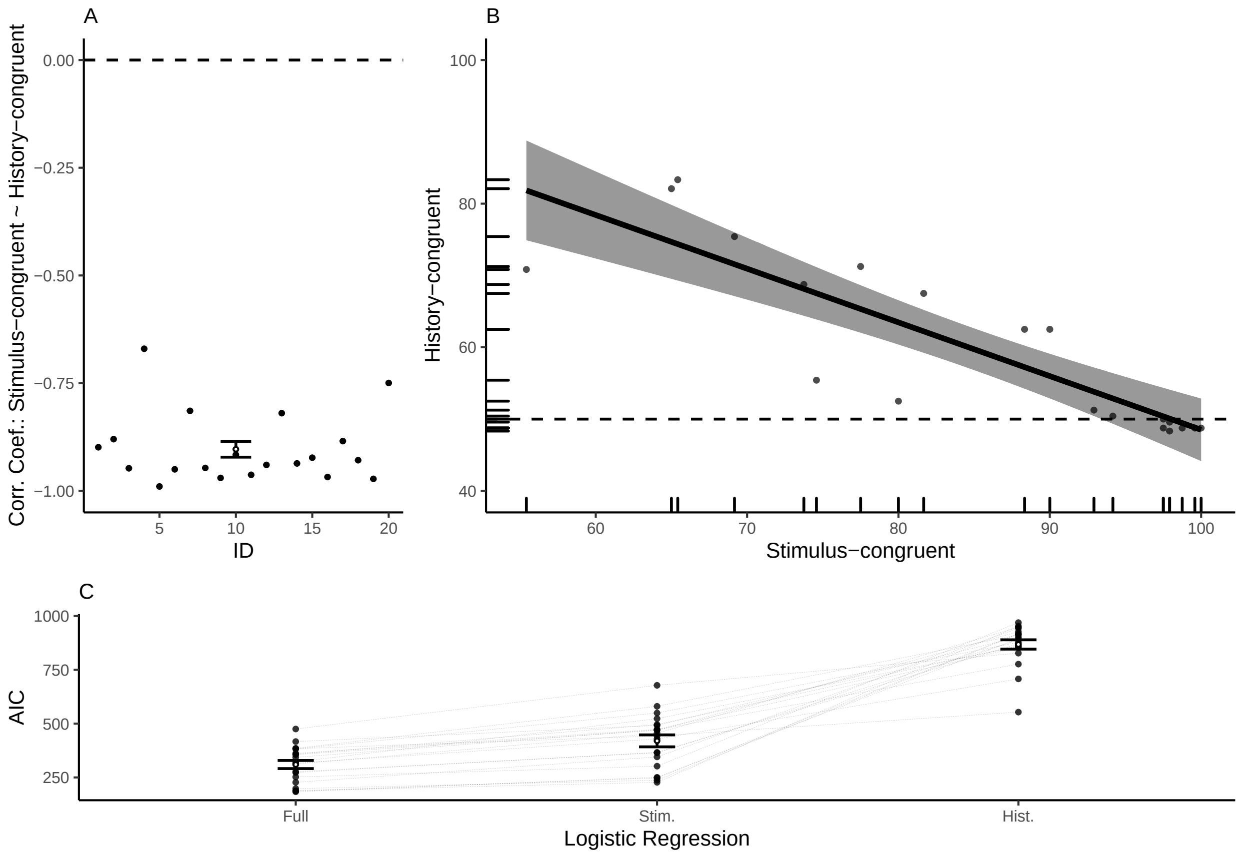
2.1 Supplementary Figure S1



Supplementary Figure S1. A. Perceptual history across runs. Related to Figure 1 and 2. As the SAR was dynamically adjusted in runs R1-4 (shown in red), we observed a progressive increase in the frequency of history-congruent percepts (F(3, 57) = 57.96, p $= 2.58 \times 10^{-17}$, $BF_{10} = 9.69 \times 10^{15}$; R1: 49.88 ± 1.41%; R2: 62.46 ± 2.1%; R3: 70.42 ± 1.5%; R4: 70.5 ± 1.43%). In runs with fixed SAR (R5-6, depicted in blue), history-congruent percepts amounted to 56.25 ± 3.74% in R5 and 63.42 ± 3.46% in R6. **B. Perceptual history across levels of SAR.** As expected, perceptual history had a stronger influence on perception at lower levels of SAR (F(1, 265.07) = 181.5, p $= 7.25 \times 10^{-32}$, $BF_{10} = 5.2 \times 10^{28}$, main effect of SAR) and ranged at chance-level when disambiguating sensory information was strong. **C. Perceptual history during internal and external mode for SARs at**

threshold. During internal mode, the frequency of history-congruent percepts approached 100% (90.57 ± 2.76%), but was reduced below chance level during external mode (48 ± 0.8%; T(19) = -2.51, p = 0.02, $BF_{10} = 2.74$, one-sample t-test). **D. History-congruent percepts during internal and external mode across the full range of SAR.** Linear mixed effects modeling indicated that frequency of history-congruent percepts was significantly affected by the factor *mode* (green: external; yellow: internal; F(2, 484.03) = 23.87, p = 1.3×10^{-10}, $BF_{10} = 2.43 \times 10^{70}$) and showed a trend for an effect of *SAR* (F(1, 188.7) = 3.42, p = 0.07, $BF_{10} = 1.1$). We observed no between-factor interaction with respect to history-congruent percepts (F(2, 469.91) = 0.07, p = 0.93, $BF_{10} = 0.05$). Please note that any main effect of *mode* was expected, since external and internal mode were defined based on the dynamic probability of stimulus-congruence. Pooled data are represented as mean ± SEM.

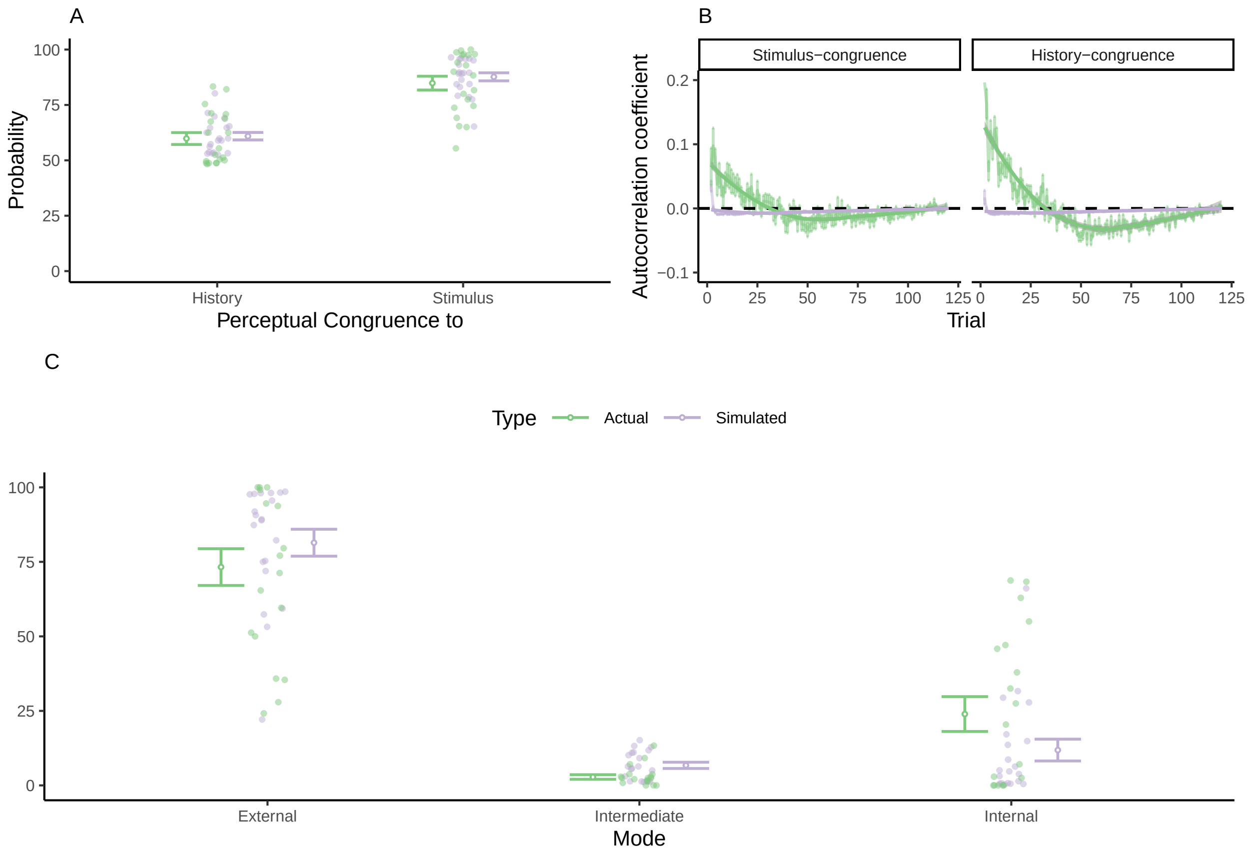
2.2 Supplementary Figure S2



Supplementary Figure S2. A. Within-participant correlations of stimulus- and history-congruent percepts. Related to Figure 1 and 2. In individual participants, Pearson correlation coefficients between the frequencies of stimulus- and history-congruent percepts (runs R5-6; fixed SAR) amounted to -0.9 ± 0.02 (T(19) = -49.25, p = 1.66×10^{-21}, $BF_{10} = 1.34 \times 10^{18}$, one-sample t-test). This strong inverse relationship suggested that perceptual history and disambiguating sensory information compete with each other to determine conscious experience. **B. Across-participants correlation of stimulus- and history-congruent percepts.** Inter-individual differences in the frequency of history-congruent percepts strongly predicted the frequency of stimulus-congruent percepts (ρ = -0.77, $p = 7.2 \times 10^{-5}$, $BF_{10} = 203.27$, Pearson correlation for runs R5-6). This negative association indicated that, overall, perceptual history had a stronger impact in participants who were less sensitive to disambiguating sensory information. **C. Predicting perceptual**

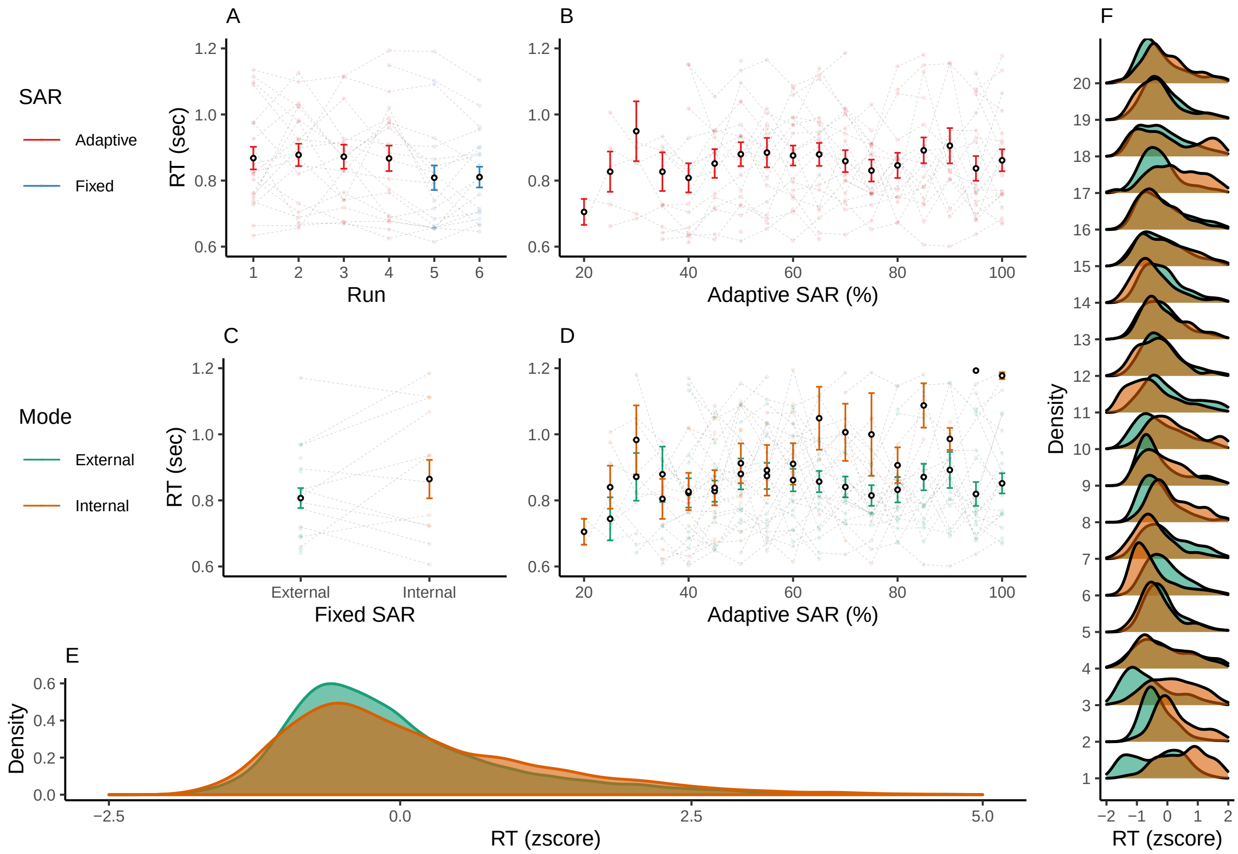
responses using logistic regression. In each participant, the Akaike Information Criterion (AIC) of logistic regression models based on both disambiguating sensory information and perceptual history (309.95 ± 18.81) was lower than the AIC for models based on sensory information only (419.95 ± 27.84; T(19) = -9.39, p = 1.45×10^{-8}, $BF_{10} = 8.89 \times 10^{5}$, paired t-test) or perceptual history only (867.86 ± 21.58; T(19) = -16.46, p = 1.06×10^{-12}, $BF_{10} = 6.54 \times 10^{9}$). Pooled data are represented as mean ± SEM.

2.3 Supplementary Figure S3



Supplementary Figure S3. A. Simulating the overall frequencies of stimulus- and history-congruent percepts with logistic regression. Related to Figure 1 and 2. We estimated logistic regression models based on both disambiguating sensory information and perceptual history in individual participants and used the regression weights to simulate perceptual responses. These simulations revealed that logistic regression reproduced the overall frequencies of history-congruent percepts observed in the actual experiment (simulated data in purple: 60.9 ± 1.69%; actual data in light green: 59.83 ± 2.69%; T(19) = 0.78, p = 0.44, BF_{10} = 0.31, paired t-test) as well as the overall frequency of stimulus-congruent percepts (simulated: 87.69 ± 1.81%; actual data: 84.85 ± 3.12%; T(19) = 1.48, p = 0.16, BF_{10} = 0.59). **B. Simulated autocorrelations of stimulus- and history-congruence.** When simulating perceptual responses from logistic regression, we detected no autocorrelation of stimulus- or history-congruence. Real trial-wise autocorrelation coefficients are plotted for comparison. **C. Simulating the relative proportions of external, internal and intermediate modes with logistic regression.** Likewise, logistic regression did not reproduce the relative proportion of trials spent in external mode (simulated: 81.43 ± 4.52%; actual: 73.25 ± 6.17%; T(19) = 2.75, p = 0.01, $BF_{10} = 4.17$, paired t-test), internal mode (simulated: 11.85 ± 3.66%; actual: 23.94 ± 5.84%; T(19) = -3.49, p = 2.44×10^{-3}, $BF_{10} = 16.92$) and intermediate mode (simulated: 6.72 ± 1.05%; actual: 2.81 ± 0.77%; T(19) = 3.73, p = 1.41×10^{-3}, $BF_{10} = 27.07$). Pooled data are represented as mean ± SEM.

2.4 Supplementary Figure S4



Supplementary Figure S4. A. RTs across runs. Related to Figure 1 and 2. In runs R1-4 (depicted in red), we adapted the SAR based on a staircase procedure, which did not affect RTs (R1: 0.87 ± 0.03 sec; R2: 0.88 ± 0.03 sec; R3: 0.87 ± 0.04 sec; R4: 0.87 ± 0.04 sec). In runs R5-6 (depicted in blue), the SAR was fixed to the average SAR from the preceding runs R1-4 (60.25 ± 2.36 sec). RTs amounted to 0.81 ± 0.04 sec in R5 and 0.81 ± 0.03 sec in R6. **B. RTs across levels of SAR.** Globally, the level of disambiguating sensory information did not have a significant effect on RT (F(1, 261.5) = 0.05, p = 0.82, BF_{10} = 0.15, main effect of SAR). **C. RTs during internal and external mode for SAR at threshold.** In Runs R5 and R6, RTs did not differ between external and internal mode (T(12) = 0.74, p = 0.48, BF_{10} = 0.35, paired t-test): **D. RTs during internal and external mode across the full range of SAR.** Linear mixed effects modeling indicated that, during internal mode, response times increased significantly for higher levels of SAR

(F(2, 476.5) = 10.73, p = 2.77×10^{-5}, $BF_{10} = 538.42$, *mode* x *SAR* interaction), driving a main effect of *SAR* in this analysis (F(1, 488.29) = 21.98, p = 3.57×10^{-6}, $BF_{10} = 1.73$). Response times were longer during internal mode (F(2, 474.05) = 5.28, p = 5.39×10^{-3}, $BF_{10} = 18.9$, main effect of *mode*). **E. Collapsed RTs.** During both internal and external mode, normalized RTs were better explained by a log-normal distribution (internal mode: AIC = 1.07×10^4, external mode: AIC = 2.79×10^4) as compared to a Gaussian distribution (internal mode: AIC = 1.08×10^4, external mode: AIC = 2.81×10^4), a gamma distribution (internal mode: AIC = 1.08×10^4, external mode: AIC = 2.8×10^4) or a Weibull distribution (internal mode: AIC = 1.24×10^4, external mode: AIC = 3.39×10^4). **F. Individiual RT distributions.** Within individual participants (y-axis), the distributions of normalized RTs were largely overlapping between internal and external mode. Pooled data are represented as mean ± SEM.