Supplementary Text: Rescorla-Wagner model on discrimination learning task

We interpret the discrimination task as a measure of value learning, specifically tracking how participants update their weights (w) for each stimulus pair over trials. Learning is assumed to start at 0.5, reflecting chance-level knowledge about which stimulus leads to reward, and progresses toward 1 as participants acquire perfect knowledge. Each Day 1 session includes five runs of 24 trials, comprising 12 distinct stimulus pairs. Each pair is presented twice per run, with left/right positions counterbalanced.

We specified and estimated three hierarchical learning models to characterize value updating during the discrimination task:

1. **Session-specific model (by-Sess):** Learning rates vary by session number (1, 2, 3).

   - Trial-level learning rates for each stimulus pair, denoted $\alpha_{j,c,p}$, are modeled as:

     $$\alpha_{j,c,p} \sim \text{Beta}(a_{j,c} \cdot k_j,\ (1 - a_{j,c}) \cdot k_j),$$

     where:
     - j indexes subjects,
     - c indexes sessions (1, 2, 3),
     - p indexes cue pairs (1 to 12),
     - $a_{j,c}$ is the subject- and session-specific mean learning rate,
     - $k_j \sim \text{Gamma}(1, 0.1)$ for each subject $j = 1, \ldots, n_{\text{subs}}$.
   - The higher-level learning rate means are drawn from:

     $$a_{j,c} \sim \text{Beta}(\mu_c \cdot \kappa,\ (1 - \mu_c) \cdot \kappa),$$

     where:
     - $\mu_c \sim \text{Beta}(8, 2)$ is the session-specific population mean,
     - $\kappa \sim \text{Gamma}(1, 0.1)$ controls the overall precision.

   For any stimulus pair used on trial i, the weight w is updated as:

   $$w_{i+1} = w_i + \alpha \cdot (1 - w_i),$$

   where α denotes the learning rate on that trial (indices omitted for brevity).

   The discrimination response (Resp_i) is modeled as:

   $$\text{Resp}_i \sim \text{Bernoulli}(w_i),$$

   where $\text{Resp}_i = 1$ if the odor-predictive stimulus was chosen, and 0 otherwise.

2. **TMS-condition model (by-TMS):** Learning rates vary by Day 1 TMS condition (sham vs. cTBS), instead of session number. The same hierarchical structure is applied, with TMS condition replacing session as the grouping variable.

3. **Constant-rate model:** Each subject has a single learning rate shared across all sessions and cue pairs. This model assumes no condition- or session-specific modulation of learning.