Supplementary Information for

A large-scale analysis of task switching practice effects across the lifespan

Mark Steyvers, Guy Hawkins, Frini Karayanidis, Scott Brown

Mark Steyvers.
E-mail: mark.steyvers@uci.edu

This PDF file includes:

Supplementary text
Figs. S1 to S8
References for SI reference citations

www.pnas.org/cgi/doi/10.1073/pnas.1906788116

Supporting Information Text

Methods

The code and data are available at https://osf.io/sxr5f

Procedure. On each trial, a set of leaves is displayed that point in one of four directions (left, right, up or down) and are moving in one of four directions (left, right, up, or down). Users have to shift focus between attending to the pointing or moving dimension depending on the task cue indicated by the colors of the leaves as well as a highlighted verbal label on the bottom of the screen. When the leaves are green, the user has to determine the direction in which the leaves are pointing (up, down, left or right) and respond accordingly. When the leaves are orange, the user has to respond based on the moving direction (up, down, left or right). Each of the four movement or pointing directions are equally likely to occur. On congruent trials (50% of trials), the orientation and direction of movement of the stimuli lead to the same response. Incongruent trials occur when the orientation and direction of movement are different, requiring the user to express behavior associated with one rule and inhibit behavior associated with the other rule. After the user responds, feedback is provided and the next trial is presented without delay. In the raw data, any trial with a response time higher than 5 seconds was coded as 5 seconds. The task cue is presented at the same time as the target display. After a task switch occurs, the game can switch after any number of trials. The maximum recorded run length of any rule was 19 trials although most run lengths are much shorter. Run lengths up to 10 trials covered 97% of cases. Among run lengths up to 10, the percentage of run lengths of 1 up 10 were 6.40%, 10.58%, 13.51%, 14.76%, 14.33%, 12.58%, 10.04%, 7.298%, 4.82%, and 5.68% respectively, indicating that the game is designed to favor run lengths between four and six trials.

Each gameplay event has a fixed duration of 60 seconds. At the end of each gameplay event, users are provided feedback on mean response time per trial, mean accuracy, and a score that is based on the total number of correct trials completed within the fixed time period as well as bonus points based on a variety of factors (e.g. streaks of correct responses). The total score is the focal point on the feedback screen, so it can be assumed that the conditions foster a combination of speed and accuracy.

Model details

Decision Process. In this subsection, we explain how the joint distribution of response time and response choice is specified in the model, which is needed for estimation. Note that parameter estimation is informed by the choice and RT data at the individual trial level.

Let x_k represent the variable that is accumulating evidence for response k. In a standard Wiener diffusion process, the change in evidence can be described by:

$$dx_k = c_k dt + \sigma dW \quad [1]$$

where t is time elapsed since the start of the trial, c_k is a drift rate given by the weighted summed perceptual evidence, σ is the standard deviation of zero-mean Gaussian distributed noise and dW is a Wiener process (Brownian motion).

For each individual accumulator described by Eq 1, the probability density for the first time to reach threshold α and make a response is given by an inverse Gaussian (Wald) distribution:

$$f_k(t|\mu_k, \lambda_k) = \left(\frac{\lambda_k}{2\pi t^3}\right)^{1/2} \exp\left(\frac{-\lambda_k(t-\mu_k)^2}{2\mu_k^2 t}\right) \quad [2]$$
$$t \sim \text{InverseGaussian}(\alpha/\nu, \alpha^2/\sigma^2)$$

where the mean and shape parameters μ and λ can be parameterized in terms of drift rate c_k, threshold α, and standard deviation of the evidence accumulation noise σ by:

$$\mu_k = \frac{\alpha}{c_k}, \ \lambda_k = \frac{\alpha^2}{\sigma^2} \quad [3]$$

In a race between K accumulators, the likelihood $g_k(t)$ that accumulator k with drift rate c_k will win the race at time t (and respond with response alternative k) is based on the density of the inverse Gaussian associated with that accumulator multiplied by the "survival" probability – the probability that none of the other accumulators have reached threshold before time t:

$$g_k(t) = f_k(t) \prod_{k' \neq k} (1 - F_{k'}(t)) \quad [4]$$

where $F_k(t)$ is the cumulative distribution function of $f_k(t)$.

Contaminant model. To explain some of the very fast responses that are unlikely to be generated from the task-switching decision process, we model the observed decisions as a mixture of the decisions from the task-switching model and a contaminant response process. This process provides a low, uniform density over the range of possible response times (0-5 seconds) and a random response over the number of response alternatives:

$$(1-p)g_k(t) + (p/K)\frac{1}{5} \quad [5]$$

where $g_k(t)$ is the density of the observed response k at time t according to the model, p is the probability of a contaminant response, and K is the number of accumulators ($K = 4$ for our data). Based on previous research, we make a simplifying assumption that there is a pre-specified proportion p of contaminant trials and set at $p = .05$.

Simulating the task activation function. As stated in the main body of the paper, the changes in task activations $w_{d,i}$ for task d, $d \in \{cued, uncued\}$, at trial i is described by exponential growth functions with the difference equations:

$$w_{d,i} = \begin{cases} w_{d,i-1} + (v - w_{d,i-1}) \cdot g & \text{if } d = cued \\ w_{d,i-1} + (u - w_{d,i-1}) \cdot h & \text{if } d = uncued \end{cases} \quad [6]$$

where u and v represent baseline and asymptote task activation values and g and h are rate parameters. It is computationally convenient to calculate the dynamics of task activation within a run without difference equations:

$$\begin{aligned} w_{c,i,r} &= w_{u,f(r-1),r-1} + (v - w_{u,f(r-1),r-1})(1 - e^{-gi}) \\ w_{u,i,r} &= w_{c,f(r-1),r-1} + (u - w_{c,f(r-1),r-1})(1 - e^{-hi}) \end{aligned} \quad [7]$$

where $w_{c,i,r}$ and $w_{u,i,r}$ represent the cued and uncued task activations at run r, index i counts the number of trials within a run, and $f(r)$ returns the number of trials associated with run r. Therefore, $w_{c,f(r-1),r-1}$ and $w_{u,f(r-1),r-1}$ represents the task activation of the cued and uncued task at the end of the previous block.

To speed up model simulations, we approximated Eq. 7 by assuming the task activations on any particular run depend only on the previous run and not the run before that. Given this assumption, the task activations at the end of the previous block become:

$$\begin{aligned} w_{c,f(r-1),r-1} &= u + (v - u)(1 - e^{-gf(r-1)}) \\ w_{u,f(r-1),r-1} &= v + (u - v)(1 - e^{-hf(r-1)}) \end{aligned} \quad [8]$$

Speed of task switching. The speed of task switching is assessed by the slope of the task activation function for the first switch trial represented by Δw_c and Δw_u for the cued and uncued tasks respectively:

$$\Delta w_c = -g(u - v)e^{-\frac{1}{2}g} \quad [9]$$

$$\Delta w_u = -h(v - u)e^{-\frac{1}{2}h} \quad [10]$$

Bayesian Inference. Bayesian inference provides posterior estimates of the model parameters on the basis of the observed data which consists of the observed choice and RT data at the individual trial level.

The task switching model has 17 parameters ($\theta_c, \theta_o, \gamma_v, \gamma_u, \tau_u, \tau_v, \gamma_g, \gamma_h, \tau_g, \tau_h, \delta_v, \delta_g, \delta_u, \delta_h, \alpha, \sigma, t_0$). To avoid identifiability issues where parameters arbitrarily trade-off during inference, we set $\tau_v = 1$ to anchor the scale of the model, leaving 16 parameters. We applied this model separately to each participant in each user sample and used Bayesian inference procedures to derive the posterior distribution over model parameters and posterior predictive distributions to get model predictions at the individual trial level.

To complete the model specification, we used the following priors for parameters where TN denotes the Half Normal prior:

$$\begin{aligned} \theta_c &\sim \text{TN}(0, 5) & \theta_o &\sim \text{TN}(0, 5) \\ \alpha &\sim \text{TN}(0, 5) & \gamma_v &\sim \text{Uniform}(0, 10) \\ \tau_u &\sim \text{Uniform}(0, 10) & \gamma_u &\sim \text{Uniform}(0, 10) \\ \tau_g &\sim \text{N}(0, 2) & \gamma_g &\sim \text{N}(0, 2) \\ \tau_h &\sim \text{N}(0, 2) & \gamma_h &\sim \text{N}(0, 2) \\ \delta_v &\sim \text{N}(0, 2) & \delta_u &\sim \text{N}(0, 2) \\ \delta_g &\sim \text{N}(0, 2) & \delta_h &\sim \text{N}(0, 2) \\ \sigma &\sim \text{Uniform}(0, 2) & t_0 &\sim \text{Uniform}(0, 1) \end{aligned}$$

The ranges of these priors were chosen after some initial experiments with the inference procedure to ensure that the likelihoods did not produce numerical instabilities and allowed the model to express a wide variety of data patterns.

The Bayesian inference procedure is based on differential evolution Markov chain Monte Carlo (DE-MCMC) described in (1) to obtain samples from the posterior distribution over parameters. The DE-MCMC procedure allows for efficient Bayesian estimation of posterior distributions when model parameters are correlated. The procedure has been applied recently to estimate response-time model parameters (2–4). We ran the DE-MCMC procedure separately for each participant with 30 chains for 500 iterations and obtained 10 samples from each chain. The chains mixed appropriately. The DE-MCMC procedure was implemented in Matlab. For the sample of 1000 users, the total time to run the DE-MCMC procedure involved around 6 hours of computing after distributing the simulations for each of the 1000 users over 180 CPUs on Amazon's Elastic Cloud.

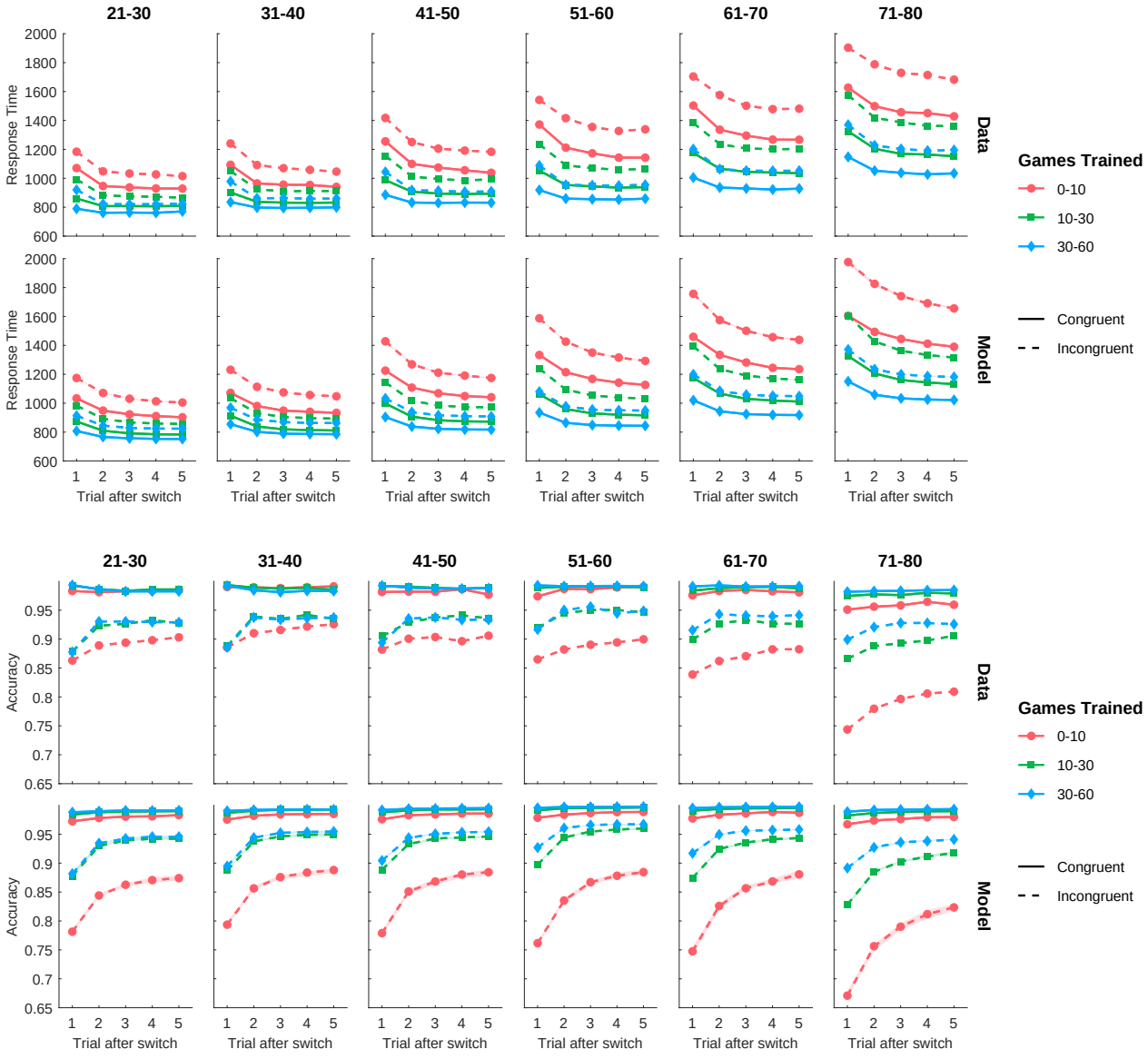

Fig. S1. Average response times and accuracy for observed data and model predictions as a function of the number of trials that occurred after a task switch. Results are broken down across age groups (columns), levels of practice (colors and markers) and congruency (solid line=congruent, dashed line=incongruent condition. Shaded areas for model predictions represent 75% credible intervals.

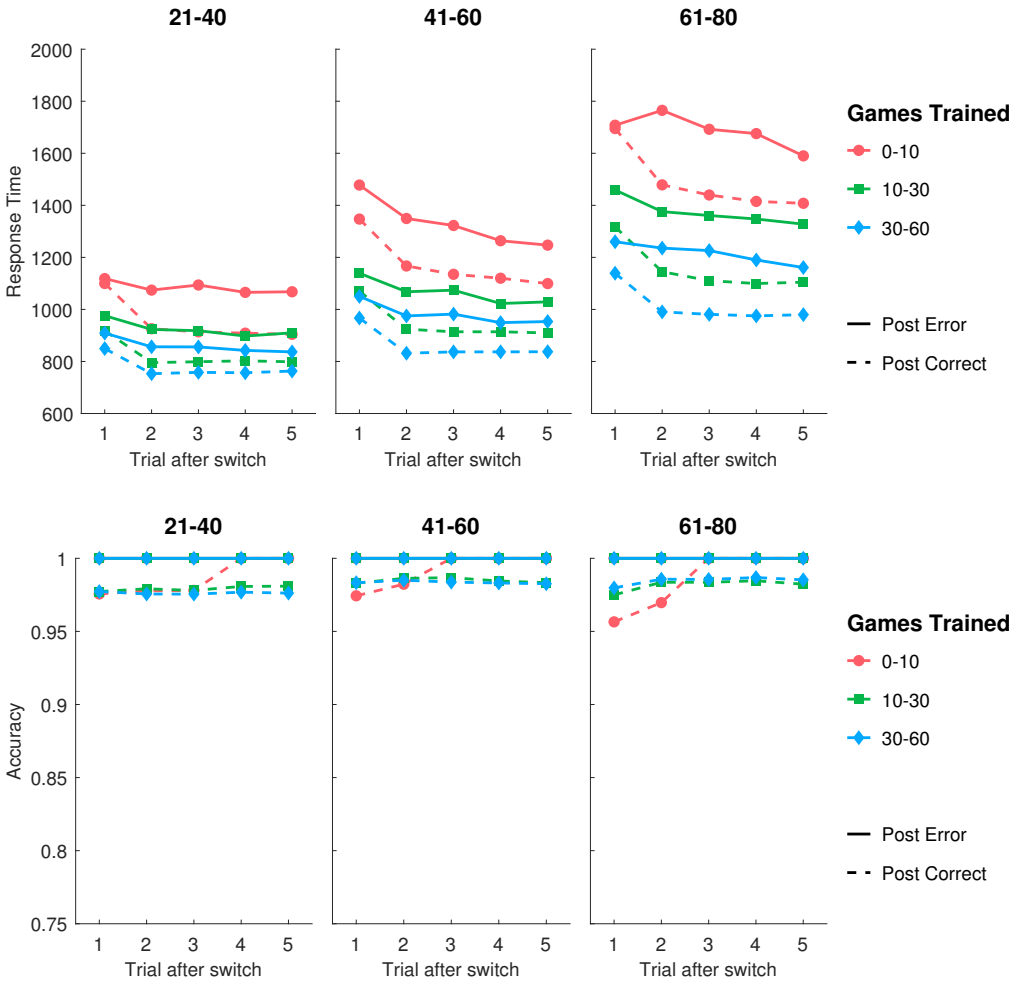

Fig. S2. Results of post-error slowing analysis. Panels show average response times and accuracy for trials following error and correct trials. Results are broken down by age group (columns), practice (colors) and trial within a run (horizontal axis). Results were collapsed over congruency as well as pairs of age groups in order to increase sample size on the post-error trials. Consistent with prior research, response times slow down after errors. Post error, response times are less affected by the position within a run. Following a task switch, the response times following an error are almost as fast as the response times following a correct trial, suggesting that task relevant representations are re-activated following an error.

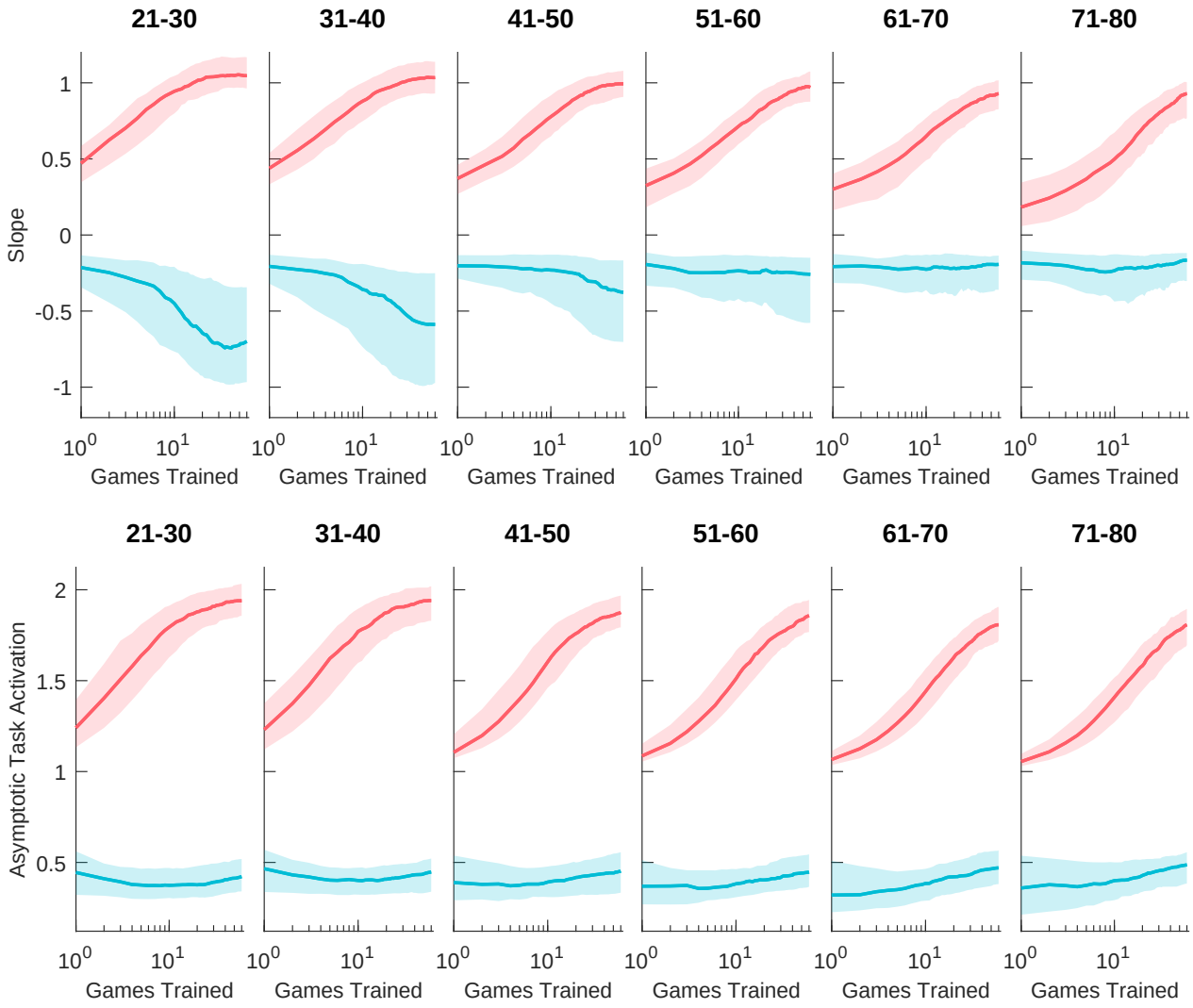

Fig. S3. Effect of practice on the average slope of the task activation function during the first trial (top row) and asymptotic task activation (bottom row). Results are broken down by age groups (columns) and cued and uncued task dimensions (red and blue lines respectively). Shaded areas show the 75% percentiles of the posterior mean across individual participants. Note that the slope is calculated on the basis on time expressed in trials and not actual elapsed time. Therefore, an older person with slower response times but with an equivalent slope relative to a younger person has a slower changes in task activation when evaluated in actual elapsed time.

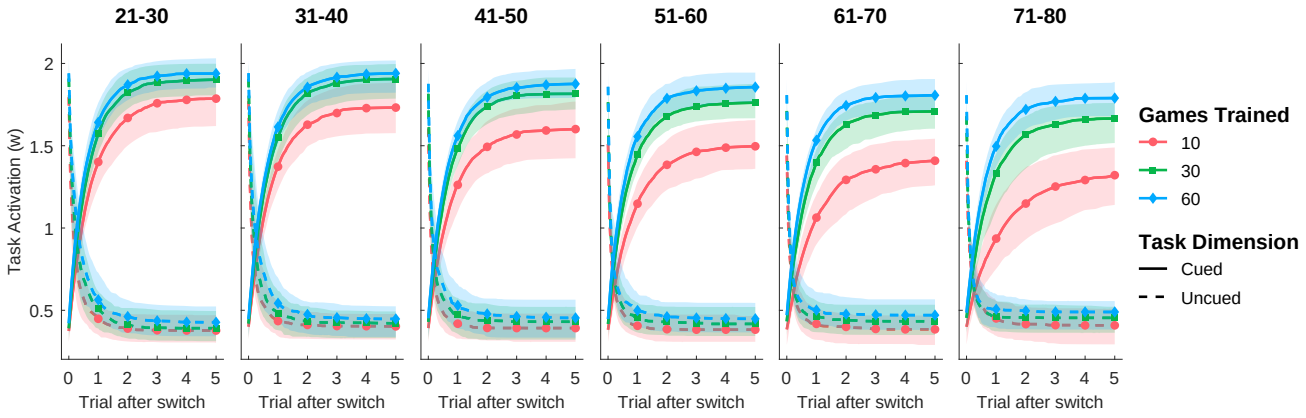

Fig. S4. Inferred dynamics of task activation levels (w) for cued and uncued tasks for each age group. Lines represent the mean across participants. Shaded areas represent 75% percentiles of the posterior mean across individual participants.

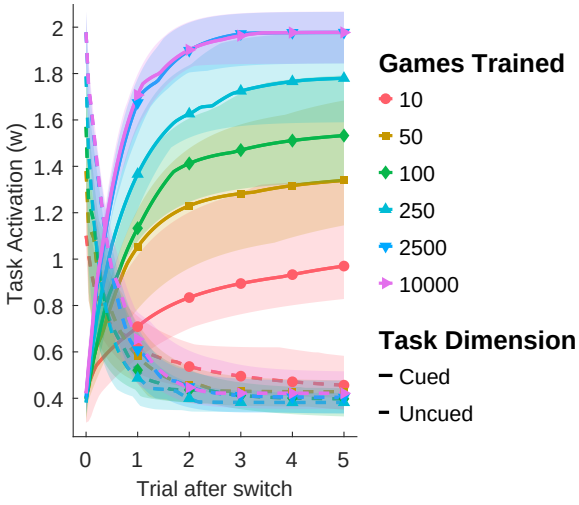

Fig. S5. Inferred dynamics of task activation levels (w) for cued and uncued tasks for a sample of participants with extensive practice. Lines represent the mean across participants. Shaded areas represent 75% percentiles of the posterior mean across individual participants.

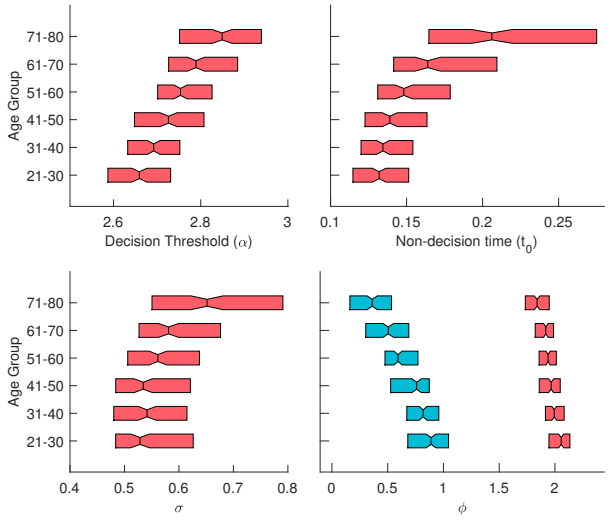

Fig. S6. Individual differences in the decision threshold α (upper left), non-decision time t_0 (upper right), standard deviation of the noise of evidence accumulation σ (lower left) and perceptual evidence for correct (ϕ_c, red, lower right) and incorrect features (ϕ_o, blue, lower right) across age groups. Boxes and notches show the 75% and 50% percentiles, respectively, of the posterior mean across individual participants.

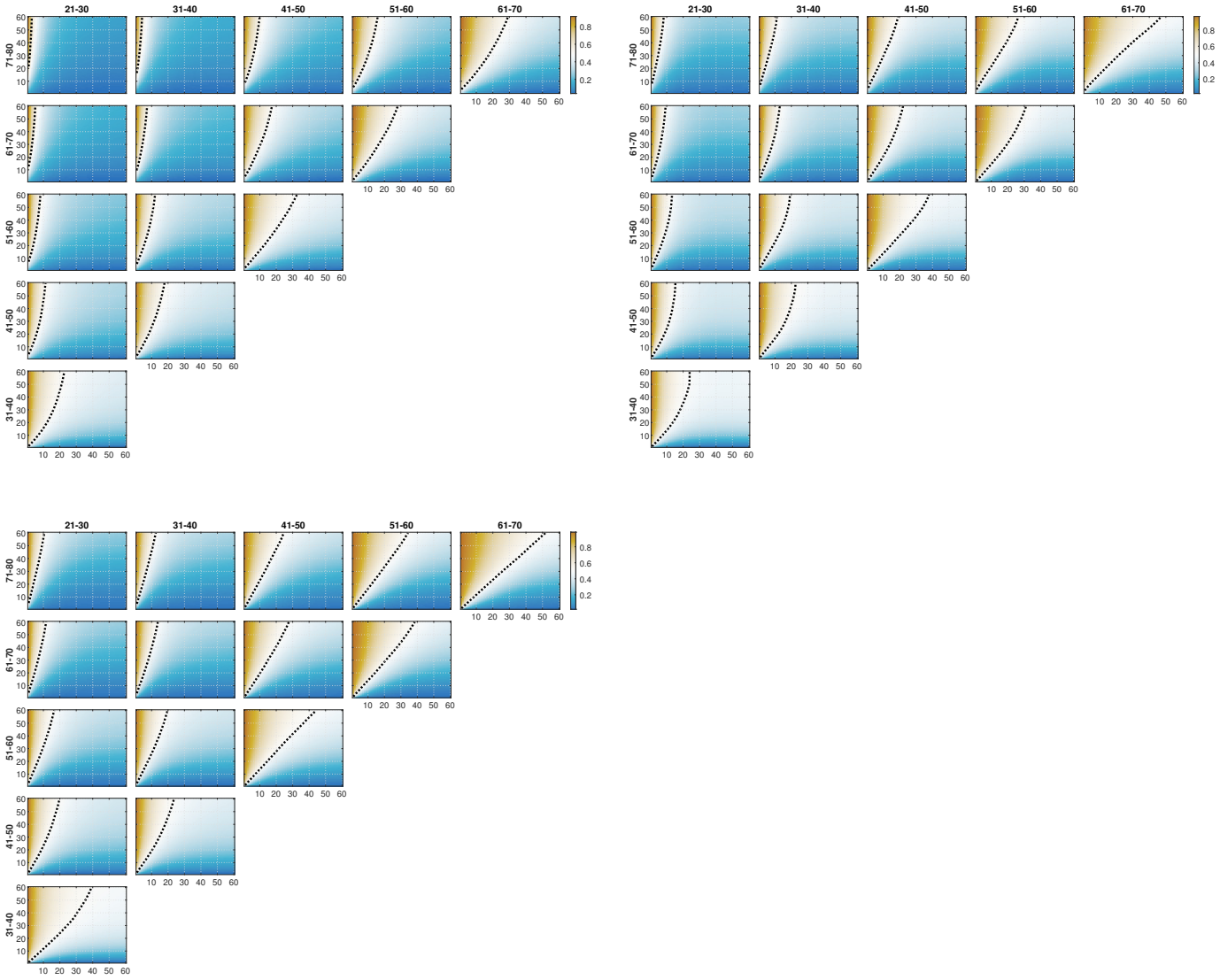

Fig. S7. Performance comparison between two age groups as a function of practice for the older age group (vertical axis) and the younger age group (horizontal axis) for three different performance metrics: drift rate (c, top left), slope of the activation function ($\Delta(w)$, top right) and asymptotic task activation (u, bottom left). Each panel shows the proportion of cases (color) where randomly sampled individuals from the older age group outperform randomly sampled individuals from the other age groups. Dashed lines show the practice levels associated with 50% odds where the age groups perform similarly. Brown colors show practice levels where the older individuals have higher odds of outperforming the younger age groups.

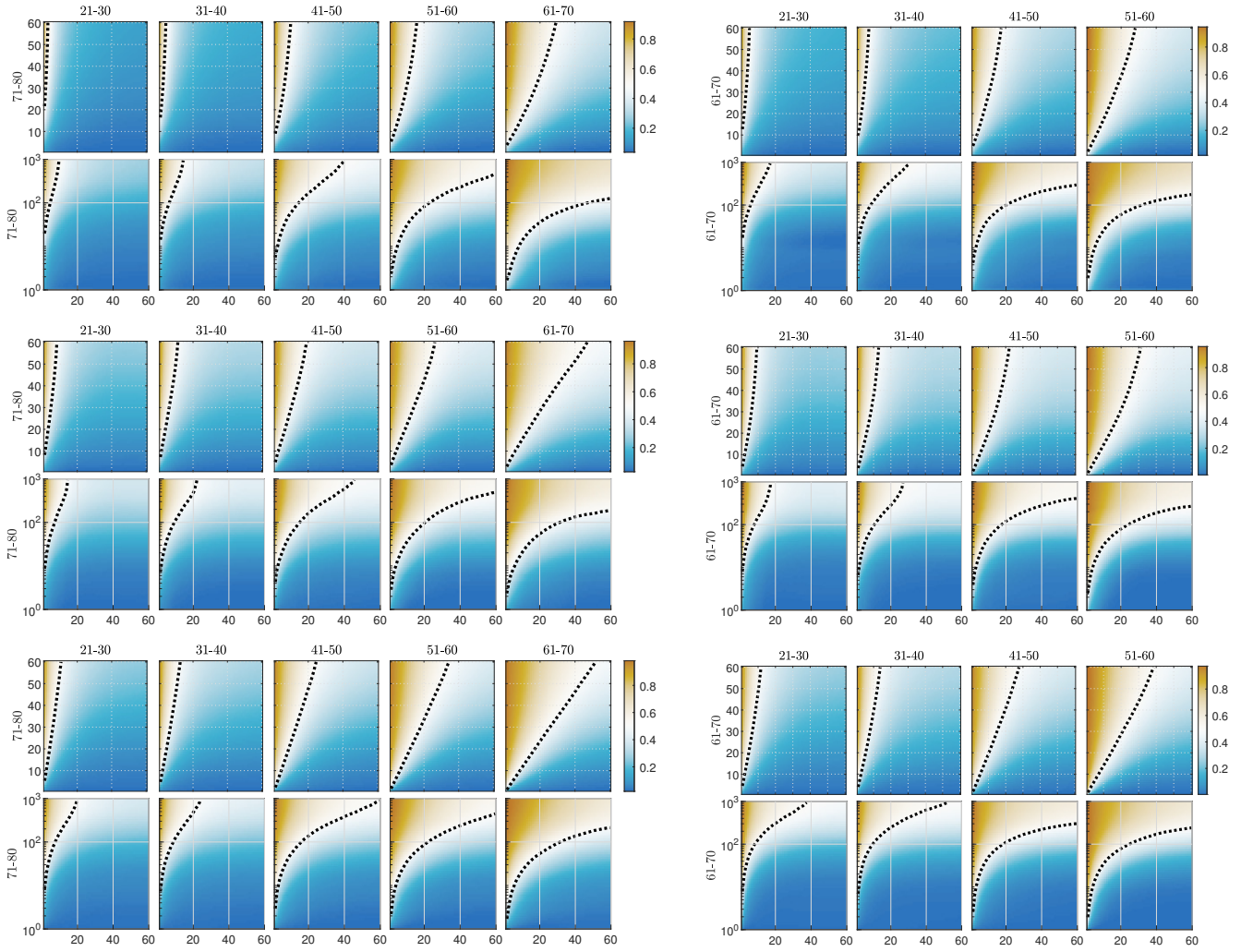

Fig. S8. Performance comparison between older and younger individuals as a function of practice for the older age group (vertical axis) and the younger age group (horizontal axis). Brown (blue) colors show the proportion of cases where randomly sampled older individuals have higher (lower) odds of outperforming randomly sampled younger individuals. Dashed lines show the practice levels resulting in equal odds of an older individual outperforming a younger individual. The top row in each pair show the performance comparisons in the sample of users with training data up to 60 games. The bottom row in each pair compares the performance of older users with training data up to 1000 games with younger users with training data up to 60 games. Results are broken down for three different performance metrics: drift rate (c, top two rows), slope of the activation function ($\Delta(w)$, middle two rows) and asymptotic task activation (u, bottom two rows).

References

1. Turner BM, Sederberg PB, Brown SD, Steyvers M (2013) A method for efficiently sampling from distributions with correlated dimensions. *Psychological methods* 18(3):368.
2. Turner BM, Van Maanen L, Forstmann BU (2015) Informing cognitive abstractions through neuroimaging: the neural drift diffusion model. *Psychological review* 122(2):312.
3. Turner BM, Rodriguez CA, Norcia TM, McClure SM, Steyvers M (2016) Why more is better: Simultaneous modeling of eeg, fmri, and behavioral data. *Neuroimage* 128:96–115.
4. Mittner M, et al. (2014) When the brain takes a break: a model-based analysis of mind wandering. *Journal of Neuroscience* 34(49):16286–16295.