

Simultaneous modeling of a Spatial Cuing Task

# Simultaneous Modeling of Reaction Times and Brain Dynamics in a Spatial Cuing Task

Simon R. Steinkamp<sup>1,\*</sup>, Gereon R. Fink<sup>1,2</sup>, Simone Vessel<sup>1,3,†</sup>, Ralph Weidner<sup>1,†</sup>

<sup>1</sup> Cognitive Neuroscience, Institute of Neuroscience & Medicine (INM-3), Research Centre Juelich, 52425, Juelich, Germany

<sup>2</sup> Department of Neurology, Faculty of Medicine and University Hospital Cologne, University of Cologne, 50937, Cologne, Germany

<sup>3</sup> Department of Psychology, Faculty of Human Sciences, University of Cologne, 50923, Cologne, Germany

† both authors contributed equally

\* **Corresponding author:**

Simon R. Steinkamp

Cognitive Neuroscience, Institute of Neuroscience and Medicine (INM-3)

Forschungszentrum Jülich,

Leo-Brand-Str. 5, 52425 Jülich, Germany

Email: [s.steinkamp@fz-juelich.de](mailto:s.steinkamp@fz-juelich.de)

## Abstract

Understanding how brain activity translates into behavior is a grand challenge in neuroscientific research. Simultaneous computational modeling of both measures offers to address this question. The extension of the dynamic causal modeling (DCM) framework for BOLD responses to behavior (bDCM) constitutes such a modeling approach. However, only very few studies have employed and evaluated bDCM, and its application has been restricted to binary behavioral responses, limiting more general statements about its validity.

## Simultaneous modeling of a Spatial Cuing Task

This study used bDCM to model reaction times in a spatial attention task, which involved two separate runs with either horizontal or vertical stimulus configurations. We recorded fMRI data and reaction times (n=29) and compared bDCM to classical DCM and a behavioral Rescorla-Wagner model using goodness of fit-statistics and machine learning methods.

Data showed that bDCM performed equally well as classical DCM when modeling BOLD responses and better than the Rescorla Wagner model when modeling reaction times. Notably, only using bDCM's parameters enabled classification of the horizontal and vertical runs suggesting that bDCM seems to be more sensitive than the other models. Although our data also revealed practical limitations of the current bDCM approach that warrant further investigation, we conclude that bDCM constitutes a promising method for investigating the link between brain activity and behavior.

*Keywords: spatial attention, fMRI, simultaneous modeling, effective connectivity, dynamic causal modeling, behavioral dynamic causal modeling*

## Acknowledgements

SV was supported by funding from the Federal Ministry of Education and Research (BMBF, 01GQ1401). We also like to thank our colleagues at the INM-3 for their valuable feedback and suggestions throughout the development of the study and analyses.

## Conflict of Interest

The authors declare no competing financial interests.

## Data Availability Statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

## Introduction

Computational modeling can deepen our understanding of how the brain processes information and produces overt behavior. In the field of psychology, computational modeling

## Simultaneous modeling of a Spatial Cuing Task

has a long history in describing and explaining behavioral concepts. For example, reinforcement learning algorithms have been used to explain classical conditioning (Rescorla et al., 1972), drift-diffusion models have been used to model reaction times in decision-making tasks (Ratcliff, 1978), and race models have been used as theoretical formulations of visual-spatial attention (Bundesen, 1990). Similarly, different computational modeling approaches have been employed in the fields of neuroscience and neuroimaging. For example, generative graphical models of brain connectivity describing blood oxygenation level-dependent (BOLD) amplitudes in response to experimental inputs can be estimated using dynamic causal modeling (Friston et al., 2003, 2017), and multivariate temporal response functions have been used to model ongoing sensory stimulation, like speech, in electrophysiological recordings (Crosse et al., 2016).

Although computational models are very prominent in the two fields, behavioral and neural responses are mostly treated separately (Turner et al., 2017). However, a combined modeling approach could provide us with deeper insights into the neural processes and the emergence of behavior. Here, different approaches have been proposed: One possibility is to correlate the parameters of neural and behavioral models to describe how the different measures are related across different participants (Vossel et al., 2016). Alternatively, in model-based fMRI, the behavioral computational model's outputs (or hidden states) are used as a factor in a classical GLM analysis. One such factor could be a participant's perceived cue validity in a probabilistic spatial cueing task, which can be recovered from reaction times (e.g., Dombert et al., 2016). Leveraging the theory-driven performance of cognitive models allowed to determine more specific brain activation patterns of cognitive processes than by using non-specific measures such as reaction times (Turner et al., 2017). A third option is a joint modeling approach (Turner et al. 2017). Here, an overarching set of parameters is used to describe both brain activity and behavior. An example is a study by Nunez et al. (2015), where the drift-diffusion model parameters were constrained with task-based brain activity, incorporating the covariation between reaction times and neural-activity on a trial-by-trial basis.

## Simultaneous modeling of a Spatial Cuing Task

Although these approaches are tremendously useful, none of them employs an integrative model describing the generation of brain activity and behavior, allowing us to investigate the hidden processes behind the two measurements directly. Rigoux and Daunizeau (2015) have provided such a framework, where DCM is extended by an additional output function to describe behavioral responses (behavioral DCM, bDCM). This simultaneous modeling does have not only high descriptive power but also allows thorough diagnostics of the model. For example, by disabling specific nodes in the network (i.e., artificial lesions), conclusions can be drawn about the contribution or necessity of different brain regions to the emergence of behavioral patterns. So far – to our knowledge – bDCM has been applied to a larger datasets in one study, which modeled binary choices in an economic decision making task (Shaw et al., 2019).

In the current study, we show that bDCM can be extended to continuous measures (i.e., reaction times). Furthermore, we provide a direct comparison between bDCM and classical DCM and between bDCM and an adjusted version of the Rescorla-Wagner model (Rescorla et al., 1972; Vossel, Mathys, et al., 2014).

As a testing ground, we modeled the effects of attentional reorientation along the horizontal and vertical meridians in a spatial cueing-paradigm, where participants had to report the orientation of a pre-cued Gabor patch. In trials in which invalid cues indicated an incorrect location of the target Gabor patch (20 % of the trials), participants had to reorient their attention to the opposite location (Posner, 1980). This paradigm has been found to elicit reliable reaction time differences between invalid and valid trials, both on the individual and the group level (Hedge et al., 2017). Additionally, it has been shown that the internal representation of cue-validity can be modeled using the Rescorla-Wagner model as a generative model of reaction times (Mengotti et al., 2017; Rescorla et al., 1972; Vossel, Mathys, et al., 2014).

Besides the reliable behavioral effects, the cortical networks involved in this task have been characterized by multiple studies. We have previously analyzed the present dataset using classical DCM (Steinkamp et al., 2020), which has also been used in similar cueing paradigms

## Simultaneous modeling of a Spatial Cuing Task

(c.f., Vossel et al., 2012). Moreover, studies in patients with stroke-induced lesions have revealed brain regions that are critically involved in spatial cueing-tasks (Corbetta & Shulman, 2011; Malherbe et al., 2018; Posner et al., 1984). It is well established that the orientation of visual-spatial attention is mediated by a dorsal fronto-parietal attention network consisting of the intraparietal sulci (IPS) and the frontal eye fields (FEF). This network interacts with a ventral fronto-parietal attention network of ventral frontal cortex, and the temporoparietal junction (TPJ) when a sudden reorientation of attention is necessary (Corbetta et al., 2005; Corbetta & Shulman, 2011).

In addition to comparing bDCM, which simultaneously models behavior and brain activity, to classical DCM and a purely behavioral model, we also investigated whether bDCM parameters encode additional information about the task that is not comprised in DCM or the Rescorla-Wagner model. For this, we followed the idea of generative embedding (Brodersen, Haiss, et al., 2011; Brodersen, Schofield, et al., 2011) and tested whether we could separate the horizontal and vertical runs of our experiments based on the parameter estimates of our models, which was not possible in previous analyses (Steinkamp et al., 2020).

## Methods

### Participants

Data were collected from 29 participants (15 female, 21-39 years old,  $M=25$ ,  $SD=3$ ) with normal or corrected-to-normal vision (all right-handed, Edinburgh handedness Inventory (Oldfield, 1971),  $M=0.86$ ,  $SD=0.14$ ), who provided written informed consent to participate in the study. One participant had to be excluded subsequently because of noncompliance. Another participant was excluded due to excessive head-movement (predefined criteria translation  $> 3\text{mm}$ , rotation  $> 3^\circ$ ). Furthermore, we could not extract the time-series for the left-TPJ VOI in one participant. Therefore, the final sample included 26 participants. The ethics board of the German Psychological Association had approved the study. Volunteers were paid 15€ per hour for their participation. The dataset has been used in a previous study (see Steinkamp et al., 2020).

## Simultaneous modeling of a Spatial Cuing Task

### Task

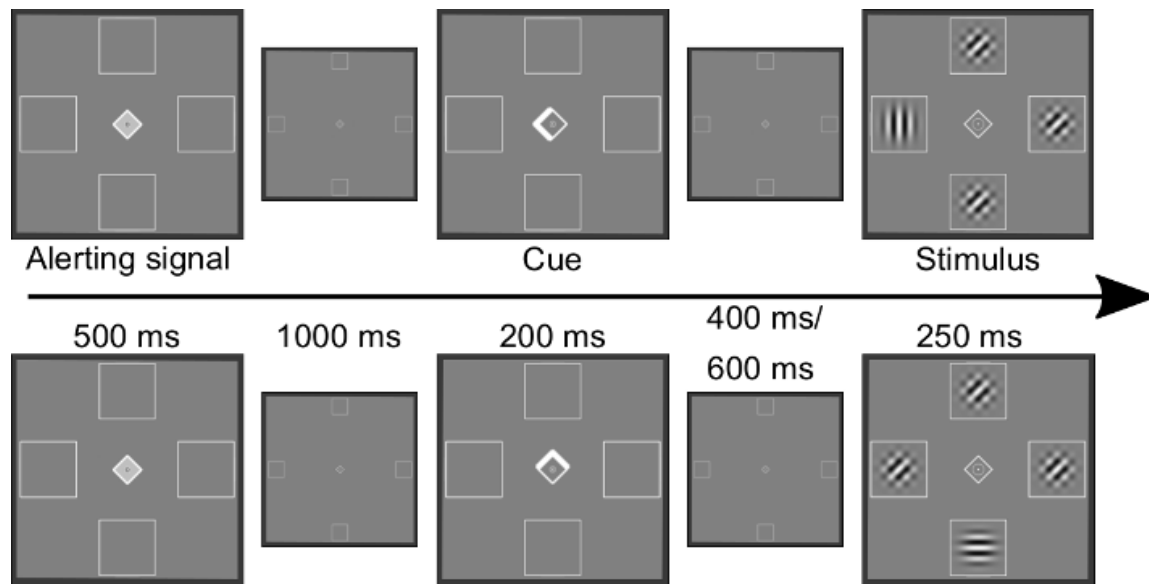


Figure 1 Illustration of the spatial cueing paradigm. In the upper row, a valid trial of the horizontal run is shown. The lower row depicts an example of an invalid trial in the vertical run. Reproduced from Steinkamp et al. (2020).

Participants performed a spatial cueing task while lying in a 3T Trio (Siemens, Erlangen) MRI scanner. Stimuli were displayed on a screen behind the scanner bore, which could be seen via a mirror (mirror to display distance: 245 cm) that was mounted on a 32-channel head coil. The participants' task was to report the orientation (horizontal/vertical) of a target Gabor patch (size  $1^\circ$  visual angle) by button presses of either the left or the right index finger while continuously fixating a diamond in the center of the screen ( $0.5^\circ$  visual angle). A brightening of the central diamond (500 ms) indicated the beginning of a trial and was followed by a spatial cue after 1000 ms (brightening of one of the diamond's edges for 200 ms) that indicated the location of the next target stimulus with 80 % probability. Participants were explicitly informed about the percentage of cue validity. The possible target locations were indicated by empty boxes ( $1^\circ$  width) located to the left, right, top, and bottom of the fixation diamond ( $4^\circ$  visual angle). After 400 ms or 600 ms, the target stimulus appeared for 250 ms at the cued location or in the box opposite to it. Distractor stimuli (constructed from two overlapping Gabor patches that were rotated by  $-45^\circ$  and  $45^\circ$ , respectively) appeared simultaneously in the remaining three locations. Participants performed two runs of the spatial cueing-paradigm. In one run,

## Simultaneous modeling of a Spatial Cuing Task

targets and cues occurred along the vertical axis, in another along the horizontal axis (see Figure 1).

Each run consisted of five blocks of 40 trials (32 valid, 8 invalid). All possible combinations of target location, target orientation, and inter stimulus interval were presented in random order within each block. The time between the trials was drawn from the set of 2.0 s, 2.7 s, 3.2 s, 3.9s, or 4.5 s with equal probability. Between the blocks, there was a break of 10 to 13 s.

Run order (vertical or horizontal first) and the response mapping (left index finger for vertical orientations/right index finger for horizontal orientations or vice versa) were counterbalanced across participants. Before the experiment, participants performed a rapid detection task to train the mapping of stimulus-response associations. Here, targets appeared rapidly in the middle of the screen, and participants had to press the corresponding button as fast as possible. Immediate feedback and a running score of their accuracy were given. Additionally, there were 20 practice trials with feedback before each run of the main experiment.

Stimulus presentation and response collection were controlled using PsychoPy (version 1.85.3, Peirce, 2007, 2008; Peirce et al., 2019).

## Behavioral analysis

The mean reaction time was calculated for each participant and cueing condition and for each target location. Before calculating the mean reaction times, we preprocessed the data for each participant separately. First incorrect, missed, and outlier trials were removed. Outliers were defined as trials with reaction times below 0.2 s and reaction times greater than the 75<sup>th</sup> percentile + 3 \* Inter Quartile Range (IQR). The higher threshold for outlier exclusion was chosen to retain as many trials as possible in the analysis (removed trials, including errors, in the horizontal run: invalid M = 2.54, SD = 2.63; valid M = 6.62, SD = 5.91; in the vertical run: invalid M = 3.12, SD = 1.8; valid M = 6.0 SD = 3.94).

For the analysis of the “validity effect” (i.e., the slowing of reaction times in invalid as compared to valid trials), the data were pooled across the two runs (horizontal/vertical). The mean

## Simultaneous modeling of a Spatial Cuing Task

reaction times of the 2 x 4 (cueing x target location) factorial design were then analyzed in a repeated-measures ANOVA. The analysis was conducted in Python 3.7 using pingouin (version 0.3.3, Vallat, 2018).

## fMRI analyses

For each participant and each run, we collected 557 T2\*-weighted images using an echo planar imaging (EPI) sequence (time of repetition (TR) 2.2 s; echo time (TE) 30 ms; flip angle 90°). Each recorded volume consisted of 36 transverse slices with a slice thickness of 3mm and a field of view of 200mm. The voxel size was 3.1 x 3.1 x 3.3 mm. The first 5 images were discarded to account for T1 equilibrium artifacts. Next to functional images, we also obtained an anatomical T1-weighted image for each participant, which was used in the preprocessing.

We preprocessed the fMRI data using fmriprep (version 1.1.1, Esteban et al., 2019), a robust and standardized pipeline, which applies slice-time correction, realignment, and normalization to MNI space. A detailed preprocessing report can be created automatically (see <http://fmriprep.readthedocs.io/en/1.1.1/workflows.html>), and has been included in the supplement.

Data was further spatially smoothed using an 8 x 8 x 8 mm FWHM Gaussian kernel. This step was done in Matlab 2018b (The MathWorks, Inc., Natick, Massachusetts, United States), using SPM12 (version 7771, Friston, 2007).

## fMRI - GLM

A classical GLM analysis was performed to identify activation peaks during attentional orientation and reorientation, which were later used to extract BOLD time-series data for the DCM analysis. The GLM analysis was conducted using SPM12. First-level models were created with four regressors of interest for each run, representing invalidly cued targets on the left (iL) and on the right (iR), as well as validly cued targets on the left (vL) and the right (vR) for the horizontal run, and invalidly and validly cued targets in the lower (iD, vD) and the upper (iU, vU) part of the screen in the vertical run.



## Simultaneous modeling of a Spatial Cuing Task

To account for other physiological noise in the BOLD signal, we added the three rotation and three translation estimates of the rigid body transform, the average white matter signal, and the average cerebral spinal fluid (CSF) signal as nuisance regressors. We further included the squared time-series of the 8 regressors, the time-shifted time-series ( $t-1$ ), as well as the square of the shifted time-series, resulting in a total of 32 nuisance regressors (Friston et al., 1996). We also applied a high pass filter at 128 s. For each run, four first-level contrasts were calculated: T-contrasts of valid and invalid trials versus baseline, an F-contrast of target onset versus baseline, which were used in the VOI analysis, and a differential contrast of invalid trials greater than valid trials. The latter contrast isolates brain regions involved in the attentional reorientation of attention.

At the group (second)- level, we investigated the differential contrast of invalid > valid trials using two planned one sample permutation t-tests against 0 using SnPM 13 (Nichols & Holmes, 2002), with default settings, 10000 permutations, and no additional variance smoothing. The cluster forming threshold was estimated during the processes with a predefined voxel-level cutoff of  $p < 0.001$ .

## Modeling Analysis

In the following, we will describe the modeling approaches used in our analysis, followed by a description of our model assessments and further analyses.

### Rescorla-Wagner Model

We employed a variant of the Rescorla-Wagner model that we already used previously (Mengotti et al., 2017). While these studies were interested in the  $\alpha$  parameter (the learning rate that describes how quickly participants adjust their internal assessment of the cue-validity), we applied this modeling approach to simulate reaction times in a trial-by-trial fashion. For parameter estimation, we defined new functions for the VBA (Variational Bayesian Analysis) toolbox (clone from master, in Jan. 2020, Daunizeau et al., 2014).

We used the following reinforcement learning formula as the evolution function, describing the hidden process governing the generation of reaction times.

## Simultaneous modeling of a Spatial Cuing Task

$$v_t = v_{(t-1)} + \alpha * \delta$$

Where  $\delta = u_t - v_{(t-1)}$  describes the prediction error at trial  $t$ . The external input  $u_t \in [0, 1]$  describes whether the cue at time  $t$  was either valid (1) or invalid (0),  $\alpha$  is the learning rate and  $v_t$  is the participant's perceived cue validity after observation of trial  $t$ .

The observation function (i.e., the mapping from perceived cue validity to reaction times) was defined as:

$$g_t = u_t * (\zeta_v + \zeta_2 * v_{t-1}) + (1 - u_t) * [\zeta_i + \zeta_2 * (1 - v_{t-1})]$$

According to this formulation, the perceived cue-validity of the previous trial governs the responses, with different bias parameters for valid and invalid trials and a general scaling parameter of the predictions.

Table 1 depicts the Gaussian priors used in our estimation:

Table 1: Overview of parameters and prior values of the Rescorla-Wagner Model.

Parameters	$\mu$	$\sigma$	
$\alpha$	0.5	0.5	To ensure $0 < \alpha \leq 1$ , $\alpha$ was logit and inverse logit transformed during parameter updating
$\zeta_v$	0	1	
$\zeta_i$	0	1	
$\zeta_2$	0	1	
$v_0$	0.5	1	Initial state of $v$

## Behavioral DCM

In the following, we will provide a short overview of key-concepts of dynamic causal modeling (DCM). For a full derivation and detailed description of DCM, see (Friston et al., 2003; Rigoux & Daunizeau, 2015; Stephan et al., 2008). DCM is a fully-Bayesian approach to create a generative model of brain dynamics and infer effective connectivity between selected brain regions. In principle, DCM describes how experimental variations (described by the input  $u$ ) drive the neural activity ( $x$ , the hidden states) in brain regions of interest in a dynamical system. The evolution function ( $\dot{x} = f(x, u)$ ) describes the temporal dynamics of the hidden states ( $\dot{x}$ )

## Simultaneous modeling of a Spatial Cuing Task

and how they are influenced by external inputs ( $u$ ). In DCM for fMRI, the evolution function  $f$  is typically described as:

$$f(x, u, \theta) \cong Ax + \sum_j u_j B^j x + Cu$$

Where  $j$  corresponds to the number of inputs and  $i$  to the number of brain regions. The neural evolution parameters in  $\theta$  correspond to the entries in  $A$  (fixed connectivity between brain regions),  $B^j$  (modulation of connection strength by input  $j$ ),  $C$  (direct effects of inputs). Hemodynamic states  $z$  (dependent on the neural states  $x$ ) are then gated through an observation function:

$$y = g(z, \phi) + \epsilon$$

This function captures BOLD signal variations based on the hemodynamic states ( $z$ ) and the hidden neural activity ( $x$ ), with hemodynamic parameters  $\phi$ . This mapping allows to observe and infer the hidden neural dynamics via the BOLD signal.

BDCM augments the described formulation of DCM by an additional evolution ( $h(x, u, \psi)$ ) and observation functions ( $g_r(r) + \epsilon_r$ ) to map the hidden neural dynamics to behavioral responses. The evolution function  $h$  of the new “behavioral” state follows the same rationale as the function  $f$  in the DCM formulation:

$$h(x, u, \psi) \cong A_r x + \sum_j u_j B_r^j x + C_r u$$

Here the parameter vector  $\psi$  describes the linear ( $A_r$ ) components of the behavioral state, as well as the direct ( $C_r$ ) and modulatory ( $B_r^j$ ) influences of experimental manipulations.  $A_r$  is an analogy of the weight vector in a regression model. In the original paper, the neural states were mapped to binary behavioral observations (button press absent or present) via a sigmoidal function:

$$s(r) = \frac{1}{1 + e^{-100*(\rho+r)}}$$

## Simultaneous modeling of a Spatial Cuing Task

Here,  $\rho$  is an unknown bias term and  $r$  is the response or decision state. In our study, we slightly adjusted the sigmoid mapping, by changing the scale on which it operates. As we are not expecting reaction times slower than 3 s, we used this as an upper bound:

$$s(r) = \frac{3}{1 + e^{-100*(\rho+r)}}$$

### Regions

As in our previous study (Steinkamp et al., 2020), we included bilateral IPS and FEF in our DCM model, which correspond to the central nodes of the dorsal fronto-parietal attention network (Vossel, Geng, et al., 2014). Additionally, as part of the ventral attention network, we included the TPJ bilaterally. As additional inclusions (e.g., the inferior/middle frontal gyrus) would have led to increasing model complexity and computational resources (and time), we did not include other brain regions, which may also play a role in attentional reorienting.

Based on our assumptions about the dorsal and ventral attention network's interplay, we created three automatic meta-analysis using Neurosynth (<https://www.neurosynth.org/>, Yarkoni et al., 2011), to define the seed coordinates for the subsequent VOI analysis (see Table 2). Our regions of interest were bilateral IPS (search term: "intraparietal sulcus"), bilateral FEF (search term: "frontal eye"), and bilateral TPJ (search term: "tpj"). We downloaded the corresponding association maps (associations,  $p < 0.01$  FDR corrected) and identified the seed location as the peak voxel in the cluster of interest, using the Anatomy toolbox (v2, Eickhoff et al., 2005). In all three maps, the two largest clusters encompassed our regions of interest in either the left or right hemispheres.

Table 2: Regions and search-terms for automated Neurosynth meta-analyses.

Region	NeuroSynth (accessed 10.10.19)	Z- Statistic	X	Y	Z
IPS – left	"intraparietal sulcus"	14.6	-30	-50	42
IPS – right	"intraparietal sulcus"	13.5	40	-38	44
FEF – left	"frontal eye"	13.9	-30	-4	52
FEF – right	"frontal eye"	14.6	32	-6	52
TPJ – left	"tpj"	8.56	-60	-54	20
TPJ – right	"tpj"	11.4	58	-50	14

## Simultaneous modeling of a Spatial Cuing Task

In each run, we used the participant level t-maps (thresholded at  $p < 0.1$  uncorrected) to search for individual local maxima in a 12 mm sphere around the seed coordinates. The first principle component of BOLD time-courses in a 9 mm VOI around the participant's maximum was extracted and adjusted based on the F-contrast defined in the first-level analysis. Task-related activity for the IPS and FEF VOIs was defined by the contrast of valid trials against baseline and for TPJ by the contrast of invalid trials against the baseline.

### *Preprocessing*

We preprocessed the BOLD signal by detrending the signal in each VOI (*spm\_detrend*) and scaling the BOLD amplitude across VOIs to a maximum value of 4 (see *spm\_dcm\_estimate*). Behavioral data were extracted from the event data, and as in the previous analyses, error trials and trials with missed responses, as well as RTs fulfilling the outlier criterion ( $RT < 0.2s$  and  $RT > 3 * IQR + UQ$ ), were excluded.

BOLD data were resampled from a TR of 2.2 s to a sampling rate of 1.1 s (by interspersing “NaN” values). The behavioral observations were set to occur at the corresponding target onset, which was also downsampled to a resolution of 1.1 s. No resampling of BOLD data was performed for the classical DCM analysis. As the Rescorla-Wagner model represents trial-by-trial dynamics, the corresponding preprocessed reaction times were used, excluding error and missed trials.

For our modeling, we assumed homogenous HRF dynamics across the six regions, fixing the initial states of the model to 0 and estimating the shape of the observation noise hyper-prior distributions. For this, we assumed that we would be able to explain 10 – 90% of the variance in both the BOLD and the reaction time data. The prior distributions over the other parameters were set to the defaults of the VBA toolbox. We used the same hyperpriors for the explained variance of the BOLD signal in the classical DCM analysis and the Rescorla-Wagner model's behavioral responses.

## Simultaneous modeling of a Spatial Cuing Task

To define the inputs into the DCM models, we created separate SPM-design matrices that were only used to define the input streams. Stream one was defined as the driving input to all six regions, containing an impulse every time a target stimulus appeared (irrespective of the cueing condition or target location). The second stream was used purely for the modulatory effects, containing an impulse only in invalidly cued targets. The input streams were extracted from the SPM design matrix and were centered before entering the model inversion (*spm\_detrend*). As mentioned above, the Rescorla-Wagner model is modeling trial-by-trial variations (rather than continuous time), so the input to this model was a vector consisting of ones and zeros, indicating whether the current trial is invalid or valid.

### *Model definition*

We used the same general model structure for behavioral and classical DCM analysis. As in our previous publication, we used IPS, FEF, and TPJ as our brain regions of interest. For our analysis, we inverted a single model. The fixed connectivity structures of our model (i.e., the A-matrix) had full connections in each hemisphere and connections between homologous regions (Figure 2). As we did not include visual areas in our modeling approach, all six regions received driving input (C-matrix). For bidirectional intra- and interhemispheric modulatory connections (B-matrix), we considered the IPS and TPJ. Connections in both hemispheres to the FEF were unidirectional, assuming that there were no feedback modulations from FEF to the other brain regions. In the case of bDCM, we also considered all six nodes as output regions.

## Simultaneous modeling of a Spatial Cuing Task

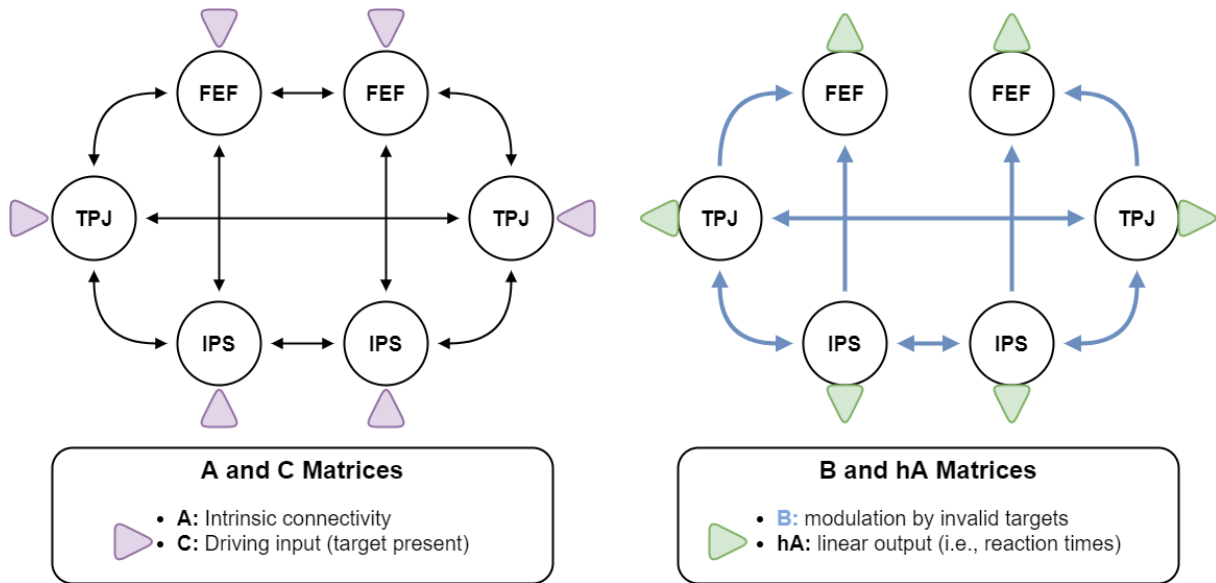


Figure 2: Basic structure of the DCM model. Regions were fully inter-connected in each hemisphere, and homologous regions were connected. All regions received driving input. We assumed that all connections between the regions were modulated by invalid trials, except for feedback and interhemispheric connections from FEF.

## Model evaluation of the Rescorla-Wagner model, classical DCM, and bDCM

As all three models are based on different underlying data at different times scales (i.e., reaction times only, BOLD time-series only, or both), we only compared the models based on their outputs, applying classical goodness of fit-statistics. The  $R^2$ -score,

$$SS_{tot} = \sum_t^n (y_t - \bar{y})^2$$

$$SS_{res} = \sum_t^n (y_t - \hat{y}_t)^2$$

$$R^2 = 1 - \left( \frac{SS_{res}}{SS_{tot}} \right)$$

where  $y_t$  describes the datapoint at  $t$ , with  $n$  timepoints in total. The average of  $y$  is defined as  $\bar{y}$ , and  $\hat{y}$  are predicted values. Similarly, we also calculated the mean absolute error (MAE)

$$MAE = \frac{1}{n} \sum_t^n |y_t - \hat{y}_t|$$

## Simultaneous modeling of a Spatial Cuing Task

Here, we estimated for each subject whether the fit-statistics were different from random for each of the model outputs by permutation testing. The predicted values  $\hat{y}$  were shuffled 10000 times (without replacement), and the two statistics were recalculated. The permutation p-value for the models is then reported as the proportion of fits greater than the model's  $R^2$ -score (smaller in case of MAE) plus one divided by the number of permutations plus one (Ojala & Garriga, 2010). At the group level, we report the proportion of significant models, based on a permutation p-value  $< 0.05$ .

To compare the model performance on reaction times, we used (Bayesian-) paired t-tests to test for differences between the fit-statistics ( $R^2$ -score and mean absolute error (MAE)), separately for the two runs. We then investigated how well bDCM and the Rescorla-Wagner model simulate the underlying reaction time distributions. This was achieved by calculating a two-sample Kolmogorov-Smirnov test between the model-derived reaction times of the Rescorla-Wagner model or bDCM and the measured reaction times. Finally, paired t-tests were used on the distance between the distributions (as determined by the KS-test) to test which simulation followed the measured data more closely (i.e., had a smaller distance at the group level).

We applied a mixed-effects linear model for each error term to compare differences in performance to the BOLD data fit between classical DCM and bDCM. The mixed-effects model followed the following formula, where "Score" either depicts the mean absolute error or the  $R^2$ -score:

$$\text{Score} \sim \text{Model} + \text{Region} + \text{Run} + \text{Model} * \text{Region} + \text{Model} * \text{Run}$$

and Model has the two factors "DCM" and "bDCM", "Run" describes either the horizontal or vertical run, and "Region" indicates the "Score" for either VOI. Each model also contained a random intercept for each participant.



## Simultaneous modeling of a Spatial Cuing Task

### Generative Embedding

We also investigated whether bDCM parameters encode additional information, which is not contained in classical DCM or in the Rescorla-Wagner model. Therefore, we tested whether the parameters derived from our computational models could be used to separate the horizontal and vertical run. Note that our previous work using classical DCM did not provide any evidence for distinct behavioral or neural processes (Steinkamp et al., 2020).

To test if a separation is possible and which sets of parameters encode the necessary information, we ran several experiments using different feature sets. Our feature sets for DCM and bDCM were: “AB”, the sum of the “A” and “B” matrices (19 parameters); “AB + C”, extending the feature set with the “C” matrix parameters (25 parameters); in case of bDCM we additionally tested the feature sets “AB + hA”, where the “hA” matrix was also included in the set (25 parameters); “AB + C + hA”, including all connectivity parameters of the bDCM (31 parameters); “C + hA”, using only the input and outputs connections of bDCM. We then also tested whether we can find any information in the parameters of the Rescorla-Wagner model (see Table 2, “RW”, 4 parameters), and whether parameters of DCM and the Rescorla-Wagner model in combination have a beneficial effect (“AB + C + RW”, 29 parameters). To be sure that there was no additional information in the BOLD time-series, we also included a feature set based on the correlation of the time-series of the six VOIs in our experiments (“Correlation”, 15 parameters) and of the time-lagged correlation with the time-series (“Correlation + Correlation<sub>t-1</sub>”, 51 parameters). Correlations were established using Nilearn’s *ConnectivityMeasure* (Abraham et al., 2014), with no variance scaling.

As we have a minimal sample size for classification approaches (26 participants, 52 instances), we applied an elaborate cross-validation procedure for our results. We applied 5-fold cross-validation to obtain an estimate of the accuracy of each feature set, whose significance was further evaluated by permutation testing (1000 iterations, scikit-learn’s *permutation\_test\_score*; Pedregosa et al. (2011)). To keep more independence between the train and test-sets, we ensured that a given participant’s instances were not distributed across

## Simultaneous modeling of a Spatial Cuing Task

training and test-sets. Due to the small sample size and high variance between participants, we repeated the cross-validation procedure 20 times, shuffling the participants order, and creating different train-test splits for each iteration.

To account for a certain amount of algorithmic variance, we performed the procedure above using two classifiers, a logistic regression, and a linear support vector machine (scikit-learn, using default parameters). Input data were normalized using robust-scaling (scikit-learn's *RobustScaler*).

## Lesion Analysis

We also applied lesion analysis to the bDCM model, as described in Rigoux and Daunizeau (2015). Here, the afferent connections towards a single brain region were reduced to 0 to simulate the absence of this region (i.e., to create an artificial lesion). The simulated data from such a lesioned model can be used to better understand behavioral changes after damage to certain brain regions.

To alleviate the problem of numerical instabilities, which resulted in values resulting in infinity or minus infinity for the hidden states, we changed the posterior self-inhibitory connection in the DCM model to  $\log(1)$  (to ensure inhibition, i.e., negativity, self-connections in DCM are exponentiated before subtraction from the diagonal of the A matrix). While increasing the self-inhibition solved the problem of instabilities, it significantly impacted the fit-statistics of the non-lesioned model. Therefore, we only provide a qualitative description of the lesion analysis. In the end, we simulated data for all 26 participants for each lesion and several levels of lesion extent. This means, rather than switching off the afferent connections (i.e., the inputs to the region) altogether, we also simulated data for connections that had 95 %, 75 %, 50 %, 25 %, 5%, and 0 % of the original incoming strength.

Since a few models still were numerically unstable, we then cleaned the simulated data by removing datasets on a per lesion basis where the variance after the 20<sup>th</sup> trial was close to 0 (i.e., the simulated reaction times flatlined at the maximum/minimum of the sigmoid function) and which returned non-values. In the qualitative analysis, we compared the group-mean

## Simultaneous modeling of a Spatial Cuing Task

validity effects for different lesions, extents, and target sides, removing participants who had an absolute validity difference greater than 2 secs.

## Results

### Behavior

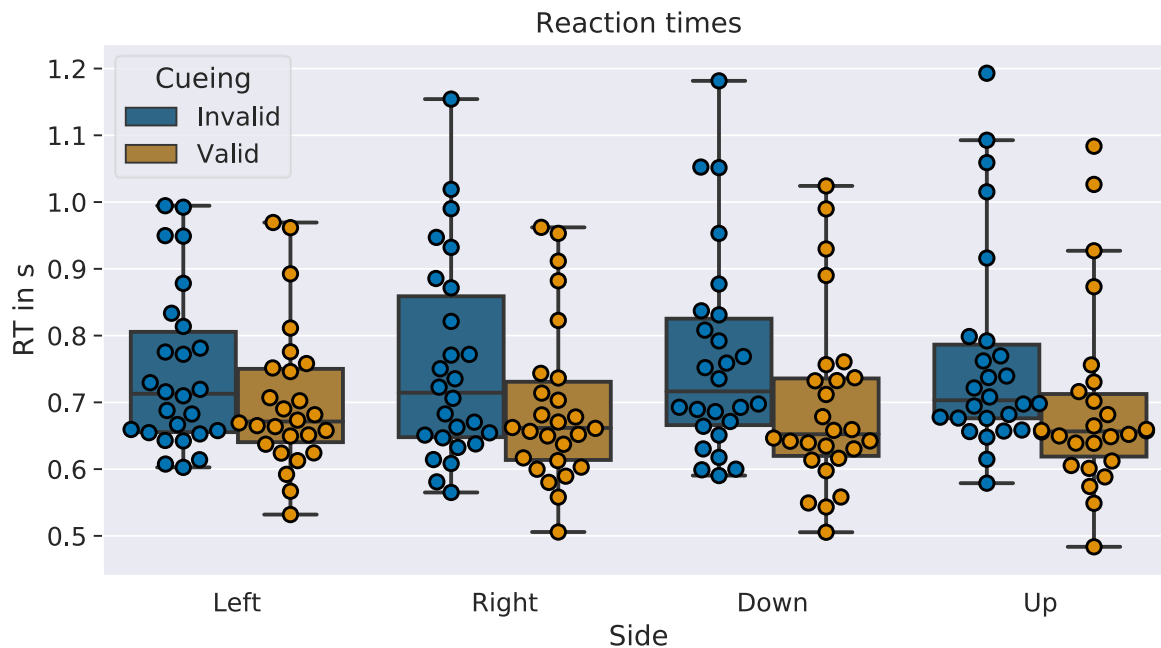


Figure 3: Box- and swarm plots of mean-reaction time data for each participant in the 8 conditions. The boxes indicate the inter-quartile range (IQR), the line in the middle the median reaction time, whiskers are extended to include the lower and upper quartiles plus 3 times the IQR. Loose points indicate outliers. The ANOVA's results are readily visible, as there are longer reaction times in invalid trials but no apparent effects between the different target-positions.

To test for reaction time effects of cueing (valid or invalid) and target-side (left, right, down, up), as well as their interactions, we applied a 2 x 4 repeated measures ANOVA (see also Figure 3). There was no significant effect of target-side ( $F(1.965, 49.125) = 0.1, p = 0.902, \eta_p^2 = 0.004, \epsilon = 0.655$ ). However, a significant main effect of cueing ( $F(1, 25) = 26.647, p < 0.001, \eta_p^2 = 0.516, \epsilon = 1.0$ ) and a weak significant interaction between target-side and cueing ( $F(2.88, 72) = 2.866, p = 0.045, \eta_p^2 = 0.103, \epsilon = 0.96$ ) were observed. All reported p-values were Greenhouse-Geisser corrected to account for a lack of sphericity.

## Simultaneous modeling of a Spatial Cuing Task

### FMRI GLM

The contrasts of invalid versus valid trials isolating reorienting-related activity for the two runs are reported in Figure 4 (group t-maps are provided on neurovault:

NV\_LINK, (Gorgolewski et al., 2015), the corresponding tables reporting global and local maxima for the different clusters are in supplement S2). We performed a one-sample permutation t-test on the first-level contrast images (invalid > valid), with a predefined cluster-forming threshold of  $p < 0.001$ , the results are reported family-wise error corrected at  $p < 0.05$

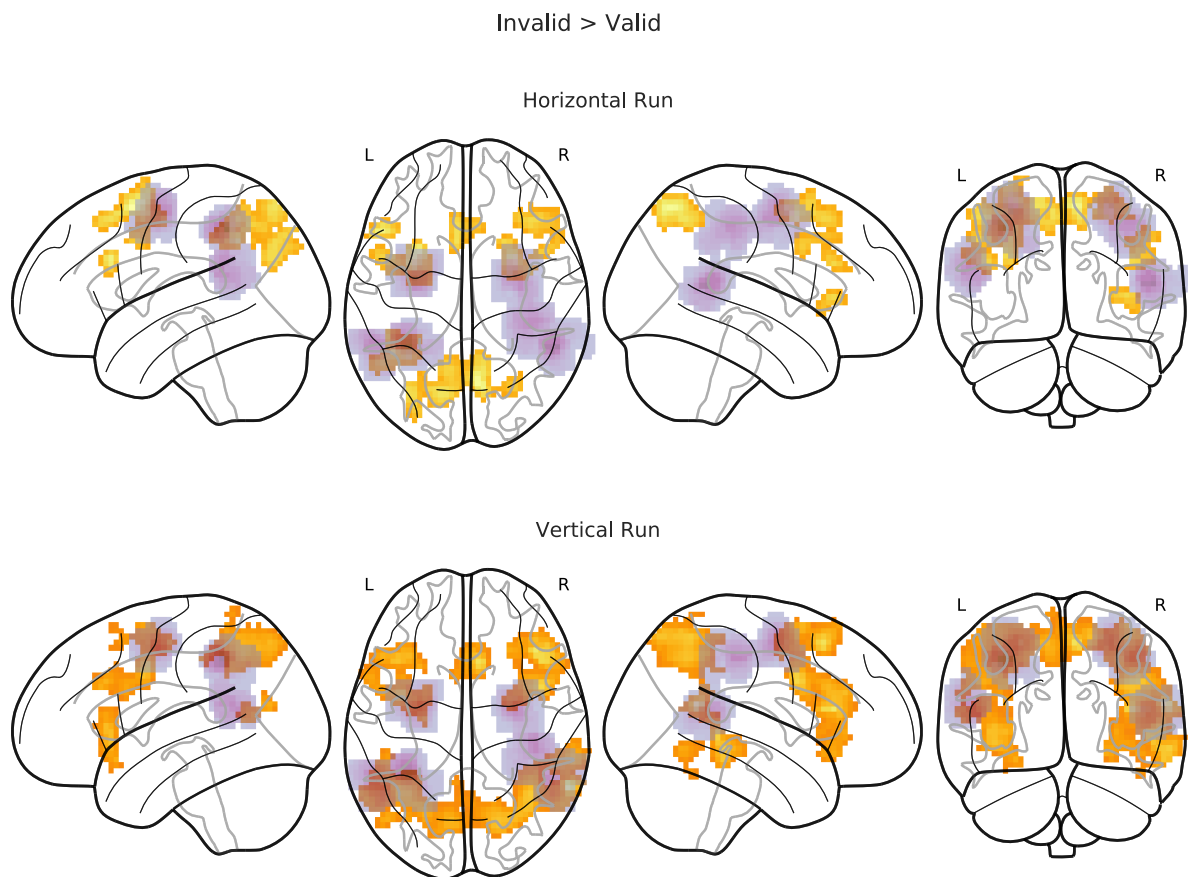


Figure 4: Non-parametric T-maps contrasting invalid > valid trials for the two runs ( $p < 0.05$  FWEc). The purple overlay indicates the regions where the 9mm VOIs for the bDCM analysis were extracted (sum of the participant's masks).

(cluster threshold horizontal  $k = 58$  voxels, vertical  $k = 70$  voxels). In both maps, we found areas classically associated with the dorsal and ventral attention networks. For example, in both runs, we observed significant activation in bilateral intraparietal sulci and frontal-eye fields. Activations of the ventral attention networks were less robust. For the horizontal run, for example, the invalid versus valid contrast revealed an involvement of the middle frontal gyrus predominantly in the right hemisphere and no significant activation close to the seed regions

## Simultaneous modeling of a Spatial Cuing Task

for the temporoparietal junction at the given threshold. However, the temporoparietal junction was significantly activated in the vertical run.

## Model Fit

### Reaction Time Data

We calculated the mean absolute error (MAE) and the  $R^2$  statistic for the Rescorla-Wagner and bDCM models and assessed their significance on a per-subject level by calculating permutation tests. In the horizontal run, the bDCM (mean absolute error,  $M = 0.091$ ,  $SD = 0.033$ , percent sig = 96.2 %;  $R^2$  score,  $M = 0.117$ ,  $SD = 0.084$ , percent. sig = 100 %) performed well when compared to the Rescorla-Wagner model (mean absolute error,  $M = 0.094$ ,  $SD = 0.033$ , percent sig. = 53.8 %;  $R^2$  score,  $M = 0.066$ ,  $SD = 0.08$ , percent. sig = 69.2 %). The

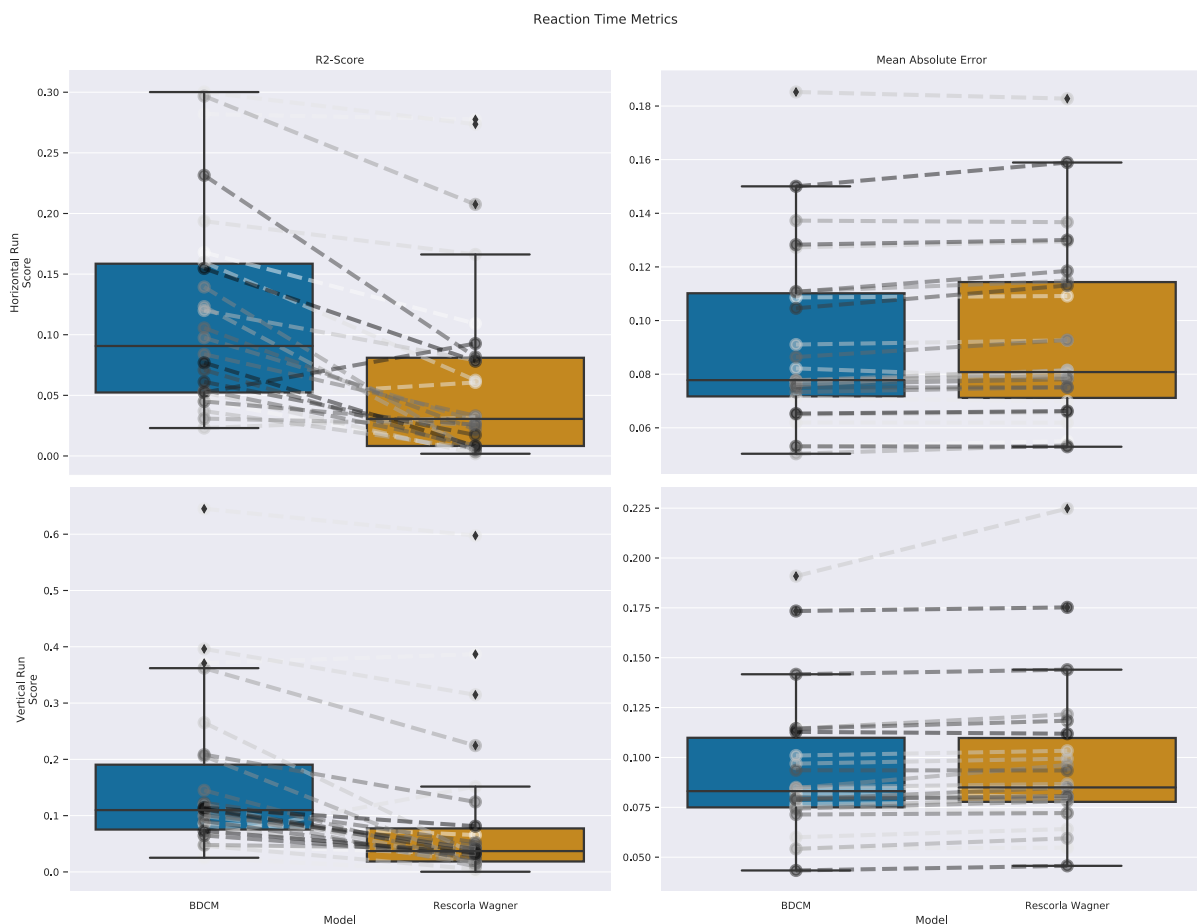


Figure 5: Boxplots comparing the different fit-statistics across models in the horizontal run. Please note that for the  $R^2$ -score a higher value is better, while the opposite is true for the mean absolute error. The dashed lines between the boxplots indicate individual participants.

results of the vertical run yielded a very similar picture, where the differences between the bDCM (mean absolute error,  $M = 0.093$ ,  $SD = 0.034$ , percent sig. = 100 %;  $R^2$  score,  $M =$

### Simultaneous modeling of a Spatial Cuing Task

0.158, SD = 0.143, percent. sig = 100 %) and the Rescorla-Wagner model (mean absolute error, M = 0.096, SD = 0.038, percent sig. = 80.8 %; R<sup>2</sup> score, M = 0.094, SD = 0.141, percent. sig = 88.5 %) were even more pronounced (see Figure 5). The paired t-tests (Table 3) confirmed this pattern and revealed a better fit for the vertical run. BDCM had a lower error and greater fit than the Rescorla-Wagner model.

Table 3: Paired t-tests between the fit-statistics of bDCM and Rescorla-Wagner models. The differences between the fit-statistics favor the bDCM model.

Run	Error	T (df 25)	p-val	CI95%	cohen-d	BF10
Horizontal	MAE	-3.557	0.002	-0.0, -0.0	0.066	23.692
Horizontal	R <sup>2</sup>	5.803	< 0.001	0.03, 0.07	0.625	4.249.729
Vertical	MAE	-2.644	0.014	-0.01, -0.0	0.098	3.571
Vertical	R <sup>2</sup>	5.204	< 0.001	0.04, 0.09	0.456	1.048.508

Furthermore, we evaluated how well the reaction time distributions of the two models' simulations matched the real reaction time distribution (Figure 6). We calculated the distance between the distributions of measured and simulated reaction times for each run, participant, and cueing-condition using the Kolmogorov-Smirnov test. We then performed paired t-tests in order to ascertain which simulation better matches the original distribution. In all cases, the bDCM simulation provided a better match (horizontal run, valid cueing,  $t(25) = -10.155$ ,  $p < 0.001$ , Cohen's  $d = 2.432$ ,  $BF_{10} = 4.39 * 10^{10}$ ; horizontal run, invalid cueing,  $t(25) = -8.121$ ,  $p < 0.001$ , Cohen's  $d = 2.392$ ,  $BF_{10} = 7.51 * 10^8$ ; vertical run, valid cueing,  $t(25) = -7.619$ ,  $p < 0.001$ , Cohen's  $d = 2.049$ ,  $BF_{10} = 2.56 * 10^8$ ; vertical run, invalid cueing,  $t(25) = -9.931$ ,  $p < 0.001$ , Cohen's  $d = 2.475$ ,  $BF_{10} = 2.87 * 10^{10}$ ). Based on visual inspection of the differences in reaction time distributions, deviations were especially pronounced at the extreme ends of the distribution.

## Simultaneous modeling of a Spatial Cuing Task

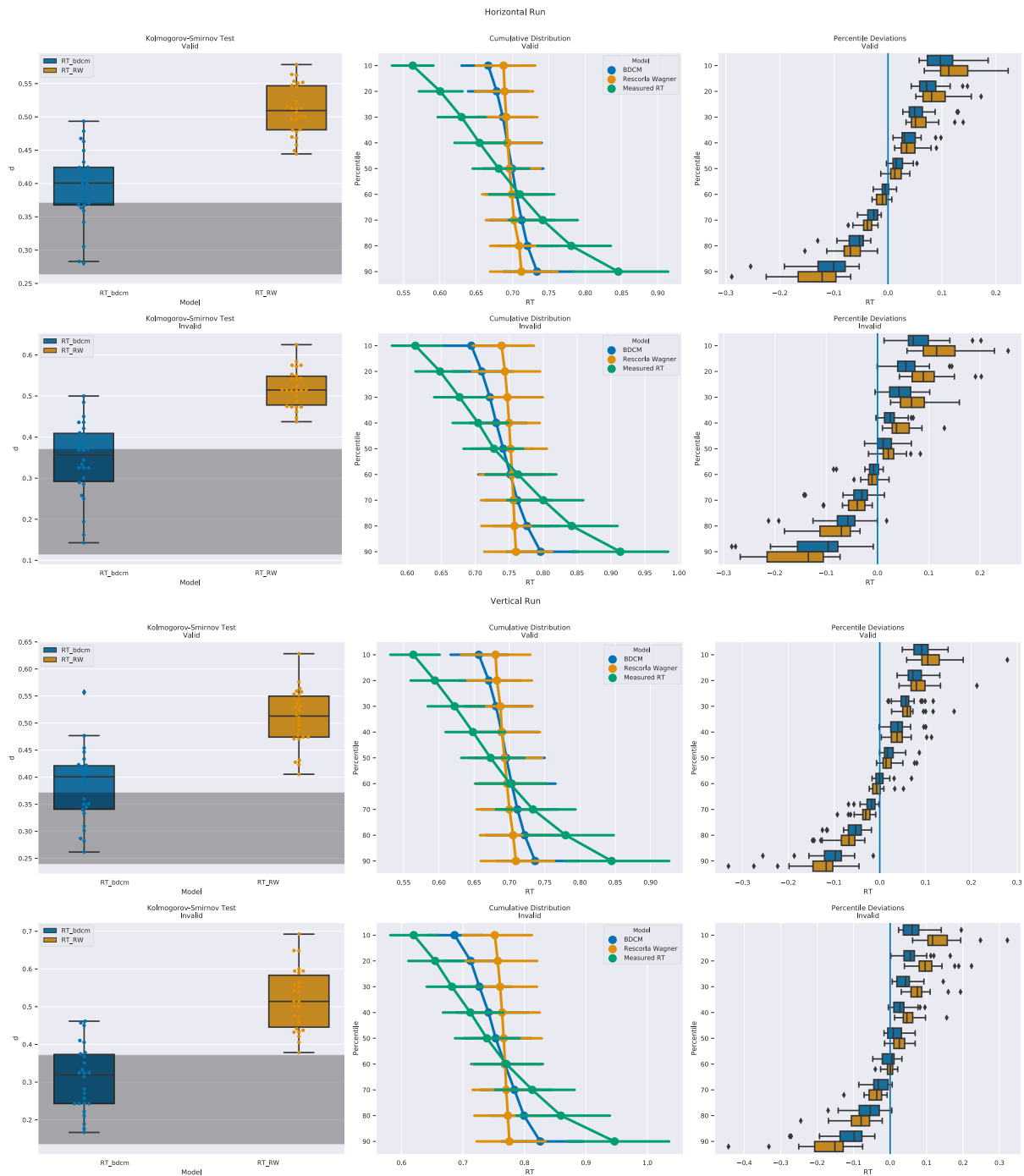


Figure 6: Left column, Kolmogorov-Smirnov distance between the simulated responses between the measured reaction times and bDCM or the Rescorla-Wagner model. The grey shading indicates non-significant differences. BDCM has, in general, a lower distance to the original distribution, and in many cases, the tests were non-significant. Both models also appeared to be better in the prediction of the invalid reaction time distribution. Middle Column: Cumulative reaction time distribution represented by the deciles of each model. The measured reaction times (green) have a more extensive spread than the reaction times from bDCM (blue) and the Rescorla-Wagner model (orange). Right column: The paired difference between the deciles of the Rescorla-Wagner model and bDCM. Differences are especially large in more extreme deciles.

## Simultaneous modeling of a Spatial Cuing Task

### BOLD Data

In a further step, we investigated whether DCM and bDCM were comparable in their fit to the measured BOLD data. For this, we calculated the fit-statistics (mean absolute error and  $R^2$ -score) for the two modeling approaches and the two runs. Since this yields fit-statistics for each brain region, we calculated two linear mixed-effects models (MLM) with participant as a random factor to test for a main-effect or interaction effects of model and fit-statistic, as well as the main effect of run and interactions between model and run (Figure 7). The results of the mixed-effects models are summarized in Table 4. Importantly, we did not find a significant main effect of model. However, there were significant main effects of brain region, these effects did not interact with the choice of model, indicating that the two models performed similarly. A similar conclusion can be drawn when looking at the different runs. Again, there were no significant main effects nor interactions of the factor model.

Table 4: Results of the MLM analysis for BOLD fit-statistics. The table is split for mean absolute error and  $R^2$  score.

	$R^2$ Score				Mean Absolute Error			
	Coef.	Std.Err.	z	P> z	Coef.	Std.Err.	z	P> z
Intercept	0.176	0.012	14.986	< 0.001	0.317	0.014	22.720	< 0.001
DCM	0	0.013	-0.017	0.986	0	0.016	-0.021	0.983
FEF Right	-0.048	0.012	-3.879	< 0.001	-0.033	0.015	-2.189	0.029
IPS Left	-0.053	0.012	-4.268	< 0.001	-0.054	0.015	-3.574	< 0.001
IPS Right	-0.063	0.012	-5.109	< 0.001	-0.072	0.015	-4.764	< 0.001
TPJ Left	-0.075	0.012	-6.074	< 0.001	0.03	0.015	1.946	0.052
TPJ Right	-0.083	0.012	-6.727	< 0.001	-0.007	0.015	-0.436	0.663
Vertical Run	-0.003	0.007	-0.411	0.681	0.031	0.009	3.593	< 0.001
DCM * FEF Right	0.007	0.017	0.428	0.669	0	0.021	0.002	0.999
DCM * IPS Left	0.008	0.017	0.483	0.629	-0.001	0.021	-0.033	0.973
DCM * IPS Right	0.005	0.017	0.301	0.763	-0.001	0.021	-0.044	0.965
DCM * TPJ Left	0.01	0.017	0.552	0.581	-0.002	0.021	-0.087	0.93
DCM * TPJ Right	0.009	0.017	0.501	0.617	-0.001	0.021	-0.059	0.953
Vertical Run * DCM	0.008	0.01	0.801	0.423	-0.002	0.012	-0.137	0.891



## Simultaneous modeling of a Spatial Cuing Task

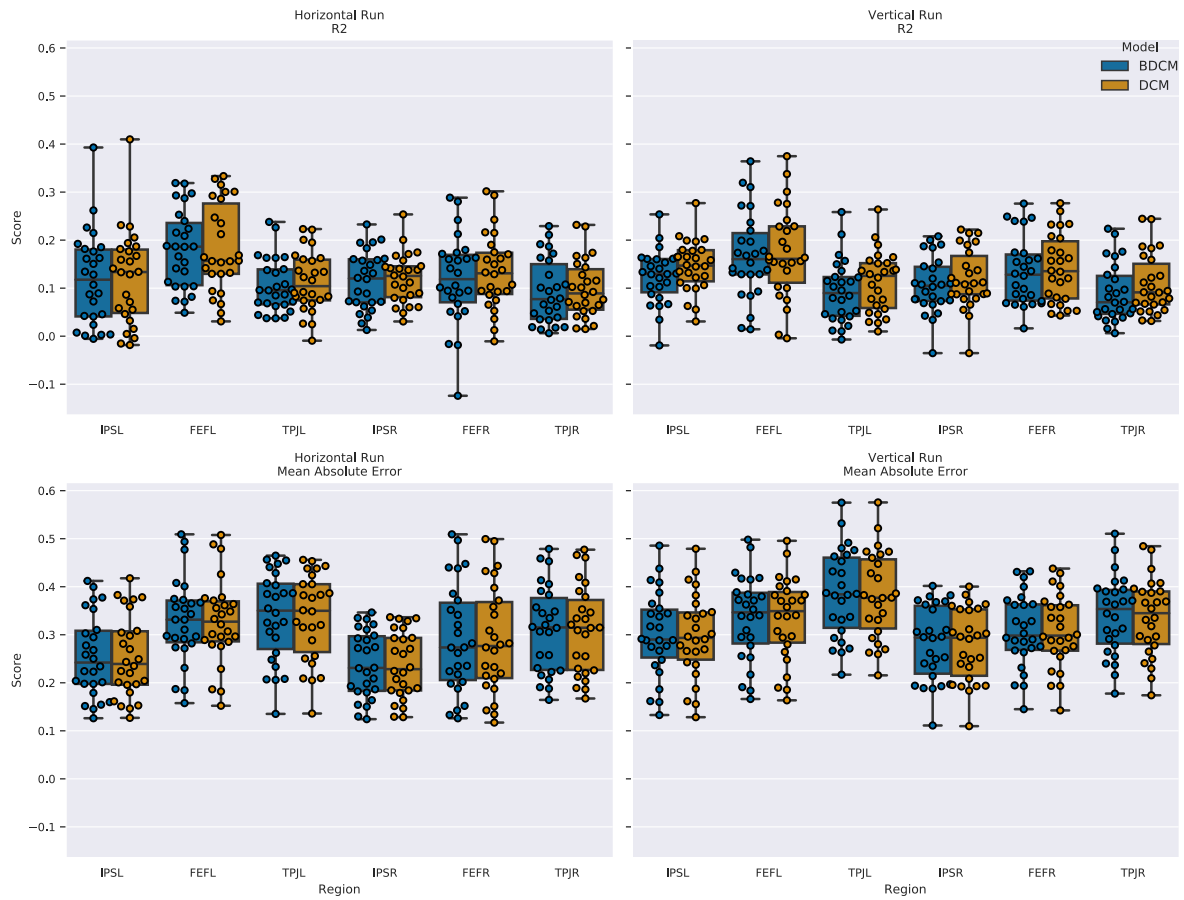


Figure 7: Box plots for the different fit-statistics. Left column shows the data for the horizontal run, the right data for the vertical run. The upper row indicates the R<sup>2</sup> scores and the lower row the mean absolute error.

## Simultaneous modeling of a Spatial Cuing Task

### Generative Embedding

In this analysis, we tested how well the different models could separate the vertical and

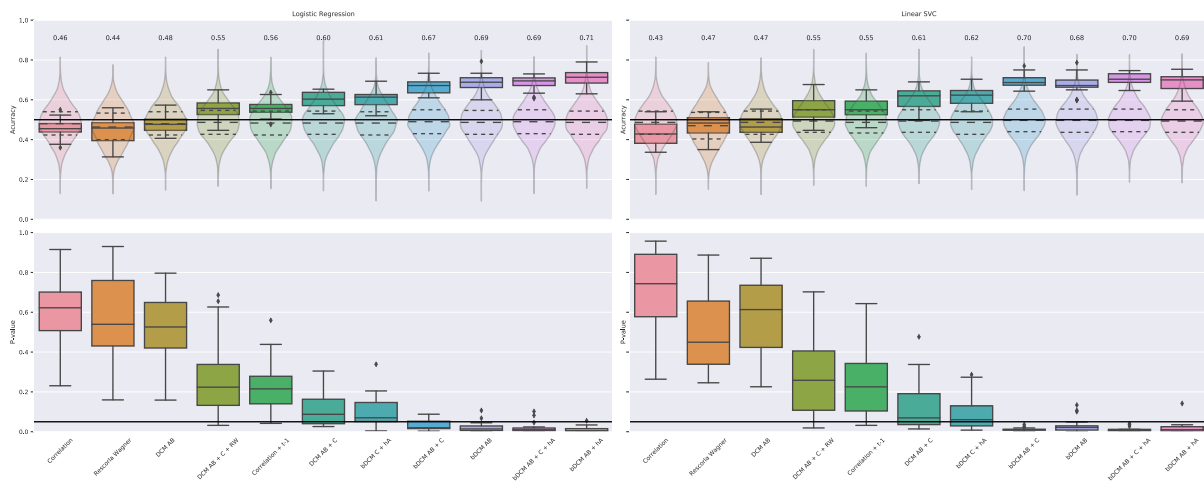


Figure 8: Classifier performance for the 20 shuffles of our dataset, sorted by average performance. In the upper row, the average performance and spread across iterations are shown, and transparent violin-plots indicate the distribution of permutation scores. In the lower panel, the distribution of permutation p-values is indicated, with the black line indicating the  $p < 0.05$  cut-off.

horizontal runs from each other using two different kinds of machine learning algorithms. Due to the high variability in our data (a very low number of training instances, difficult task), we decided to shuffle the assignment of train-test folds so that the cross-validation was performed multiple times. We assessed how often predictions of a classifier were significant (permutation P-Value  $< 0.05$ ) and calculated the average predictive accuracy. The horizontal run could not be separated from the vertical run using combinations of features that were not derived from bDCM parameters (see Figure 8). This indicated that bDCM provided us with some very minute differences between the two runs, that were not included in other models or the BOLD signal.

## Simultaneous modeling of a Spatial Cuing Task

### Lesion Analysis

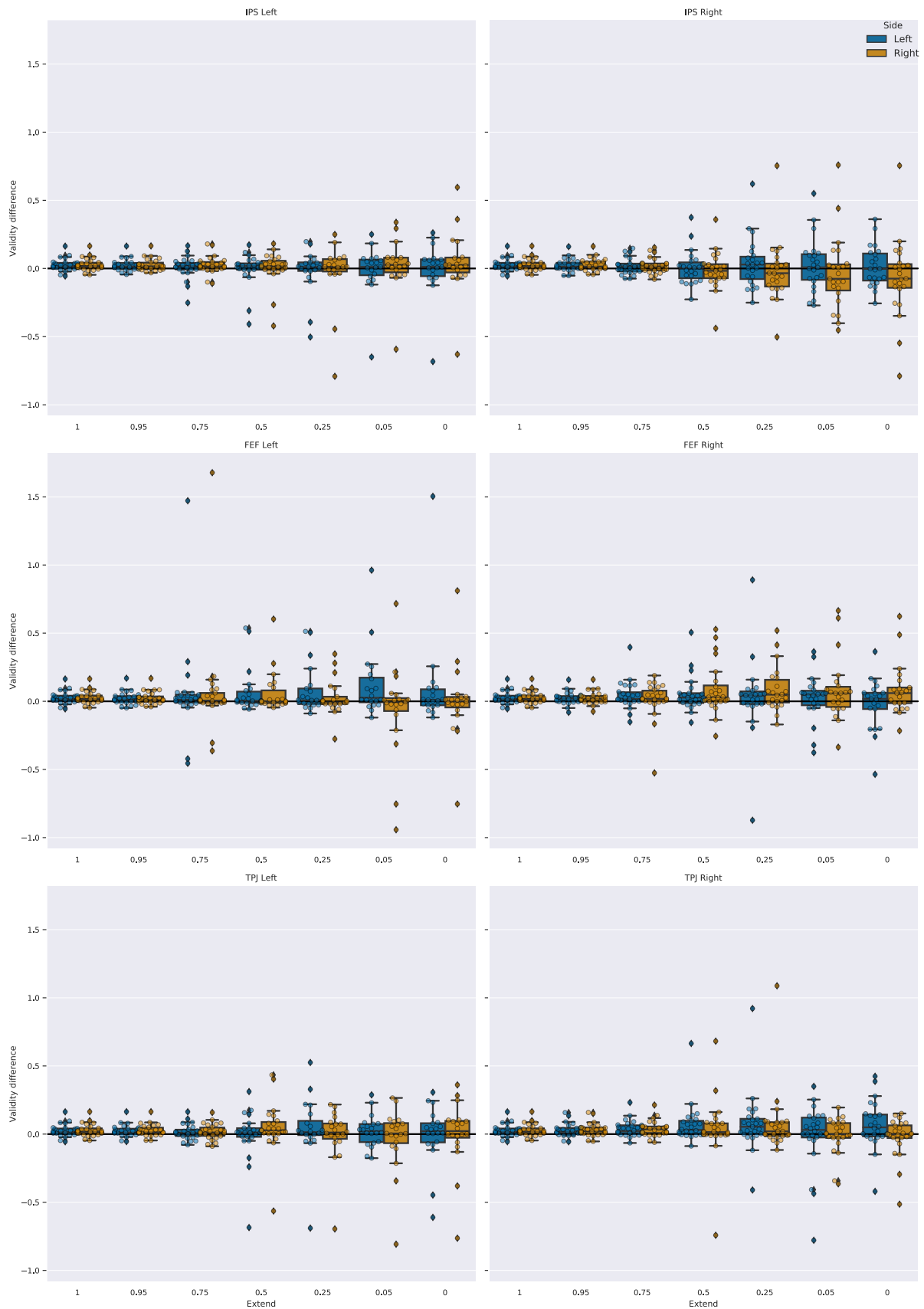


Figure 9: Validity effect for simulated reaction times for the horizontal run, after brain lesions of different extent. An extent of 1 means that no lesion occurred. An extent of 0 indicates that the region was fully disabled. Boxplots indicate the median of the data, the IQR, and the minimum and maximum values. Outliers exceed the  $1.5 * IQR$  criterion.

## Simultaneous modeling of a Spatial Cuing Task

Figure 9 depicts the results of the lesion analysis for the horizontal run (the vertical run can be found in the supplement). Constraining the self-connections in the model resulted in significant decreases in model fit of reaction time data in both the horizontal run (paired t-tests; MAE,  $t(25) = 5.054$ ,  $p < 0.001$ , Cohen's  $d = 0.197$ ;  $R^2$ -score,  $t(25) = -6.573$ ,  $p < 0.001$ , Cohen's  $d = 1.260$ ) and the vertical run (paired t-tests; MAE,  $t(25) = 4.159$ ,  $p < 0.001$ , Cohen's  $d = 0.175$ ;  $R^2$ -score,  $t(25) = -7.185$ ,  $p < 0.001$ , Cohen's  $d = 0.898$ ). Despite this note of caution, there were some interesting trends in the data. The effect of lesion extent on the validity effect seemed highly specific for the different network nodes with some lesions increasing and other lesion decreasing the validity effect, depending on the extend of the artificial lesion.

From a computational anatomy perspective, lesioning the TPJ yielded plausible effects, with a gradual increase in the contralesional validity effect following larger lesions.

## Discussion

We applied bDCM (Rigoux & Daunizeau, 2015) to simultaneously model neural responses and reaction times in a spatial cueing task. We here demonstrated that bDCM could not only be applied to binary responses, but also to continuous read-outs (i.e., reaction times).

After reproducing previously published effects of cue validity at the behavioral and neural level, we modeled behavioral and functional imaging data in three different ways. BDCM, as a novel approach, was compared to both classical DCM and the behavioral Rescorla-Wagner model. As all three models serve different purposes and rely on different data on different timescales, we restricted the model comparison to the models' outputs and fit-statistics.

Although the original paper on bDCM suggested that incorporating behavior also leads to an advantage in representing the BOLD response of bDCM over classical DCM, we did not find significant differences between both modeling approaches. The benefit of including behavioral measures might only be prevalent when BOLD recordings are noisier than behavioral recordings (Rigoux & Daunizeau, 2015).

## Simultaneous modeling of a Spatial Cuing Task

Furthermore, we compared simulated reaction times of bDCM and our implementation of the Rescorla-Wagner model (Vossel, Mathys, et al., 2014). BDCM had in general a better fit to reaction time data (reflected in higher  $R^2$ -score and lower error) and represented the distribution of reaction times more closely in both valid and invalid trials (reflected in significantly lower distances, which were calculated by the Kolmogorov-Smirnov test). Both the Rescorla-Wagner model and bDCM did not model the extreme ends of the reaction distributions well, however, bDCM deviated less from the measured data.

Note that this comparison was not performed to favor one model over the other. Instead, it was conducted to evaluate bDCM against the performance of a highly specialized, validated, and less complex model in a cueing task. Despite the superior fit of bDCM, the Rescorla-Wagner model performed extremely well, given the small number of parameters. Hence, if we penalized for model complexity, the Rescorla-Wagner model would probably be identified as the preferred model for reaction times. BDCM also incorporates the dynamics of the BOLD response and operates on a timescale of seconds, rather than trials. Thus, having only 8 parameters more than the classical DCM (69 parameters) seems to be an adequate increase in complexity. The resulting more detailed representation of reaction time distributions in bDCM might be useful to uncover relevant aspects for assessing cognitive functions as previously demonstrated for other modeling approaches. For example, parameters of drift-diffusion models of reaction times (Smith & Ratcliff, 2009) were found to be related to general intelligence (van Ravenzwaaij et al., 2011) and working memory (Schmiedek et al., 2007). Furthermore, distributional reaction time analysis may categorize healthy participants and patients suffering from psychiatric disorders (Kaiser et al., 2019; Karalunas et al., 2014; Vinogradov et al., 1998). The Rescorla-Wagner model could also be used for such differentiations, especially in the domain of belief-updating (Mengotti et al., 2017). By modeling a single cognitive process, however, the Rescorla-Wagner model is very dependent on the presence and size of a participant's validity effects (see analysis in S3, showing that the correlation between model fit and cue-validity are higher for Rescorla-Wagner than bDCM).

## Simultaneous modeling of a Spatial Cuing Task

BDCM, in contrast, simulated smoother reaction time distributions (larger number of non-significant p-values in KS-test), possibly providing a richer representation of the underlying processes. Although bDCM may reflect a portion of variance in the reaction time data that is not task-related, this variance could reflect the processes of belief-updating in a more complex brain-dynamics-dependent matter. BDCM is a *model of brain dynamics* that can, in principle, be applied to any task, while the Rescorla-Wagner model represents a specialized *model of a cognitive process*.

As bDCM can be applied to model different behavioral read-outs in various tasks, it can enhance our understanding of how DCM's connectivity parameters relate to behavior. So far, this link could only be established using indirect methods, such as correlations between DCM parameters and behavioral measures across participants. For example, DCM's task connectivity parameters have been related to symptoms of depression and schizophrenia (Desseilles et al., 2011; Schlösser et al., 2008; Wu et al., 2014), and have been correlated with behavioral measures before and after interventions using non-invasive neurostimulation (Grefkes et al., 2010). Although the investigation of such associations does not allow causal interpretations, bDCM enables more firm conclusions how brain dynamics in selected brain regions impact behavior.

Furthermore, brain and behavioral dynamics both regularize bDCM, so that the model parameters encode the most reliable set of information from both sources (Rigoux & Daunizeau, 2015). This procedure could yield more robust and stable connectivity estimates but also encode more specific information. This may be particularly relevant for so-called “generative embedding” approaches, where a generative model and its estimated parameters are used as a form of dimensionality reduction (Brodersen, Haiss, et al., 2011; Brodersen, Schofield, et al., 2011). In fact, this was confirmed by our findings, where it was possible to differentiate the horizontal from the vertical run only when using the bDCM model's connectivity parameters. This makes bDCM a unique approach for the identification of biomarkers that are relevant for certain behaviors – provided that they are stable across participants and sessions (Elliott et al., 2020).

## Simultaneous modeling of a Spatial Cuing Task

Since bDCM is a generative model, it can also be used to simulate how alterations to the underlying brain network might change behavior (Rigoux & Daunizeau, 2015). This allows simulating the behavioral effects of neuromodulatory interventions and the generation of new hypotheses and experiments. The guidance and information of computational models will eventually lead to a better understanding of the neural mechanisms underlying behavioral outcomes (Kriegeskorte & Douglas, 2018; Turner et al., 2017).

Unfortunately, applying artificial lesions to the network model in our study revealed technical problems of this approach. More specifically, the estimated models lacked numerical stability and required manual intervention, which substantially changed the model's output. Even though some of the resulting patterns were consistent with the literature (e.g., an increase of the contralesional validity effect after a lesion to right TPJ and to a lesser extent in left TPJ (Malherbe et al., 2018; Posner et al., 1984)), other simulations were highly variable. Hence, the relatively novel bDCM approach's potential problems, such as over-fitting and non-generalizability, need to be considered in future studies.

## Conclusion

bDCM was applied for the first time to reaction time data of a larger sample of participants. Our findings provided evidence for a considerable additional value of the method compared to a purely behavioral model and classical DCM and identified practical use issues. Data suggest that bDCM is indeed a promising tool to enhance our understanding of how brain dynamics generate specific behavioral patterns.

## References

- Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., Gramfort, A., Thirion, B., & Varoquaux, G. (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics*, 8. <https://doi.org/10.3389/fninf.2014.00014>
- Brodersen, K. H., Haiss, F., Ong, C. S., Jung, F., Tittgemeyer, M., Buhmann, J. M., Weber, B., & Stephan, K. E. (2011). Model-based feature construction for multivariate

## Simultaneous modeling of a Spatial Cuing Task

decoding. *NeuroImage*, *56*(2), 601–615.

<https://doi.org/10.1016/j.neuroimage.2010.04.036>

Brodersen, K. H., Schofield, T. M., Leff, A. P., Ong, C. S., Lomakina, E. I., Buhmann, J. M., & Stephan, K. E. (2011). Generative Embedding for Model-Based Classification of fMRI Data. *PLoS Computational Biology*, *7*(6), e1002079.

<https://doi.org/10.1371/journal.pcbi.1002079>

Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, *97*(4), 523–547.

<https://doi.org/10.1037/0033-295x.97.4.523>

Corbetta, M., Kincade, M. J., Lewis, C., Snyder, A. Z., & Sapir, A. (2005). Neural basis and recovery of spatial attention deficits in spatial neglect. *Nature Neuroscience*, *8*(11), 1603–1610.

<https://doi.org/10.1038/nn1574>

Corbetta, M., & Shulman, G. L. (2011). Spatial Neglect and Attention Networks. *Annual Review of Neuroscience*, *34*(1), 569–599. <https://doi.org/10.1146/annurev-neuro-061010-113731>

Crosse, M. J., Di Liberto, G. M., Bednar, A., & Lalor, E. C. (2016). The Multivariate Temporal Response Function (mTRF) Toolbox: A MATLAB Toolbox for Relating Neural Signals to Continuous Stimuli. *Frontiers in Human Neuroscience*, *10*.

<https://doi.org/10.3389/fnhum.2016.00604>

Daunizeau, J., Adam, V., & Rigoux, L. (2014). VBA: A Probabilistic Treatment of Nonlinear Models for Neurobiological and Behavioural Data. *PLoS Computational Biology*, *10*(1), e1003441. <https://doi.org/10.1371/journal.pcbi.1003441>

Desseilles, M., Schwartz, S., Dang-Vu, T. T., Sterpenich, V., Ansseau, M., Maquet, P., & Phillips, C. (2011). Depression alters “top-down” visual attention: A dynamic causal modeling comparison between depressed and healthy subjects. *NeuroImage*, *54*(2), 1662–1668. <https://doi.org/10.1016/j.neuroimage.2010.08.061>

Dombert, P. L., Kuhns, A., Mengotti, P., Fink, G. R., & Vossel, S. (2016). Functional mechanisms of probabilistic inference in feature- and space-based attentional



## Simultaneous modeling of a Spatial Cuing Task

systems. *NeuroImage*, 142, 553–564.

<https://doi.org/10.1016/j.neuroimage.2016.08.010>

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., & Zilles, K. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage*, 25(4), 1325–1335.

<https://doi.org/10.1016/j.neuroimage.2004.12.034>

Elliott, M. L., Knodt, A. R., Ireland, D., Morris, M. L., Poulton, R., Ramrakha, S., Sison, M. L., Moffitt, T. E., Caspi, A., & Hariri, A. R. (2020). What Is the Test-Retest Reliability of Common Task-Functional MRI Measures? New Empirical Evidence and a Meta-Analysis. *Psychological Science*, 095679762091678.

<https://doi.org/10.1177/0956797620916786>

Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., Kent, J. D., Goncalves, M., DuPre, E., Snyder, M., Oya, H., Ghosh, S. S., Wright, J., Durnez, J., Poldrack, R. A., & Gorgolewski, K. J. (2019). fMRIPrep: A robust preprocessing pipeline for functional MRI. *Nature Methods*, 16(1), 111–116.

<https://doi.org/10.1038/s41592-018-0235-4>

Friston, K. J. (Ed.). (2007). *Statistical parametric mapping: The analysis of functional brain images* (1st ed). Elsevier/Academic Press.

Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. *NeuroImage*, 19(4), 1273–1302. [https://doi.org/10.1016/S1053-8119\(03\)00202-7](https://doi.org/10.1016/S1053-8119(03)00202-7)

Friston, K. J., Preller, K. H., Mathys, C., Cagnan, H., Heinzle, J., Razi, A., & Zeidman, P. (2017). Dynamic causal modelling revisited. *NeuroImage*.

<https://doi.org/10.1016/j.neuroimage.2017.02.045>

Friston, K. J., Williams, S., Howard, R., Frackowiak, R. S. J., & Turner, R. (1996). Movement-Related effects in fMRI time-series. *Magnetic Resonance in Medicine*, 35(3), 346–355. <https://doi.org/10.1002/mrm.1910350312>

Gorgolewski, K. J., Varoquaux, G., Rivera, G., Schwarz, Y., Ghosh, S. S., Maumet, C., Sochat, V. V., Nichols, T. E., Poldrack, R. A., Poline, J.-B., Yarkoni, T., & Margulies,

## Simultaneous modeling of a Spatial Cuing Task

- D. S. (2015). NeuroVault.org: A web-based repository for collecting and sharing unthresholded statistical maps of the human brain. *Frontiers in Neuroinformatics*, 9. <https://doi.org/10.3389/fninf.2015.00008>
- Grefkes, C., Nowak, D. A., Wang, L. E., Dafotakis, M., Eickhoff, S. B., & Fink, G. R. (2010). Modulating cortical connectivity in stroke patients by rTMS assessed with fMRI and dynamic causal modeling. *NeuroImage*, 50(1), 233–242. <https://doi.org/10.1016/j.neuroimage.2009.12.029>
- Hedge, C., Powell, G., & Sumner, P. (2017). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods*. <https://doi.org/10.3758/s13428-017-0935-1>
- Kaiser, D., Haerberle, G., & Cichy, R. M. (2019). *Cortical sensitivity to natural scene structure* [Preprint]. Neuroscience. <https://doi.org/10.1101/613885>
- Karalunas, S. L., Geurts, H. M., Konrad, K., Bender, S., & Nigg, J. T. (2014). Annual Research Review: Reaction time variability in ADHD and autism spectrum disorders: measurement and mechanisms of a proposed trans-diagnostic phenotype. *Journal of Child Psychology and Psychiatry*, 55(6), 685–710. <https://doi.org/10.1111/jcpp.12217>
- Kriegeskorte, N., & Douglas, P. K. (2018). Cognitive computational neuroscience. *Nature Neuroscience*. <https://doi.org/10.1038/s41593-018-0210-5>
- Malherbe, C., Umarova, R. M., Zavaglia, M., Kaller, C. P., Beume, L., Thomalla, G., Weiller, C., & Hilgetag, C. C. (2018). Neural correlates of visuospatial bias in patients with left hemisphere stroke: A causal functional contribution analysis based on game theory. *Neuropsychologia*, 115, 142–153. <https://doi.org/10.1016/j.neuropsychologia.2017.10.013>
- Mengotti, P., Dombert, P. L., Fink, G. R., & Vossel, S. (2017). Disruption of the Right Temporoparietal Junction Impairs Probabilistic Belief Updating. *The Journal of Neuroscience*, 37(22), 5419–5428. <https://doi.org/10.1523/JNEUROSCI.3683-16.2017>

## Simultaneous modeling of a Spatial Cuing Task

- Nichols, T. E., & Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Human Brain Mapping, 15*(1), 1–25.
- Nunez, M. D. (2015). Individual differences in attention influence perceptual decision making. *Frontiers in Psychology, 6*. <https://doi.org/10.3389/fpsyg.2015.00018>
- Ojala, M., & Garriga, G. C. (2010). Permutation Tests for Studying Classifier Performance. *J. Mach. Learn. Res., 11*, 1833–1863.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97–113. [https://doi.org/10.1016/0028-3932\(71\)90067-4](https://doi.org/10.1016/0028-3932(71)90067-4)
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., & others. (2011). Scikit-learn: Machine learning in Python. *The Journal of Machine Learning Research, 12*, 2825–2830.
- Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods, 162*(1–2), 8–13. <https://doi.org/10.1016/j.jneumeth.2006.11.017>
- Peirce, J. W. (2008). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics, 2*. <https://doi.org/10.3389/neuro.11.010.2008>
- Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods, 51*(1), 195–203. <https://doi.org/10.3758/s13428-018-01193-y>
- Posner, M. I. (1980). Orienting of attention. *The Quarterly Journal of Experimental Psychology, 32*(1), 3–25.
- Posner, M. I., Walker, J. A., Friedrich, F. J., & Rafal, R. D. (1984). Effects of parietal injury on covert orienting of attention. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 4*(7), 1863–1874.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*(2), 59–108. <https://doi.org/10.1037/0033-295X.85.2.59>

## Simultaneous modeling of a Spatial Cuing Task

Rescorla, R. A., Wagner, A. R., & others. (1972). A theory of Pavlovian conditioning:

Variations in the effectiveness of reinforcement and nonreinforcement. *Classical Conditioning II: Current Research and Theory*, 2, 64–99.

Rigoux, L., & Daunizeau, J. (2015). Dynamic causal modelling of brain–behaviour relationships. *NeuroImage*, 117, 202–221.

<https://doi.org/10.1016/j.neuroimage.2015.05.041>

Schlösser, R. G. M., Wagner, G., Koch, K., Dahnke, R., Reichenbach, J. R., & Sauer, H.

(2008). Fronto-cingulate effective connectivity in major depression: A study with fMRI and dynamic causal modeling. *NeuroImage*, 43(3), 645–655.

<https://doi.org/10.1016/j.neuroimage.2008.08.002>

Schmiedek, F., Oberauer, K., Wilhelm, O., Süß, H.-M., & Wittmann, W. W. (2007). Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *Journal of Experimental Psychology: General*, 136(3), 414–429. <https://doi.org/10.1037/0096-3445.136.3.414>

Shaw, D., Czekóová, K., Gajdoš, M., Staněk, R., Špalek, J., & Brázdil, M. (2019). Social decision-making in the brain: Input-state-output modelling reveals patterns of effective connectivity underlying reciprocal choices. *Human Brain Mapping*, 40(2), 699–712.

<https://doi.org/10.1002/hbm.24446>

Smith, P. L., & Ratcliff, R. (2009). An integrated theory of attention and decision making in visual signal detection. *Psychological Review*, 116(2), 283–317.

<https://doi.org/10.1037/a0015156>

Steinkamp, S. R., Vessel, S., Fink, G. R., & Weidner, R. (2020). Attentional reorientation along the meridians of the visual field: Are there different neural mechanisms at play?

*Human Brain Mapping*, hbm.25086. <https://doi.org/10.1002/hbm.25086>

Stephan, K. E., Kasper, L., Harrison, L. M., Daunizeau, J., den Ouden, H. E. M., Breakspear, M., & Friston, K. J. (2008). Nonlinear dynamic causal models for fMRI. *NeuroImage*, 42(2), 649–662. <https://doi.org/10.1016/j.neuroimage.2008.04.262>

## Simultaneous modeling of a Spatial Cuing Task

Turner, B. M., Forstmann, B. U., Love, B. C., Palmeri, T. J., & Van Maanen, L. (2017).

Approaches to analysis in model-based cognitive neuroscience. *Journal of Mathematical Psychology*, 76, 65–79. <https://doi.org/10.1016/j.jmp.2016.01.001>

Vallat, R. (2018). Pingouin: Statistics in Python. *Journal of Open Source Software*, 3(31), 1026. <https://doi.org/10.21105/joss.01026>

van Ravenzwaaij, D., Brown, S., & Wagenmakers, E.-J. (2011). An integrated perspective on the relation between response speed and intelligence. *Cognition*, 119(3), 381–393. <https://doi.org/10.1016/j.cognition.2011.02.002>

Vinogradov, S., Poole, J. H., Willis-Shore, J., Ober, B. A., & Shenaut, G. K. (1998). Slower and more variable reaction times in schizophrenia: What do they signify? *Schizophrenia Research*, 32(3), 183–190. [https://doi.org/10.1016/S0920-9964\(98\)00043-7](https://doi.org/10.1016/S0920-9964(98)00043-7)

Vossel, S., Geng, J. J., & Fink, G. R. (2014). Dorsal and Ventral Attention Systems: Distinct Neural Circuits but Collaborative Roles. *The Neuroscientist*, 20(2), 150–159. <https://doi.org/10.1177/1073858413494269>

Vossel, S., Mathys, C., Daunizeau, J., Bauer, M., Driver, J., Friston, K. J., & Stephan, K. E. (2014). Spatial Attention, Precision, and Bayesian Inference: A Study of Saccadic Response Speed. *Cerebral Cortex*, 24(6), 1436–1450. <https://doi.org/10.1093/cercor/bhs418>

Vossel, S., Weidner, R., Driver, J., Friston, K. J., & Fink, G. R. (2012). Deconstructing the Architecture of Dorsal and Ventral Attention Systems with Dynamic Causal Modeling. *Journal of Neuroscience*, 32(31), 10637–10648. <https://doi.org/10.1523/JNEUROSCI.0414-12.2012>

Vossel, S., Weidner, R., Moos, K., & Fink, G. R. (2016). Individual attentional selection capacities are reflected in interhemispheric connectivity of the parietal cortex. *NeuroImage*, 129, 148–158. <https://doi.org/10.1016/j.neuroimage.2016.01.054>

Wu, G., Wang, Y., Mwansisya, T. E., Pu, W., Zhang, H., Liu, C., Yang, Q., Chen, E. Y. H., Xue, Z., Liu, Z., & Shan, B. (2014). Effective connectivity of the posterior cingulate

## Simultaneous modeling of a Spatial Cuing Task

and medial prefrontal cortices relates to working memory impairment in schizophrenic and bipolar patients. *Schizophrenia Research*, 158(1–3), 85–90.

<https://doi.org/10.1016/j.schres.2014.06.033>

Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods*, 8(8), 665–670. <https://doi.org/10.1038/nmeth.1635>