Computational modeling of neural tuning during arithmetic problem solving

A computational model was used to account for spatial functional activation patterns of addition and subtraction problems in DD and TD children. The dynamical model originally introduced by Usher and Cohen (1999) has been previously used to show that artificial neural ensembles can implement a salience map for visual stimuli, and it has been suggested that the PPC may demonstrate similar tuning properties for numerosities (Knops et al., 2014; Roggeman et al., 2010; Sengupta et al., 2014). Here, we employed a static version of this model to assess the contributions of model parameters associated with neural tuning to the spatial activation patterns observed during task-based fMRI. ROIs were obtained from significant clusters resulting from the group-level MRS analysis (**Figure 1, Table 2**). Voxel-wise effects for these ROIs were next extracted from each individual's maps for the contrast: Complex Addition - Simple Addition; Complex Subtraction - Simple Subtraction (**Figure 2a**). We model only voxels within each ROI that showed an effect for Complex problems. To estimate neural activation patterns, each resulting voxel, x_i, was input to the following equation:

$$\hat{x}_i = \alpha F(x_i) - \beta \sum_{j=1, j \neq i}^{N} F(x_j) + noise$$

such that optimal parameter values minimize the difference between $\hat{x}_i$ and x_i. Here, α refers to the strength of self-excitation for each voxel and β to the strength of lateral inhibition between voxels. Voxels x_j are defined as the N neighboring voxels within a 6mm sphere, regardless of effect directionality, around x_i to account for spatial smoothing. Edge cases were handled by extending beyond the boundary defined by each ROI mask. Noise was generated for each voxel from a normal distribution with a mean of zero and a standard deviation of 0.05 (Knops et al., 2014). $F(x)$ refers to the sigmoidal activation function defined by:

$$F(x) = \begin{cases} 0, for\ x \leq 0 \\ \dfrac{x}{1+x}, for\ x > 0 \end{cases}$$

Given the high degree of individual variability in DD (Iuculano, Tang, Hall, & Butterworth, 2008; Rubinsten & Henik, 2009; Skagerlund & Traff, 2016; Traff, Olsson, Ostergren, & Skagerlund, 2016), we used hierarchical Bayesian inference, which better characterizes heterogeneous samples compared to other estimation methods (Lee, 2011). A graphical example using plate notation for this procedure is shown in **Figure 2b**. We first defined weakly informative priors as $\alpha \sim Normal(\alpha_\mu, \alpha_\sigma)$, $0 \leq \alpha \leq 5$ and $\beta \sim Normal(\beta_\mu, \beta_\sigma)$, $0 \leq \beta \leq 1$ such that α and β are truncated at boundaries within the numerical ranges found by previous simulations of these parameters (Roggeman et al., 2010; Sengupta et al., 2014). Hyperpriors $\alpha_\mu,\ \beta_\mu \sim Normal(0,1)$ were then used to estimate a sample mean. Model likelihood was defined as $x \sim Normal(x_\mu, x_\sigma)$. Uninformative uniform priors were used to estimate σ for all parameters.

Bayesian inference was performed using PyMC3 (Salvatier, Wiecki, & Fonnesbeck, 2016) using No-U-Turn Sampling (NUTS; Hoffman & Gelman, 2014) initialized with auto-diff variational inference (ADVI) with 10 chains at 5,000 iterations each. This approach estimates possible parameter values until convergence on values, which most closely approximate the input data.

Model convergence was assessed using the Gelman-Rubin convergence diagnostic $\widehat{R}$ which compares the estimated between-chains and within-chain variances for each model parameter (Gelman & Rubin, 1992). Because hierarchical models violate the assumption of independence mandated by frequentist statistics, a standard t-test using the individual subject parameters was not performed. Instead, Bayesian hypothesis testing (Krushke, 2010; Cavanagh, Wiecki, Kochar, & Frank, 2014) was conducted by comparing the probability mass of each parameter between conditions, as estimated by the group mean posterior distributions (**Figure 2c**). Model fit was assessed via posterior predictive checks, which evaluate the model's ability to generate new data similar to the input data (Gelman, Meng, & Stern, 1996). Here we generated 500 posterior predictive datasets from the fit model to assess significant effects.