Supplemental figures

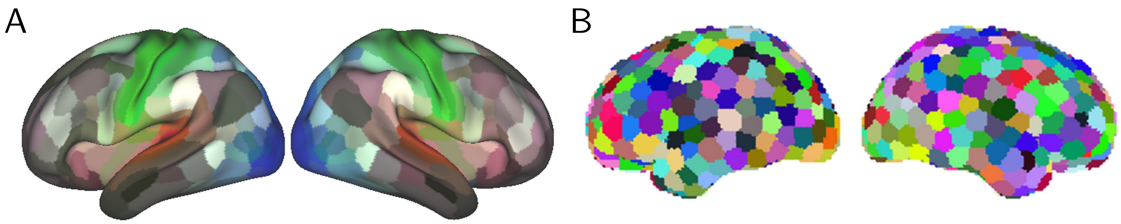

Fig. S1: Region of interest (ROI) masks. (A) For Human Connectome Project data, 360 cortical ROIs as provided by [Glasser et al., 2016] were used. (B) For movie data, 500 cortical ROIs obtained from using k-means clustering on the spatial coordinates $\{x, y, z\}$ of cortical voxels were used. Two amygdala regions (one per hemisphere) were also included but are not show here.

Type	Run	Name (Year)	Length
Funny	1	Step Brothers (2008)	2:09
	2	Wedding Crashers (2005)	2:03
	3	Night at the Museum: 2 (2009)	3:01
	4	Patch Adams (1998)	1:40
	5	The Elevator (2010)	3:02
	6	Bruno (2009)	2:21
Scary	1	The Eye (2002)	3:01
	2	Paranormal Activity 3 (2011)	2:05
	3	Entry 18 (2009)	2:24
	4	The Blair Witch Project (1999)	3:02
	5	The Others (2001)	2:03
	6	Shutter Island (2010)	1:42

Table S1: Film names and clip duration for the "scary" and "funny" conditions.

Working memory: "2-back" vs. "0-back"		
	Accuracy on the "first" set	Accuracy on the "second" set
Reservoirs	86.5%	86.3%
Raw activation	78.9%	77.6%
Concatenation	82.0%	82.8%
Auto-regressive model	81.2%	81.1%
Theory of mind: "social" vs. "random"		
	Accuracy on the "first" set	Accuracy on the "second" set
Reservoirs	91.8%	91.9%
Raw activation	83.4%	84.2%
Concatenation	85.9%	87.8%
Auto-regressive model	87.2%	86.6%

Table S2: Comparison of mean cross-validation accuracy using the "first" dataset and classification accuracy on the "second" dataset. The similar results on the two datasets indicate the robustness of the cross-validation scheme.

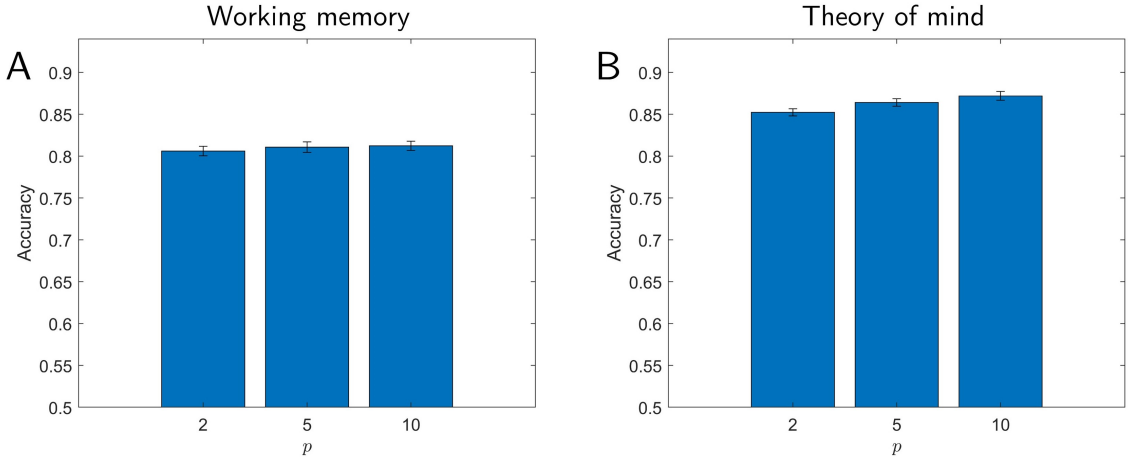

Fig. S2: Classification accuracy using autoregressive models for working memory (A) and theory of mind (B). Results are shown as a function of model order, p.

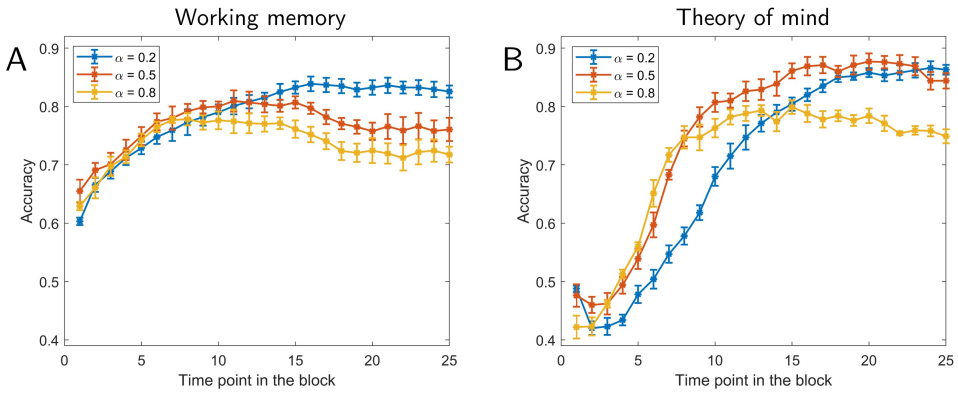

Fig. S3: Classification accuracy as a function of time using only the low-dimensional data (10 top/bottom principal components for working memory, and 12 top/bottom principal components for theory of mind). Results for working memory (A) and theory of mind (B). Accuracy is shown as a function of time point within a task block. Different curves show results for different forgetting rates, α. The values of τ were based on the parameters exhibiting highest accuracy in Fig. 3. Error bars show the standard error of the mean.

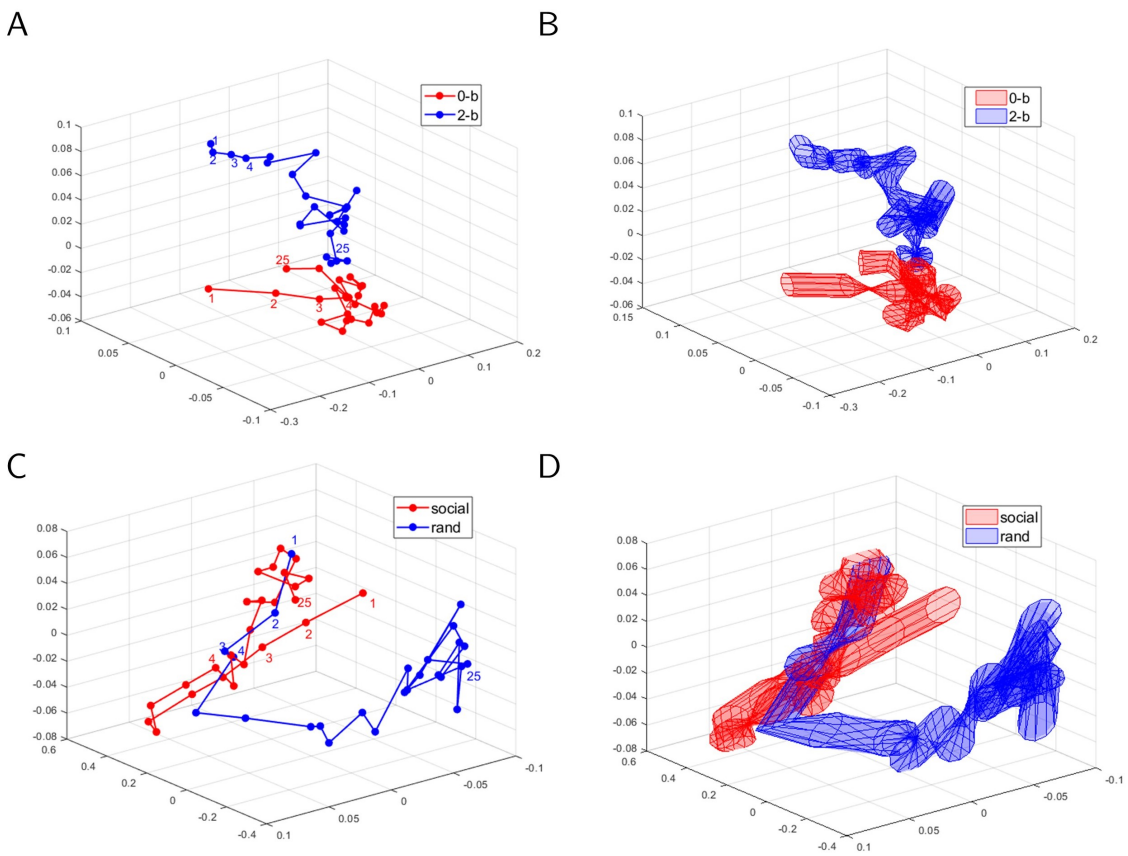

Fig. S4: Temporal trajectories based on the "top" three principal components of the input time series (that is, no reservoir) for fMRI task data. Mean trajectories are displayed in (A) for working memory and (C) for theory of mind. Variability (standard error across participants) is displayed in (B) and (D), respectively. For working memory data (A-B), these trajectories were well separated throughout the block. However, for theory of mind data (C-D), the trajectory for the social condition did not evolve temporally as seen when using reservoirs; note that the final states (see point 25) were close to the initial ones (see points 1-2). Therefore, it appears that a low-dimensional representation based directly on input activations does not adequately capture the temporal evolution structure associated with the social condition.