The Surprising Role of the Default Mode Network

T. Brandman[1]*, R. Malach[1], E. Simony[1,2].

[1]Department of Neurobiology and the Azrieli National Institute for Human Brain Imaging and Research, Weizmann Institute of Science.

[2]Faculty of Engineering, Holon Institute of Technology.

*Correspondence to: talli.brandman@gmail.com

Abstract: The default mode network (DMN) is a group of high-order brain regions recently implicated in processing external naturalistic events, yet it remains unclear what cognitive function it serves. Here we identified the cognitive states predictive of DMN fMRI coactivation. Particularly, we developed a *state-fluctuation pattern analysis*, matching network coactivations across a short movie with retrospective behavioral sampling of movie events. Network coactivation was selectively correlated with the state of *surprise* across movie events, compared to all other cognitive states (e.g. emotion, vividness). The effect was exhibited in the DMN, but not dorsal attention or visual networks. Furthermore, surprise was found to mediate DMN coactivations with hippocampus and nucleus accumbens. These unexpected findings point to the DMN as a major hub in high-level prediction-error representations.

The default mode network (DMN) is a group of high-order brain regions, so-called for its decreased activation during tasks of high attentional demand, relative to the high baseline activation of the DMN at rest (*1-3*). Much research has been conducted in the pursuit of the enigmatic role of this network, consistently pointing to DMN activity during internal processes such as mind wondering, mental time travel and perspective shifting (*4-6*). However, recent neuroimaging studies suggest that the DMN is important not only for internally-driven processes, but remarkably, for long-time scale naturalistic processing of real-life events (*7-11*), making it central to understanding how our brain tackles incoming information during everyday life. This discovery was enabled by computational advancements in the analysis of neuroimaging signals, which now allow us to track the dynamics of continuous naturalistic processing in healthy human brains, noninvasively (*9,*

12). Such studies have shown that dynamic responses of the DMN carry information about long timescales of narrative content, and may be associated with subsequent memory of it (*7-10*). Yet it remains unknown what are the specific roles of the DMN in naturalistic cognition.

The difficulty in pinpointing the cognitive processes reflected by DMN responses during naturalistic stimulation, lies in connecting between the dynamic cognitive state and DMN activity. Here we developed a new approach of *state-fluctuation pattern analysis* (SFPA) to directly relate the two. Specifically, we modeled the cognitive state along the time-course of a movie stimulus using a technique we term *retrospective behavioral sampling* (see Materials and Methods), and compared each cognitive measure to the temporal patterns of neural responses evoked by the same movie. Critically, we employed our previous discovery that task-driven DMN coactivation can be revealed by employing inter-subject functional correlation (ISFC) (*9*). Using the new methodology to systematically link ISFC to behavior, we were able to show that the cognitive measure that best fits DMN coactivation dynamics is the level of *surprise* induced by movie events. We further demonstrate surprise-dependent DMN coactivation with subcortical regions implicated in predictive processing (*13-16*). This study therefore highlights a new role of the DMN, as a central hub in prediction-error representation of ongoing real-life events, likely involving the temporal integration of incoming information with representations stored in memory.

Results

Cognitive dynamics were modeled from behavioral responses of 45 participants to the first episode of Sherlock (BBC series, 2010), sampling 49 events of the movie on measures of surprise, vividness of memory, emotional intensity and valence, perceived importance, episodic memory and theory of mind. Neural dynamics of coactivation (i.e. activity correlations across brain regions), were modeled from functional magnetic resonance imaging (fMRI) responses of 35 participants to the same movie (*8, 10*), in regions of the DMN and hippocampus, as well as the dorsal attention network (DAN) and visual-processing areas (Vis). Since the DMN manifests spontaneous fluctuations both at rest and at task (*17-19*), we used ISFC to eliminate these spontaneous signals and extract the shared component of stimulus-induced coactivation across brain regions and across individuals (*9*). Our approach was thus optimized for matching across temporal response patterns of brain and behavior to a dynamic naturalistic input (Figure 1).

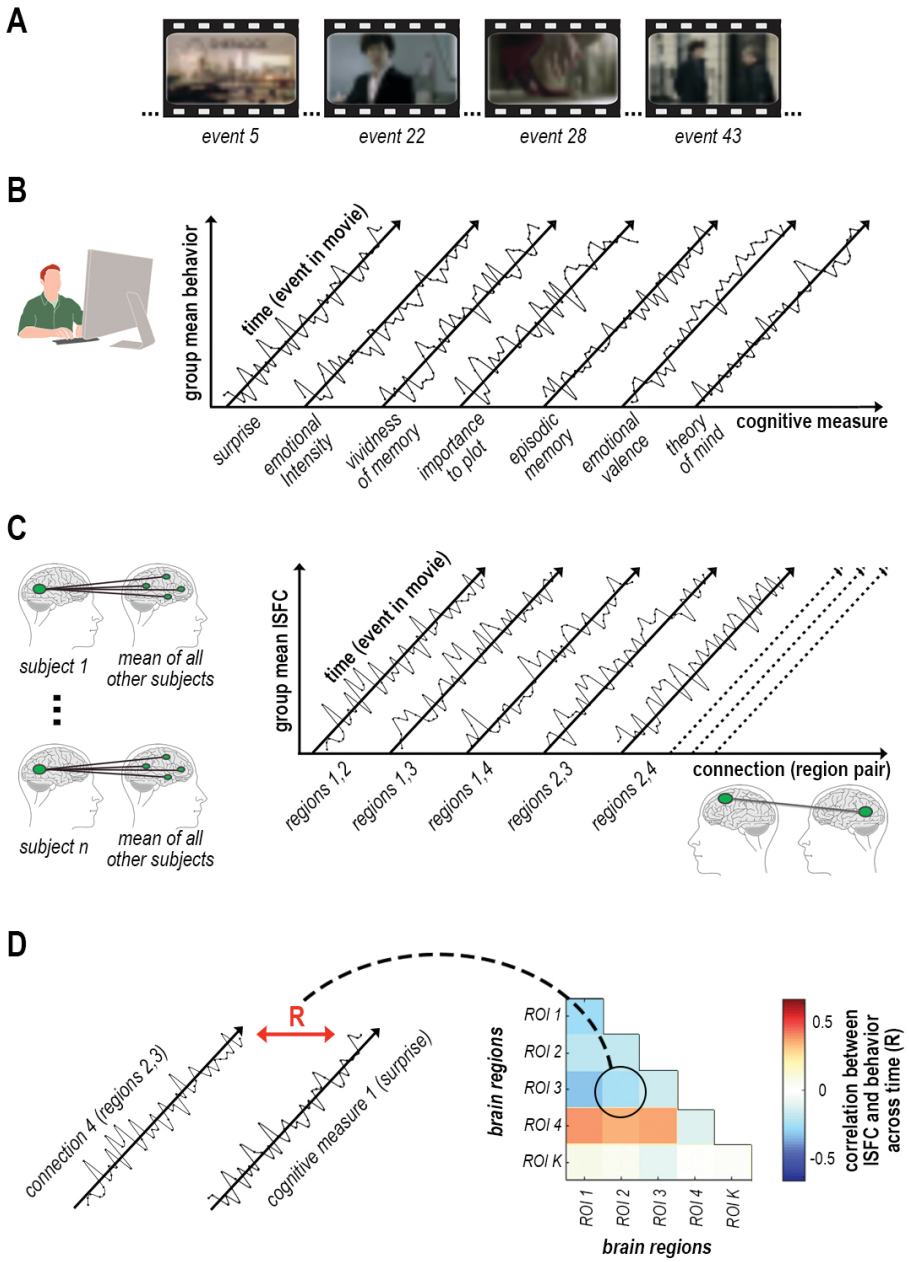

Figure 1. SFPA of naturalistic narrative processing. (A) A 23-minute-long excerpt of Sherlock (BBC, 2010) was viewed during either fMRI or web-based participation. Retrospective behavioral sampling was performed on 49 movie events. (B) Following web-based viewing, each event was probed verbally in a questionnaire, retrospectively measuring the fluctuation in 7 cognitive states throughout the movie. (C) ISFC corresponding to each movie event ±7 TR was measured by correlating the time-window fMRI signal between every region in each subject and every other region in all other subjects, measuring the fluctuation in coactivation throughout the movie. (D) Temporal patterns of cognitive states and ISFC were tested for correlation across movie events.

DMN coactivation, both within cortical DMN regions and between DMN and hippocampus, fluctuated proportionally to the magnitude of surprise, but not other behavioral measures. Particularly, SFPA revealed significant correlations (via permutation test; $p < 0.05$, corrected)

between surprise ratings and ISFC among DMN region pairs (Figure 2). The overall correlation between surprise ratings and ISFC mean across all DMN region pairs was r(47) = 0.44 (p = 0.001, 95% CI [0.18, 0.64]). In addition, surprise ratings were significantly correlated with ISFC between DMN regions and hippocampus (perm. p < 0.05, corrected). By contrast, surprise ratings did not correlate with pairwise ISFC in DAN and Vis (perm. p > 0.05, corrected). The overall correlation between surprise ratings and ISFC mean across all regions in DAN was r(47) = 0.03 (p = 0.859, 95% CI [-0.25, 0.31]) and in Vis was r(47) = -0.08 (p = 0.589, 95% CI [-0.35, 0.21]). Furthermore, no significant correlations were found between ISFC in the DMN and other behavioral measures (perm. p > 0.05, corrected; Supplementary Figure 1). Notably, a similar pattern of correlations between ISFC and surprise was also found in participants who had watched a short thriller movie (perm. p < 0.05; Supplementary Figure 2), yet it was confounded with emotional intensity (r(34) = 0.92, p < 0.001, 95% CI [0.85, 0.96]), thus making it uninformative for continued analysis (see Supplementary Note 1).

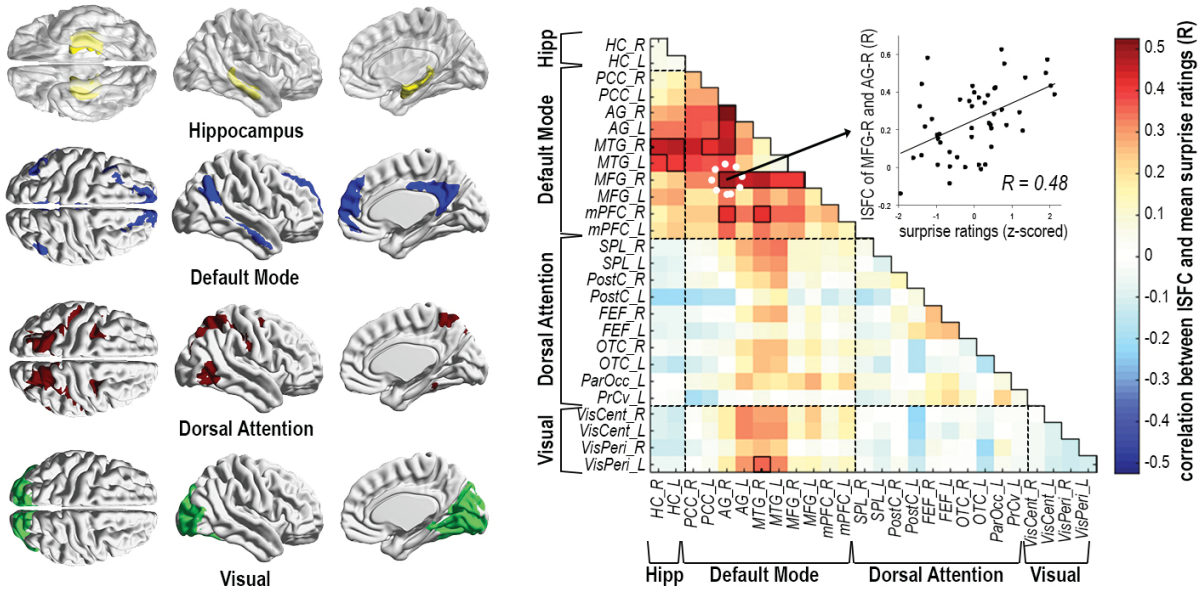

Figure 2. Correlation SFPA of surprise ratings and coactivations in Sherlock. Left: brain-maps denoting Hippocampus (yellow), DMN (blue), DAN (red) and Vis (green). Right: Fluctuations in surprise ratings were tested for correlation with fluctuations in ISFC between each pair of regions. Black outlines denote above-chance correlations at $p < 0.05$ (corrected), determined by random permutation testing (1000 iterations). Scatterplot illustrates the correlation between mean surprise and mean ISFC of the right middle frontal gyrus (MFG) and angular gyrus (AG) across movie events.

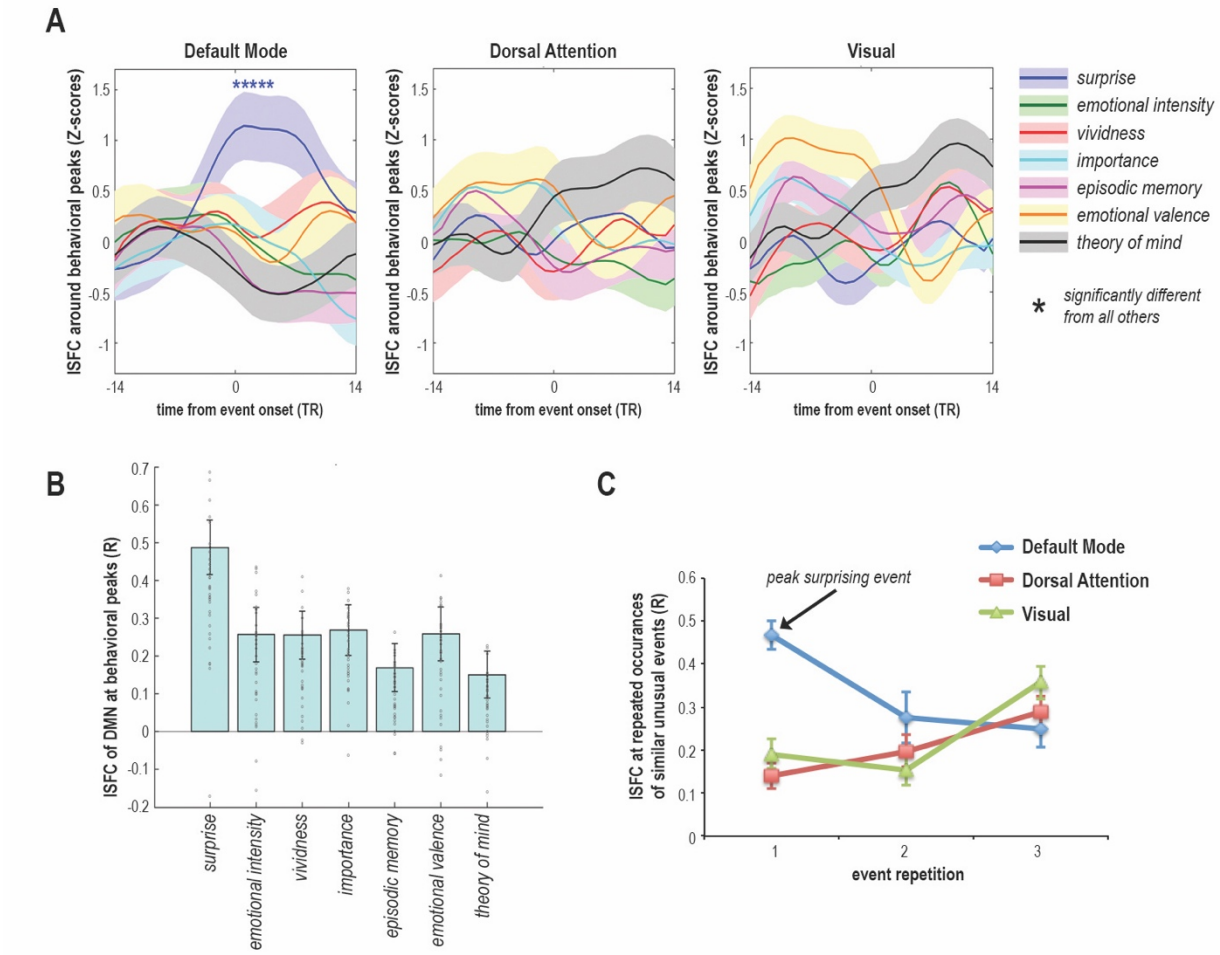

Figure 3. Peak-state SFPA of coactivations in Sherlock. ISFC was averaged across the 5 peak events on each behavioral measure. DMN regions were selectively coactivated during peak surprise, compared to all other peak states, as revealed by random permutation testing (1000 iterations) at $p < 0.05$. Network ISFC is plotted as mean ± SEM across subjects. (A) Peak-state SFPA plotted around the scan volume corresponding to the event (time 0) ± 14 volumes adjacent to the event. (B) DMN ISFC corresponding to the event alone (time 0 in A). Surprise 0.49 ± 0.07, Emotional Intensity 0.26 ± 0.07, Vividness 0.25 ± 0.06, Importance 0.27 ± 0.07, Episodic Memory 0.17 ± 0.06, Emotional Valence 0.26 ± 0.07, Theory of Mind 0.15 ± 0.06. Gray circles denote individual-subject ISFC values. (C) ISFC corresponding to repetitions of a similar event within the same scene, the first of which was found to generate peak surprise. The scene depicted a press conference in which the same initially-surprising text-message was sent to all attendees 3 times, corresponding to a decrease in coactivation of the DMN, but not DAN or Vis.

Independent of the correlation SFPA, peak-state SFPA examined the relationship between DMN coactivation and peak cognitive states, in an event-triggered analysis. To this end, we extracted the mean network ISFC over time-windows corresponding to the 5 highest behaviorally-scored events in the movie, separately for each behavioral measure. DMN coactivation was selectively enhanced during peak surprising events (Figure 3). Particularly,

mean ISFC during surprise peaks was higher than during other cognitive peaks in the DMN (perm. $p < 0.05$), but not DAN or Vis (perm. $p > 0.05$). The 3 networks were significantly different in their ISFC as measured by peak-state SFPA ($F(12) = 43.94$, $p < 0.001$, $\eta_p^2 = 0.56$; see full ANOVA report in Methods), and particularly during peak surprise ($F(2) = 98.94$, $p < 0.001$, $\eta_p^2 = 0.33$). This suggests that surprise ratings are unlikely to reflect a large attentional shift or low-level perceptual processing typical to Vis and DAN (*20*), and thus more likely to reflect a higher-order response to an unexpected occurrence.

Neuro-computational theories of predictive processing describe the brain as a Bayesian inference machine, which optimizes its predictions of future events by calculating the mismatch between expectation and reality, termed prediction error (*21, 22*). If surprising events in the movie triggered a prediction error, exhibited by increased DMN coactivation, then we would expect the prediction error to decrease with repetition of an initially surprising event. Indeed, we see an example for this during a scene in the movie depicting a press conference, in which an initial surprising mass text message is sent to all attendees, triggering a peak surprising event. Within the same scene, the same mass text message is sent twice more. As demonstrated in Figure 3C, mean ISFC of the DMN plummeted during the second occurrence of this event, and remained low during the third, whereas DAN and Vis exhibited different response patterns. This suggests that after processing an unusual event for the first time, the prediction error reflected in DMN coactivation is diminished, consistent with error-driven prediction updating (*21-23*).

To further understand the link to predictive processing in our data, we specifically examined striatal regions, which have been previously shown to respond to unexpected stimuli during trial-by-trial learning tasks, and towards novel contexts in naturalistic stimulation (*13-16*). Indeed, SFPA revealed that DMN coactivation with striatal regions, primarily the Nucleus Accumbens (NAcc), fluctuated proportionally to the state of surprise, as revealed by a significant correlation between surprise and ISFC of NAcc and DMN regions (perm. $p < 0.05$, corrected; Figure 4). Despite this, surprise did not modulate coactivation among striatal regions themselves, nor between striatum and hippocampus (perm. $p > 0.05$, corrected), consistent with these regions' involvement in a wider range of learning and memory functions (*24*). Thus, the unique element connecting surprise to hippocampus and NAcc here is the DMN. This result did not extend to nearby Thalamus, suggesting that surprise-dependent coactivation with DMN is unique to hippocampus and striatum.

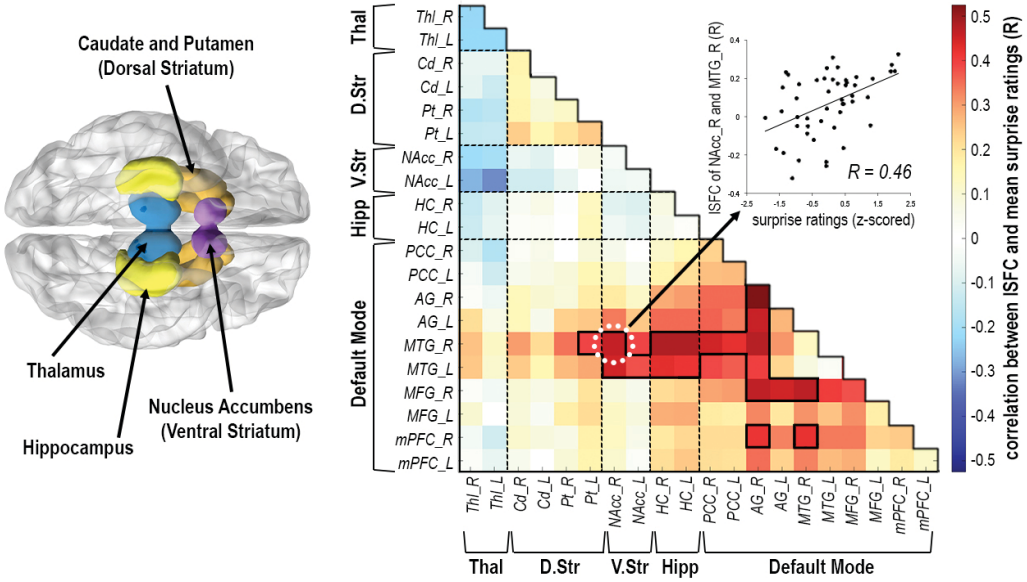

Figure 4. Correlation SFPA of surprise ratings and cortical-subcortical coactivations in Sherlock. Left: brain-map denoting Hippocampus (yellow), NAcc (ventral striatum: purple) Caudate and Putamen (dorsal striatum: Orange), Thalamus (blue). Right: Fluctuations in surprise ratings were tested for correlation with fluctuations in ISFC between each pair of subcortical regions and DMN regions. Black outlines denote above-chance correlations at $p < 0.05$ (corrected), determined by random permutation testing (1000 iterations). Scatterplot illustrates the correlation between mean surprise and mean ISFC of the right middle temporal gyrus (MTG) and NAcc across movie events.

Notably, the current results cannot be explained by overall DMN activation or deactivation during surprising events, as no significant correlations were found between surprise ratings and mean univariate responses of DMN regions (perm. $p > 0.05$, corrected), nor did we find univariate effects during surprise peaks (perm. $p > 0.05$). Whole-brain analysis of surprise-dependent univariate activity ($p < 0.05$, corrected) similarly revealed little to no overlap with DMN voxels (Supplementary Figure 3). Predictive processing is thus reflected in DMN shared patterns of activity fluctuations rather than in a DMN on/off response. In addition, low-level stimulus features did not modulate DMN coactivation, as revealed by correlation of visual saliency and luminance with ISFC in DAN and Vis (perm. $p < 0.05$, corrected), but not DMN (perm. $p > 0.05$, corrected; Supplementary Figure 4). Thus, DMN coactivations are unlikely to reflect low-level sensory processing, and more likely correspond to higher semantic processing of movie-narrative content.

Discussion

Altogether, our findings reveal interactions among DMN regions, hippocampus and NAcc, which are selectively coactivated during the processing of high-level prediction errors. The DMN is central to this process, acting as a hub for surprise-dependent responses of subcortical regions. To better understand these functional interactions, we must first consider the role of surprise in semantic comprehension of unfolding events.

A surprising event forces us to update our internal model of reality, or in the case of a fictional movie - the narrative, to fit contradictive incoming information. This requires, first, an internal model, second, detection of a mismatch between the internal model and incoming information, and third, integration of incoming information with previously acquired information to improve model predictions. For the first prerequisite, the DMN is a suitable candidate to carry an internal model of the narrative, as its regions have been shown to carry information about narrative content (*7-10*), and have been hypothesized to represent event models and contextual schemas (*4, 25, 26*). Our findings offer evidence in support of the second prerequisite, by showing surprise-dependent coactivation of DMN regions and NAcc. Given previous accounts of NAcc in predictive processing (*13-15, 27*), this may point to predictive error detection reflected in reported surprise. The third step towards model updating, i.e. integration across concurrent and past events, requires the process of memory retrieval. Previous findings have linked DMN regions, as well as their coactivation with hippocampus, to memory recall (*10, 28, 29*). Thus, surprise-dependent coactivation of DMN and hippocampus found here may relate to retrieval processes needed for temporal narrative integration.

Finally, our proposed interpretation corresponds with the hypothesis that switching between internal and external based processing modes is necessary for error-driven learning, and involves the DMN and hippocampus (*30*). In this case, surprise may lead to switching between unexpected incoming information (external mode), memory of previous events and our internal model (internal modes), as we integrate across all 3. Thus, coactivation of DMN and subcortical regions may support integration across external information and internal representations, as we experience the mental state of surprise.

Methods

Stimuli

We examined human behavioral responses and functional magnetic resonance imaging (fMRI) responses to two movies. The first movie was a 23-minute excerpt (10) from the first episode of the BBC television series Sherlock (2010). The second movie was an 8-minute edited excerpt (31,32) from Bang! You're Dead, from the television series Alfred Hitchcock Presents (1961).

Behavioral Participants

Forty-five participants (19 female, age 33.2 ± 8.7 years) were included in the behavioral data for the movie Sherlock, and 41 participants (17 female, age 31.3 ± 7.7 years) were included in the behavioral data for the movie Bang! You’re Dead. All participants reported normal or corrected-to-normal vision and hearing, and gave informed consent. Two additional participants for Sherlock, and 3 additional participants for Bang! You’re Dead, were excluded from behavioral-data analysis because they did not complete the task as instructed.

Behavioral Experimental Procedure

We collected behavioral responses to each of the movies using Amazon Mechanical Turk. Experimental procedures were approved by the institutional review board (IRB; approval reference # 533-2) of the Weizmann Institute of Science.

Participants were first screened for technical compatibility (e.g. operating system, internet connection, screen size and sound) and fluent English writing ability, in order to enable successful video viewing and questionnaire completion. In addition, participants ability to properly hear and see the video was tested before beginning the experiment, in a short audiovisual clip followed by auditory and visual catch questions. Participants were instructed to sit at a distance of 1 foot (12 inches) from the screen. Sherlock was presented at 200 mm over 112.5 mm, and Bang! You’re Dead was presented at 180 mm over 135 mm. Participants viewed the movie from start to end without pausing, skipping or rewinding. Single continuous viewing was additionally monitored via recorded viewing times.

We developed a novel method of retrospective behavioral sampling in order to measure the fluctuations in cognitive states throughout the movie experience. After viewing the movie, participants first typed a brief free recall describing the content of the movie. Next, participants completed a questionnaire recording their self-reported experience referring to

various events of the movie. The questionnaire for Sherlock referred to 49 events, sampling the time-course of the movie at intervals of ~30 sec. The questionnaire for Bang! You're Dead referred to 39 events, sampling the movie at intervals of ~15 seconds. Participants were randomly assigned to respond to 1 of 3 subsets of events, chronologically interleaved. Included events were probed in random order throughout the questionnaire. The reminder for each event was presented as a timestamp with a short description of something that happened at a particular moment in the movie (e.g. 10:14 - Sherlock (in lab): "Mike, can I borrow your phone?"). Participants were then asked to focus their memory on that particular event, including no more than a few seconds before and after it. They rated how vividly they remembered the event, typed a detailed free recall of the event, and rated to what extent the event was surprising, emotionally intense, emotionally negative or positive, and important to the plot. All ratings were collected on scales from 1 to 7. Instructions for the free recall of each event resembled the autobiographical interview method (33,34), asking participants to recall every detail they remembered about what happened at that moment of the movie, what they saw and heard, their thoughts, emotions and physical sensations while viewing the event.

Behavioral Data Processing

We extracted measures of episodic memory and theory of mind (TOM) from the open answers of the free recall for each event separately, as follows. Episodic memory was measured as the number of mentions (memory units) of remembered facts about things that happened in the movie during or adjacent to the event (e.g. "Sherlock was wearing a white shirt and stepped into the apartment" would be counted as two memory units). This score excluded the information already given in the reminder for the event, as well as facts that did not match the actual movie content. In addition, TOM was (orthogonally) measured as the number of references to the state of mind of movie characters during said event (e.g. "Sherlock seemed happy about discovering a fourth victim" would be counted as 1 TOM unit).

Episodic memory units, TOM units and each of the behavioral ratings were z-scored (demeaned and divided by standard deviation), within each participant and each behavioral measure separately, across the time-course of responses. Thereafter, responses were averaged across subjects, resulting in a single temporal pattern per behavioral measure, describing the group fluctuation in each cognitive state throughout movie events.

fMRI Data Sources

fMRI data for Sherlock included 17 participants obtained with permission from Chen et al. (10) and 18 participants obtained with permission from Zadbood et al. (8). These data consisted of preprocessed 3-T fMRI T2*-weighted echo-planar imaging (EPI) blood-oxygen-level-dependent (BOLD) responses with whole-brain coverage (TR 1,500 ms), in Montreal Neurological Institute (MNI) standard volume space (8,10). Both datasets included the responses to the target stimulus, i.e. the first half of the episode, consisting of 946 volumes. Additional data (10) contained responses to the second half of the episode (used for ROI localization - see below), and consisted of 1030 volumes.

fMRI data collection and sharing for Bang! You're Dead was provided by the Cambridge Centre for Ageing and Neuroscience (CamCAN; 31,32). From the repository data, we randomly sampled 30 participants within an age range of 20-50 years. These data consisted of 3-T fMRI T2*-weighted EPI raw BOLD responses with whole-brain coverage (31). Movie-scan data consisted of 193 volumes (TR 2,470 ms). In addition, resting-state data (used for ROI localization - see below) of the same CamCAN participants consisted of 261 volumes (TR 1,970 ms).

fMRI Data Processing

fMRI data were analyzed using MATLAB (MathWorks) with statistical parametric mapping (SPM) for preprocessing, NeuroElf for region-of-interest (ROI) organization and BrainNet for ROI visualization.

Preprocessing was performed on raw signals only (CamCAN data) and included slice-timing correction, spatial realignment, transformation to MNI space (voxel size 3 mm x 3 mm x 3 mm), and spatial smoothing with a 6 mm full-width at half-maximum (FWHM) Gaussian kernel. Thereafter, all data underwent voxel-wise detrending and z-scoring (demeaned and divided by standard deviation) across scan volumes.

Functional network and ROI localization was performed in two steps, constraining selection first by response correlation within tested participant sample, and second by previous functional network definitions based on vast samples (35). To measure response correlations within our current sample, we calculated the seed-based functional connectivity during rest and non-target movie scans, independent of the target movie data later used for inter-subject functional correlation (ISFC) analysis. Spherical 80-voxel seeds were defined anatomically in MNI space, based on locations validated in previous reports (36), for the default mode

network (DMN) in the posterior cingulate cortex (PCC: 0, −53, 26), for the dorsal attention network (DAN) in the intraparietal sulcus (IPS: 22, −58, 54), and for visual network (Vis) in the primary visual cortex (V1: 30, −88, 0). Functional connectivity was calculated separately for each participant by correlating the signal time-course within each voxel with the average signal time-course of the seed region. Pearson coefficients within each voxel were averaged across participants, resulting in 3 correlation maps corresponding to the 3 seeds. Voxels with mean correlation values of at least 0.3 were included in the second selection step. In order to maintain comparable network sizes across datasets (DMN, 4418-4768 voxels; DAN 3705-3541 voxels; Vis 4604-4983 voxels), and due to shorter duration (less degrees of freedom) a higher cutoff value of 0.35 was used for CamCAN data. The second selection step utilized a predefined parcellation of 14 functional networks (35), by discarding voxels outside the predefined DMN (a, b, or c), DAN (a or b) and Vis networks from each of the corresponding correlation maps. Finally, remaining voxels were allocated to gross anatomical regions based on the atlas definition of each network. Voxels of the DMN were allocated to the PCC/precuneus, angular gyrus (AG), middle temporal gyrus (MTG), middle frontal gyrus (MFG), and medial prefrontal cortex (mPFC). Voxels of the DAN were allocated to the superior parietal lobe (SPL), postcentral gyrus (PostC), frontal eye field (FEF), occipital temporal cortex (OTC), parietal occipital cortex, and precentral ventral region (PrCv). Voxels of Vis were allocated to visual central areas (VisCent) and visual peripheral areas (VisPeri). In addition, subcortical ROIs hippocampus (HC), nucleus accumbens (NAcc), Caudate (Cd), Putamen (Pt), and thalamus (Thl) were defined anatomically via the automated anatomical labelling atlas (AAL; 37).

To prepare for ISFC analysis, for each participant we extracted the average across voxels within each ROI, along the response time-course of the target-movie scan. We then calculated the average across all other participants (excluding reference participant) across ROI voxels. ISFC between two ROIs was calculated in a sliding time-window of 15 scanning volumes, as the Pearson correlation between the signal time-course of each participant in the first ROI, and the average time-course of all other participants in the second ROI. Correlation values were Fisher-transformed and averaged across participants, resulting in a single mean correlation value per window per ROI pair. This was repeated for every volume (±7 TR in time-window), yielding a single time-course of ISFC values per ROI pair, describing the fluctuation in coactivation among each pair of tested regions throughout movie events.

State-Fluctuation Pattern Analysis (SFPA)

We developed a method of state-fluctuation pattern analysis (SFPA) to examine how cognitive states are dynamically linked to functional network coactivation during continuous naturalistic stimulation. The first component of this method is the novel technique of retrospective behavioral sampling and modeling participants' natural experience into temporal patterns of cognitive states, as described above. The second component of SFPA tests whether dynamic coactivation among brain regions is predicted by each cognitive state. To this end, we present 2 complementary analyses, which examine the correlation across temporal patterns of coactivation and behavior, and the coactivation corresponding to peak cognitive states. As these analyses were performed across the means of independent groups, for behavior and for coactivation, the temporal patterns of one modality serve as independent predictors for the other.

For the correlation SFPA, we first down-sampled the ISFC time-course to match the behavioral time-course, by selecting the ISFC scores centered on each of the behaviorally-tested events in the movie. Thus, each event was assigned a single ISFC score calculated, as described in the previous section, across the 15-TR time-window centered around the behavioral event onset (event TR ±7). Very early or late events, with less than 7 TRs available for ISFC scoring before and after event onset, were discarded, resulting in 49 events for Sherlock, and 36 events for Bang! You're Dead. For each behavioral measure, we then calculated the Pearson correlation between the temporal pattern of cognitive state and the temporal pattern of corresponding ISFC scores, separately for each pair of ROIs. This resulted in a matrix of correlation coefficients as illustrated in Figure 1D. Permutation testing was performed by random shuffling of the time-series of ISFC and correlating again with the cognitive state, repeated 1000 times, thus resulting in the null distribution for significance testing. Significance was determined at $p < 0.05$ (2-tailed) by testing the original correlation value against the permutation distribution. Because permutation testing was repeated per ROI pair, p values were corrected for multiple comparisons using the false detection rate (FDR; 38).

For the peak-state SFPA, we examined the event-triggered ISFC during peak cognitive states. To this end, we first identified the top 5 peaks along the temporal patterns of cognitive states, for each behavioral measure separately. ISFC values for each region pair were z-scored (demeaned and divided by their standard deviation) across the time-course of the movie. We then averaged the ISFC z-scores across all network ROIs, and across the 5 peak events,

within an event window of 29 time-bins centered around the event onset (event TR ± 14). To clarify, the value assigned to each time-bin in the event window is the ISFC score, as calculated across the 15-TR time-window centered around the time-bin TR (for example, time-bin 16 in the event window corresponds to the event TR +1, and the ISFC score assigned to this time bin was calculated between event TR -6 and event TR +8). This resulted in a time-course of mean network ISFC, describing the overall network coactivation corresponding to each type of peak cognitive state. Permutation testing was performed to compare between the ISFC of each cognitive state relative to every other state. This was done by measuring the maximum absolute difference between ISFC mean across 5 randomly-selected events, and ISFC mean across an additional 5 randomly-selected events, repeated 1000 times. Significance at $p < 0.05$ (1-tail) was tested against a single critical threshold of difference, determined by the 95th percentile of the distribution of the maximum differences.

We tested inter-network differences in peak-state SFPA in a repeated-measures ANOVA of ISFC with network (DMN, DAN, Vis) and cognitive state (surprise, emotional intensity, vividness, importance, episodic memory, emotional valence, theory of mind) as within-subject factors. Results revealed a significant main effect of cognitive state ($F(6) = 13.79$, $p < 0.001$), a marginal main effect of network ($F(2) = 2.85$, $p = 0.065$), and a highly significant two-way interaction between network and cognitive state ($F(12) = 43.94$, $p < 0.001$). We thus further tested the inter-network differences in peak-state SFPA specific to surprise, in a repeated-measures ANOVA of ISFC during peak surprise, with network as the only factor, revealing the effect reported in results.

Control I: SFPA of Univariate Activations

To test whether univariate activation or deactivation may explain our results with ISFC, we also performed the SFPA using the mean ROI BOLD time-course in place of the ISFC time-course. The value assigned to each time-bin was the mean BOLD across subjects, in the single TR corresponding to it in time. Notably, similar results were found when assigning to each time-bin the mean of the time window corresponding to the ISFC analysis (15 TR) as well as with a 5-TR window. All other analysis steps were the same as described in the previous section. In addition, we conducted a whole-brain analysis by correlating, for each participant, the voxel-wise BOLD with the behavioral time-course of surprise ratings. Pearson coefficients were Fisher-transformed and voxel-wise (Bonferroni-corrected) T-test was performed to test group effect. Significant results were plotted on a brain map of T-

values, describing the magnitude of correlation between BOLD activation and surprise ratings in each voxel.

Control II: SFPA of Visual Attributes

We tested whether low-level visual features of the movie stimuli were correlated with network coactivation across the same movie events probed in the behavioral experiment. To this end, we extracted the mean levels of visual luminance and spectral saliency from each movie frame, and calculated the average across all movie frames within every time-window corresponding, in temporal range, to the ISFC time-windows. This yielded 2 time-courses, describing the fluctuations in visual saliency and visual luminance throughout movie events. We then performed SFPA as described above, using the luminance and saliency time-courses in place of the behavioral cognitive-state time-courses.

References

1. R. L. Buckner, J. R. Andrews-Hanna, D. L. Schacter, The brain's default network: anatomy, function, and relevance to disease. *Ann N Y Acad Sci* **1124**, 1-38 (2008).
2. M. E. Raichle *et al.*, A default mode of brain function. *Proc Natl Acad Sci U S A* **98**, 676-682 (2001).
3. G. L. Shulman *et al.*, Common Blood Flow Changes across Visual Tasks: II. Decreases in Cerebral Cortex. *J Cogn Neurosci* **9**, 648-663 (1997).
4. R. L. Buckner, L. M. DiNicola, The brain's default network: updated anatomy, physiology and evolving insights. *Nat Rev Neurosci* **20**, 593-608 (2019).
5. M. E. Raichle, The brain's default mode network. *Annu Rev Neurosci* **38**, 433-447 (2015).
6. M. D. Rugg, K. L. Vilberg, Brain networks underlying episodic memory retrieval. *Curr Opin Neurobiol* **23**, 255-260 (2013).
7. C. Baldassano, U. Hasson, K. A. Norman, Representation of Real-World Event Schemas during Narrative Perception. *J Neurosci* **38**, 9689-9699 (2018).
8. A. Zadbood, J. Chen, Y. C. Leong, K. A. Norman, U. Hasson, How We Transmit Memories to Other Brains: Constructing Shared Neural Representations Via Communication. *Cereb Cortex* **27**, 4988-5000 (2017).
9. E. Simony *et al.*, Dynamic reconfiguration of the default mode network during narrative comprehension. *Nat Commun* **7**, 12141 (2016).

10. J. Chen *et al.*, Shared memories reveal shared structure in neural activity across individuals. *Nat Neurosci* **20**, 115-125 (2017).
11. Y. Lerner, C. J. Honey, L. J. Silbert, U. Hasson, Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. *J Neurosci* **31**, 2906-2915 (2011).
12. U. Hasson, Y. Nir, I. Levy, G. Fuhrmann, R. Malach, Intersubject synchronization of cortical activity during natural vision. *Science* **303**, 1634-1640 (2004).
13. A. Pine, N. Sadeh, A. Ben-Yakov, Y. Dudai, A. Mendelsohn, Knowledge acquisition is governed by striatal prediction errors. *Nat Commun* **9**, 1673 (2018).
14. N. Axmacher *et al.*, Intracranial EEG correlates of expectancy and memory formation in the human hippocampus and nucleus accumbens. *Neuron* **65**, 541-549 (2010).
15. P. F. Rodriguez, A. R. Aron, R. A. Poldrack, Ventral-striatal/nucleus-accumbens sensitivity to prediction errors during classification learning. *Hum Brain Mapp* **27**, 306-313 (2006).
16. A. Ben-Yakov, Y. Dudai, Constructing realistic engrams: poststimulus activity of hippocampus and dorsal striatum predicts subsequent episodic memory. *J Neurosci* **31**, 9032-9042 (2011).
17. M. D. Greicius, B. Krasnow, A. L. Reiss, V. Menon, Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proc Natl Acad Sci U S A* **100**, 253-258 (2003).
18. M. W. Cole, D. S. Bassett, J. D. Power, T. S. Braver, S. E. Petersen, Intrinsic and task-evoked network architectures of the human brain. *Neuron* **83**, 238-251 (2014).
19. F. M. Krienen, B. T. Yeo, R. L. Buckner, Reconfigurable task-dependent functional coupling modes cluster around a core functional architecture. *Philos Trans R Soc Lond B Biol Sci* **369**, (2014).
20. M. Corbetta, G. L. Shulman, Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci* **3**, 201-215 (2002).
21. K. Friston, A theory of cortical responses. *Philos Trans R Soc Lond B Biol Sci* **360**, 815-836 (2005).
22. A. Clark, Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav Brain Sci* **36**, 181-204 (2013).
23. R. A. Rescorla, A. R. Wagner, in *Classical conditioning II: current research and theory,* A. H. Black, W. F. Prokasy, Eds. (Appleton-Century-Crofts, New York, 1972), pp. 64-99.

24. C. M. Pennartz, R. Ito, P. F. Verschure, F. P. Battaglia, T. W. Robbins, The hippocampal-striatal axis in learning, prediction and goal-directed behavior. *Trends Neurosci* **34**, 548-559 (2011).
25. D. Stawarczyk, M. A. Bezdek, J. M. Zacks, Event Representations and Predictive Processing: The Role of the Midline Default Network Core. *Top Cogn Sci*, (2019).
26. D. Hassabis, E. A. Maguire, Deconstructing episodic memory with construction. *Trends Cogn Sci* **11**, 299-306 (2007).
27. S. Ikemoto, Dopamine reward circuitry: two projection systems from the ventral midbrain to the nucleus accumbens-olfactory tubercle complex. *Brain Res Rev* **56**, 27-78 (2007).
28. A. D. Wagner, B. J. Shannon, I. Kahn, R. L. Buckner, Parietal lobe contributions to episodic memory retrieval. *Trends Cogn Sci* **9**, 445-453 (2005).
29. W. Huijbers, C. M. Pennartz, R. Cabeza, S. M. Daselaar, The hippocampus is coupled with the default network during memory retrieval but not during memory encoding. *PLoS One* **6**, e17463 (2011).
30. C. J. Honey, E. L. Newman, A. C. Schapiro, Switching between internal and external modes: A multiscale learning principle. *Netw Neurosci* **1**, 339-356 (2018).
31. M. A. Shafto *et al.*, The Cambridge Centre for Ageing and Neuroscience (Cam-CAN) study protocol: a cross-sectional, lifespan, multidisciplinary examination of healthy cognitive ageing. *BMC Neurol* **14**, 204 (2014).
32. J. R. Taylor *et al.*, The Cambridge Centre for Ageing and Neuroscience (Cam-CAN) data repository: Structural and functional MRI, MEG, and cognitive data from a cross-sectional adult lifespan sample. *Neuroimage* **144**, 262-269 (2017).
33. D. Berntsen, Tunnel memories for autobiographical events: central details are remembered more frequently from shocking than from happy experiences. *Mem Cognit* **30**, 1010-1020 (2002).
34. B. Levine, E. Svoboda, J. F. Hay, G. Winocur, M. Moscovitch, Aging and autobiographical memory: dissociating episodic from semantic retrieval. *Psychol Aging* **17**, 677-689 (2002).
35. A. Schaefer *et al.*, Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cereb Cortex* **28**, 3095-3114 (2018).
36. K. R. Van Dijk *et al.*, Intrinsic functional connectivity as a tool for human connectomics: theory, properties, and optimization. *J Neurophysiol* **103**, 297-321 (2010).

37. E. T. Rolls, C. C. Huang, C. P. Lin, J. Feng, M. Joliot, Automated anatomical labelling atlas 3. *Neuroimage* **206**, 116189 (2020).
38. Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 289-300 (1995).

Acknowledgments: We thank Avigail Mirsky for data curation contribution, Chen et al. (*10*) and Zadbood et al. (*8*) for resource contributions, and Aya Ben-Yakov, Talya Sadeh, Michal Bernstein and Galit Yovel for their useful advice. This work was supported by the Israel Science Foundation grant 1458/17 to ES. CamCAN funding was provided by the UK Biotechnology and Biological Sciences Research Council (grant number BB/H008217/1), together with support from the UK Medical Research Council and University of Cambridge, UK.

Data and materials availability: All materials, code and data directly collected for this study will be made freely available upon request.

Competing interests: Authors declare no competing interests.