Contributions of episodic retrieval and mentalizing to autobiographical thought: Evidence from functional neuroimaging, resting-state connectivity, and fMRI meta-analyses

Jessica R. Andrews-Hanna, Rebecca Saxe, Tal Yarkoni

Abstract

A growing number of studies suggest the brain’s “default network” becomes engaged when individuals recall their personal past or simulate their future. Recent reports of heterogeneity within the network raise the possibility that these autobiographical processes comprised of multiple component processes, each supported by distinct functional–anatomic subsystems. We previously hypothesized that a medial temporal subsystem contributes to autobiographical memory and future thought by enabling individuals to retrieve prior information and bind this information into a mental scene. Conversely, a dorsal medial subsystem was proposed to support social-reflective aspects of autobiographical thought, allowing individuals to reflect on the mental states of one’s self and others (i.e., “mentalizing”). To test these hypotheses, we first examined activity in the default network subsystems as participants performed two commonly employed tasks of episodic retrieval and mentalizing. In a subset of participants, relationships among task-evoked regions were examined at rest, in the absence of an overt task. Finally, large-scale fMRI meta-analyses were conducted to identify brain regions that most strongly predicted the presence of episodic retrieval and mentalizing, and these results were compared to meta-analyses of autobiographical tasks. Across studies, laboratory-based episodic retrieval tasks were preferentially linked to the medial temporal subsystem, while mentalizing tasks were preferentially linked to the dorsal medial subsystem. In turn, autobiographical tasks engaged aspects of both subsystems. These results suggest the default network is a heterogeneous brain system whose subsystems support distinct component processes of autobiographical thought.

Introduction

Though merely a decade has elapsed since the “default network” (DN) was introduced as a large-scale brain system (Greicius et al., 2003; Raichle et al., 2001), its widespread engagement and clinical relevance has sparked a surge of interest concerning the network’s adaptive functions. Important insight comes from studies demonstrating overlapping patterns of DN activity across a wide range of autobiographical and social tasks, supporting the notion of a “core network” (e.g., Buckner and Carroll, 2007; Rabin et al., 2010; Schacter et al., 2007, 2012; Spreng and Grady, 2010; Spreng et al., 2009). These findings prompted a recent wave of debates surrounding an overarching function or specific mechanism suitable to accommodate shared activity patterns across varied tasks. While some accounts hold that a broad function of the DN is to support internally-directed or perceptually-decoupled processes (Andrews-Hanna, 2012; Buckner et al., 2008; Smallwood et al., 2012), other accounts emphasize its role in mental simulation or self-projection across time, space, and body (Buckner and Carroll, 2007; Spreng et al., 2009), mental construction of novel or familiar scenes (Hassabis and Maguire, 2007, 2009; Schacter et al., 2007), associative prediction (Bar, 2007, 2009), mentalizing (Mars et al., 2012; Mitchell, 2009; Schilbach et al., 2008, 2012), semantic and conceptual retrieval (Binder et al., 2009), and so on.

One way to reconcile these seemingly-conflicting views is to examine the DN on a finer scale, considering the possibility that the network comprises functionally-distinct components that co-activate or interact during more complex forms of cognition (Andrews-Hanna, 2012; D’Argembeau et al., in press; Kim, 2012; Schacter et al., 2012; Szpunar et al., in press). Under this view, periods of passive rest may encourage participants to default to a variety of mental processes (i.e., mnemonic, self-referential, social), each of which may be supported by distinct DN components. However, since these introspective processes often co-occur, the DN components will activate together, creating the perception of a functionally homogeneous network.

These principles might also extend to complex experimentally-directed tasks that activate the DN. Tasks requiring individuals to recall and reflect on specific, personal past experiences
autobiographical memory), or imagine future events for which individuals might experience (termed autobiographical future thought), might invoke multiple component processes including 1) retrieval of episodic elements and contextual details contributing to the autobiographical experience, 2) integration of such elements with conceptual knowledge about the self, and 3) meta-cognitive reflection on the feelings, emotions, and/or beliefs of one’s self or other people involved in the experience (termed mentalizing) (Cabeza et al., 2004; Conway, 2009; Prebble et al., 2013; Rubin, 2006; Schacter et al., 2012). In contrast to autobiographical tasks, laboratory-based tasks of episodic retrieval and mentalizing isolate a subset of these component processes, and the neural overlap of these more specific tasks may be largely distinct.

Support for the heterogeneity account of the DN comes from recent findings suggesting that the DN fractionates into at least two anatomically and functionally dissociable subsystems, including a dorsal medial subsystem and a medial temporal subsystem, both converging on a midline core (Fig. 1; Andrews-Hanna et al., 2010a; Yeo et al., 2011). In addition to the involvement of the dorsal medial prefrontal cortex (dMPFC), other key regions within the dorsal medial subsystem include the temporoparietal junction (TPJ) and lateral temporal cortex, with more recent whole-brain parcellations including inferior and superior frontal gyrus (Yeo et al., 2011). In contrast, the medial temporal subsystem comprises the hippocampal formation (HF), parahippocampal cortex (PHC), retrosplenial cortex (Rsp), and posterior inferior parietal lobule (pIPL). Based on a broad literature review, we recently hypothesized that the medial temporal subsystem allows individuals to mentally simulate events by retrieving episodic information and binding that information together in a spatially-coherent manner, while the dorsal medial subsystem plays more of a social-reflective role, allowing individuals to infer the mental states of other people and reflect on their own mental states (Andrews-Hanna, 2012; Andrews-Hanna et al., in press; Buckner et al., 2008). With these ideas in mind, the goals of the present manuscript were three-fold. First, we sought to compare the neural underpinnings of mentalizing to those of episodic recollection, testing the prediction that these processes differentially engage the dorsal medial and medial temporal subsystems. Next, we sought to examine relationships among task-evoked regions in the absence of an overt task. Finally, we sought to test the hypothesis that more complex tasks such as autobiographical memory and autobiographical future thought engage aspects of both subsystems. Across three studies, we tested our predictions using a multi-method approach including task-related fMRI, resting-state functional connectivity MRI (rs-fcMRI), and automated fMRI meta-analyses.

Material and methods

Overview

The role of the default network components in episodic retrieval, mentalizing, and autobiographical memory/future thought were explored in a three-part experiment. In Part 1, participants performed laboratory-based episodic memory and mentalizing tasks while scanned with high-resolution fMRI. In Part 2, we examined whether patterns of task-related activity were reflected in the brain’s functional architecture at “rest” by comparing rs-fcMRI with task-defined regions of interest in a subset of participants from Part 1. We next sought correspondence between our findings and the wider neuroimaging literature in Part 3 by employing a novel automated meta-analysis approach (Yarkoni et al., 2011) to identify brain regions that most strongly predicted the presence of episodic retrieval and mentalizing across an extensive database of 5809 published neuroimaging articles. A meta-analysis of autobiographical tasks was also conducted to test the hypothesis that these processes are jointly engaged during autobiographical memory and prospect.

Participants

Neuroimaging and behavioral data was available for 43 participants in Part 1, though eight participants were excluded from subsequent analysis due to excessive motion, weak signal-to-noise, scanner artifacts or poor behavioral performance. Thus, 35 participants (mean age = 21.8 years; age range = 19–28; 13 males) contributed useable imaging and behavioral data for Part 1. Resting state data during a separate resting task was acquired in 33 of these participants (mean age = 21.8 years; age range = 19–28; 12 males), thus contributing to Part 2. All participants were right handed as measured by the Edinburgh Handedness Inventory, and reported no history of psychiatric or neurological conditions, nor use of psychoactive medications. Procedures were carried out according to the Partners Health Care Institutional Review Board, and participants were reimbursed for participation.

Matlab-compatible Psychophysics Toolbox software (Brainard, 1997) was used to program each task paradigm, and stimuli were projected onto a computer screen, viewed through a MRI-compatible mirror placed on top of the head coil. Participants used a button box placed in their right hand to relay their responses. All participants wore plastic glasses with either neutral or corrective lenses and were given earplugs to dampen scanner noise.

MRI data acquisition and preprocessing

Scanning for Parts 1 and 2 was performed in a single session on a 3T Siemens Tim Trio system (Siemens, Erlangen, Germany) using a 12-channel phased-array head coil. High resolution, 3-D, T1-weighted anatomical images were collected using a magnetization prepared rapid acquisition gradient echo (MPRAGE) sequence with the following parameters: repetition time (TR) = 2530 ms, echo time (TE) = 3.44 ms, inversion time (TI) = 1100 ms, flip angle = 7°, field of view (FOV) = 256 × 256 mm, matrix size = 256 × 256, 1 × 1 × 1.133 mm voxel size, and 256 slices. Sequence parameters for functional data varied across studies.

Part 1 employed high-resolution whole-brain asymmetric spin-echo blood-oxygenated level dependent (BOLD) sequences with the following scanning parameters: TR = 5000 ms (trial onset jittered with respect to the TR onset), TE = 30 ms, flip angle = 90°, FOV = 288 mm × 288 mm, matrix size = 144 × 144, 2 mm³ voxel size, 55 axial oblique slices aligned to the anterior commissure/posterior commissure plane. Resting state runs from 33 participants in Part 2 who also completed Part 1 utilized the following BOLD sequence parameters: TR = 2500 ms, TE = 30 ms, flip angle = 90°, FOV = 288 mm × 288 mm, matrix size = 96 × 96, 3 mm³ voxel size, 36 axial oblique slices aligned to the anterior commissure/posterior commissure plane, 156 acquisitions.

Functional data in Parts 1 and 2 underwent a series of common preprocessing steps carried out using FSL (FMRIB, Oxford, UK) and SPM2 (Wellcome Trust Center for Neuroimaging, London, UK) tools. The first three frames of each run were discarded to allow for image intensity stabilization (FLS), followed by skull stripping using the Brain Extraction Tool (FSL), slice timing correction (SPM), motion correction within and between runs using a rigid-body motion correction algorithm (MCFLIRT/FSL), co-registration of the functional images to the structural MP-RAGE (SPM), nonlinear atlas registration to a T2*-weighted MNI template (re-sampled to 2-mm cubic voxels; SPM), and spatial smoothing using either a 4 mm (Part 1) or a 6 mm (Part 2) full-width half-maximum Gaussian kernel (SPM).

For exploratory analyses across Parts 1 and 2, we corrected for multiple comparisons using Monte-Carlo simulations computed with the AFNI programs 3DFWHMx and 3dCLUSTSIM, by using a voxel-wise threshold of \( p < 0.0001 \), combined with an alpha \(< 0.05\) cluster-extent threshold (Ward, 2000). Voxel and cluster thresholds were determined to minimize the frequency of false positives. Results were projected onto surface projections using Caret software (Van Essen, 2005).
Part 1: Task-related fMRI of episodic memory and mentalizing

Task paradigms

Participants were scanned as they performed two separate laboratory-based tasks of episodic memory retrieval and mentalizing: a “remember/know” recognition task (Tulving, 1985), and a “false belief” task assessing the presence of a “theory of mind” (Saxe and Kanwisher, 2003; Zaitchik, 1990). While an alternative approach would have been to design a single task with multiple experimental control conditions, our choice for using separate tasks was motivated by several reasons. First, the use of two commonly-employed tasks of episodic memory and mentalizing allowed us to directly compare our results to previous studies employing similar tasks. Second, disparate tasks encourage distinct retrieval-related and mentalizing modes of cognition. Since both episodic retrieval and mentalizing are thought to comprise automatic, bottom-up processes that may be difficult to inhibit (Cabeza et al., 2008; Lieberman, 2007), participants may have found it difficult to ignore the task-irrelevant process if scanned while performing a single task with multiple conditions. Third, each experimental condition of interest was compared to carefully-matched control conditions, ensuring that differences between tasks isolated episodic memory and mentalizing rather than other factors that differed between paradigms, such as trial length, language comprehension, and reasoning demands. Also note that while the findings in Part 1 might be influenced by the precise nature of the tasks chosen (i.e. a first order False Belief task as opposed to a more complex mentalizing task), we find our approach particularly powerful because we extend the results of Part 1 to patterns of brain connectivity in the absence of an overt task (Part 2), as well as to the wider neuroimaging literature for which a variety of task paradigms are used (Part 3). Converging findings across these analyses would provide strong support for the presence of functional–anatomic brain systems that play fundamental roles in episodic retrieval and mentalizing.

In a behavioral session approximately 20 min prior to the fMRI study, participants completed a deep encoding semantic classification task structured to promote recollection-rich recognition. Participants viewed 80 pictures of living and non-living objects, each with a single word description below the picture. For each object, participants judged its living/non-living status and used a Likert scale to rate how much they “liked” the object. Picture-word pairs were presented twice throughout the encoding task. Participants were then scanned while performing an episodic retrieval task and a separate mentalizing (“theory of mind”) task. The order of the tasks was counterbalanced across participants.

In the retrieval task, two rapid event-related fMRI recognition runs were acquired while participants completed a recognition paradigm based on the remember/know procedure (Tulving, 1985). Instructions were adapted from Rajaram (1993) and Wheeler and Buckner (2004), with an emphasis on differentiating the episodic process of recollection from the arguably non-episodic process of familiarity (see Rugg and Yonelinas, 2003; Vilberg and Rugg, 2008; Yonelinas, 2002 for a discussion of the dual processes account of recognition memory). Specifically, participants were presented with 80 words previously viewed in the encoding session, as well as 40 novel words. Each word appeared for 3 s, followed by a jittered inter-trial interval (ITI) during which participants were instructed to stare at a fixation crosshair. The ITI ranged from 1 to 12 s as determined by the optseq2 optimization algorithm (Dale, 1999) such that fixation trials comprised 1/3 of the task. For each word, participants made a one-step decision selecting a “remember” response if they recalled the word itself as well as additional details about their prior encounter with the word (i.e. the picture, their feelings, their responses on the rating task, etc.), a “know” response if they only had a sense of familiarity that the word was previously-encountered, a “new” response if the word was judged to be novel, or a “guess” response if they were guessing.

To assess the neural underpinnings of mentalizing, participants completed the “False Belief/False Photograph paradigm” across three task runs (Saxe and Kanwisher, 2003; Zaitchik, 1990). The “False Belief” (FB) condition required participants read short stories outlining a series of events in which a character holds a belief conflicting with reality. The participant was asked to infer the false mental state of the character. The “False Photograph” (FP) condition required participants read short stories outlining a series of events in which a photograph or a map contains a record of the location or content of an inanimate object, which has since become outdated. FP stories serve as useful control stimuli because they are structurally identical to the false belief stories, ensuring the conditions are matched on visual input, motor responses, narrative comprehension, and executive function/reasoning demands. The critical difference between the conditions is that only the FB stimuli require participants to infer the mental states of other people. Examples of the two types of stimuli are shown below:

False Belief: Jenny put her chocolate in the cupboard. Then she went outside. While Jenny was outside, Alan moved her chocolate into the fridge. When Jenny returns, she will look for her chocolate in the: Fridge/Cupboard

False Photograph: A photograph was taken of an apple hanging on a branch. The film took 30 minutes to develop. In the meantime, a strong wind blew the apple from the branch to the ground. When the film is developed, it shows the apple on the: Branch/Ground

Nineteen stimuli were selected from Saxe and Kanwisher (2003), and 17 additional stimuli were created for the purposes of the present study, yielding a total of 18 FB and 18 FP stimuli. Story types did not significantly differ with respect to mean reading time (FB = 7.37 s, FP = 7.77 s, independent samples t-test: t(46) = −1.42, p = 0.16) and number of words (Story: FB = 31.5, FP = 30.0, independent samples t-test: t(46) = 1.49, p = 0.14; Question: FB = 10.8; FP = 10.5; independent samples t-test: t(46) = 0.44, p = 0.67). Participants were given 10 s to read each story, followed by a 5 s question period. To encourage participants to closely attend to the stories, half of the questions required participants infer the false belief of the protagonist (or the outdated record or content), while the remaining questions asked about the present reality of the situation. Fourteen seconds of fixation followed each trial.

Statistical analyses

SPM2 (Wellcome Trust Center for Neuroimaging, London, UK) was used to implement the general linear model for fMRI analysis. In addition to task-related regressors (described below), each model included separate regressors to capture the mean intensity within each task run, the mean intensity across all task runs, and linear trends within each run.

For the episodic retrieval task, separate task regressors modeled trials in which participants submitted a correct “Remember” response (i.e. “Hits–Remember”), a correct “Know” response (i.e. “Hits–Know”), and trials in which participants submitted a correct “New” response (i.e. “Correct Rejections”), concatenated across task runs. A single additional task regressor that was not subsequently analyzed modeled trials in which participants submitted infrequent “Guess” responses, incorrect “Remember” or “Know” responses (i.e. “False Alarms”), and incorrect “New” responses (i.e. “Misses”). For each task regressor, a canonical hemodynamic response function was convolved at the onset of each trial. At the single-subject level, parameter estimate images were calculated for each task regressor, and the contrast parameter estimates for remembered Hits versus Correct Rejections were extracted and carried to the next level for random-effects group analysis (one-sample t-test). Note that the contrast between “Hits–Remember” and “Hits–Know” was not permitted since there were too few “Know” trials for reliable analysis (Inline Supplementary Table S1).
For the mentalizing task, a separate model was created, comprising a regressor for FB trials and a second regressor for FP trials, concatenated across task runs. For each task regressor, a canonical hemodynamic response function was convolved with a 15 s boxcar function at the onset of each trial. At the single-subject level, parameter estimate images were calculated for each task regressor, and the contrast between the FB and FP parameter estimates was carried to the group level for random-effects analysis (one-sample t-test). Whole brain contrasts and voxel- and cluster-wise corrections for multiple comparisons were implemented as described above.

Part 2: Resting-state functional connectivity of task-evoked regions

The objective of Part 2 was to compare patterns of task-evoked activity with patterns of whole-brain connectivity in the absence of an overt task (i.e. during awake “rest”). All participants who completed Part 2 also completed Part 1 in the same experimental session. Examining patterns of rs-fcMRI in this same group of participants – as opposed to in an independent sample of participants – was a particular advantage of our study because it enabled us to explore the functional connectivity profiles of goal-defined seed regions created from the tasks in Part 1.

rs-fcMRI preprocessing and correlation analysis

In addition to standard fMRI preprocessing steps described earlier, a series of additional preprocessing steps were implemented to prepare the data for rs-fcMRI analysis (Van Dijk et al., 2010). Data were temporally filtered to retain frequencies between 0.08Hz and 0.08Hz, and a series of nuisance regressors designed to capture motion, respiratory, and cardiac noise were modeled as covariates of non-interest. These included the 6 translation/rotation motion parameters along with their temporal derivatives, and the timeseries and temporal derivatives averaged across regions of interest in the deep white matter, ventricles, and across the whole brain. Since our primary analyses for Part 2 were within-subject comparisons between different ROIs and we did not attempt to interpret the biological basis of negative correlations, we chose to remove the whole brain signal. Individual subject, seed-based correlation procedures were performed on the residual data by cross-correlating the averaged timeseries within an a priori ROI (seed) with the timeseries of each voxel within the entire brain. The resulting individual correlation maps were Fisher z-transformed to obtain normal distributions, and random-effects analyses were performed on the z-transformed maps to compute group results using the same thresholding procedures described above. In addition to generating whole-brain maps reflecting functional correlations with each task-defined seed, we also explored which regions exhibited significant differences in rs-fcMRI between the two seeds by performing whole-brain paired t-tests.

Task-defined seed regions were created by extracting nearby, non-overlapping clusters of activity within the default network that were differentially recruited during the episodic retrieval and mentalizing tasks in Part 1. We select regions that were nearby in spatial location to permit exploration of differential patterns of connectivity across the rest of the brain in an unbiased manner, particularly since the two seeds exhibit similar signal-to-noise properties. The left parietal cortex satisfied these criteria; episodic memory and mentalizing differentially activated portions of the left lateral cortex in a manner corresponding to the pPL and TPJ seeds defined in our previous study (Andrews-Hanna et al., 2010a). Seed regions were defined from the episodic memory and mentalizing activation maps by selecting significant activation clusters in the left parietal cortex for each task separately and retaining voxels that did not overlap between the two tasks. In order to ensure that these seeds regions fell solely within the default network, we masked these clusters with a default network mask identified by Yeo et al. (2011), which clustered patterns of rs-fcMRI from 1000 participants into seven cortical networks.

Part 3: Large-scale automated meta-analyses of episodic retrieval, mentalizing, and autobiographical thought

While Parts 1 and 2 provide strong evidence for different contributions of the DN subsystems in episodic retrieval and mentalizing, these findings may have been influenced by the nature of the task paradigms selected to assess these processes. As a stronger test of the heterogeneity account of the default network, we next used large-scale meta-analyses to examine the neural underpinnings of episodic retrieval and mentalizing across the wider neuroimaging literature. Another objective of Part 3 was to examine episodic retrieval and mentalizing in relationship to autobiographical tasks, including tasks of autobiographical memory and autobiographical future thought. If autobiographical thought is a complex process comprising retrieval-related and mentalizing component processes, these tasks should be predicted by patterns of brain activity that overlap with that elicited by episodic retrieval and mentalizing.

To explore these objectives, we used the Neurosynth framework (Yarkoni et al., 2011; www.neurosynth.org) to perform an automated meta-analysis of fMRI studies exhibiting high-frequency usage of a priori terms reflecting processes similar to episodic memory, mentalizing, and autobiographical thought. As the database contains the largest corpus of fMRI activations extracted to date (nearly 200,000 foci from 5809 studies), it provides a unique opportunity to quantify both the consistency with which the use of a term implies activation in a particular region (i.e. P(Task|Activation)), as well as the specificity with which the presence of activation in a given region provides information about the nature of that study (P(Activation|Task)). This distinction is analogous to the distinction between “forward” and “reverse” inference (Poldrack, 2006; for further discussion, see Yarkoni et al., 2011). Here we focus on patterns of reverse inference (P(Task|Activation)) because they reveal neural systems that discriminate cognitive functions rather than being non-specifically invoked by many different functions (i.e. language, attention, etc.).

Meta-analytic feature search

We performed separate meta-analyses for episodic retrieval, mentalizing, and autobiographical thought by aggregating coordinates linked to a series of a priori search terms associated with each mental state at a frequency of 1 positive term for every 1000 words of text. For the episodic retrieval meta-analysis, we searched the Neurosynth database for coordinates linked to terms indicative of retrieval based on episodic, rather than non-episodic, processes: “remember,” “recollected,” “recollection,” and “recollective.” As these search terms could also identify tasks of autobiographical memory and future thought (which were analyzed in a subsequent meta-analysis), we excluded studies linked to the terms and/or phrases: “autobiographical,” “autobiographical memory,” and “autobiographical recall.” This meta-analysis resulted in a total of 95 studies and 5714 coordinate peaks that meet the above criteria. For the meta-analysis of mentalizing tasks, we searched the database for coordinates linked either to the phrase “theory of mind” or the term “mentalizing,” isolating 79 studies with 2825 peaks. Finally, a meta-analysis of autobiographical tasks used the search terms “autobiographical,” “autobiographical memory,” and “autobiographical recall,” and excluded studies linked to the terms “remember,” “recollect,” “recollected,” “recollection,” and “recollective.” This analysis isolated 62 studies with 2849 peaks.

Statistical analyses

As described in more detail in Yarkoni et al., 2011, whole-brain binary activation images were generated for each published manuscript containing the a priori search term(s) at high frequency by setting the intensity of a voxel to “1” if it fell within 4 mm of an fMRI coordinate reported in the manuscript, or “0” otherwise. Statistical analyses were calculated by aggregating the articles using \( \chi^2 \) tests of statistical independence, testing the dependency between the presence of the search term(s) and the
Results

Part 1: Task-related fMRI of episodic retrieval and mentalizing reveal patterns of functional-dissociation corresponding to DN subsystems

Behavioral results

Behavioral results from the episodic retrieval task indicate a high proportion of recollection-based recognition: participants exhibited a predominance of “Remember” responses as compared to “Know” responses (Inline Supplementary Table S1), with near-ceiling corrected recognition rates (%Hits: 0.97 ± 0.01; %FA = 0.08 ± 0.02; %Hits − %FA = 0.89 ± 0.08). Response time varied across trials, with “Hits-Remember” trials being significantly faster than “Correct Rejections” (paired t-test: t(33) = −7.04; p = 0.000). Accuracy on the mentalizing task was also high (FB: 91 ± 0.02; FP: 0.89 ± 0.02), and did not significantly differ between conditions (paired t-test: t(33) = 1.06; p = 0.30) (Inline Supplementary Table S1). However, response time was significantly faster for FB trials (3121 ms ± 102 ms) compared to FP trials (3287 ms ± 100 ms; paired-t-test: t(33) = −3.81; p = 0.001).

Inline Supplementary Table S1 can be found online at http://dx.doi.org/10.1016/j.neuroimage.2014.01.032.

Neuroimaging results

To test the heterogeneity account of the DN and the hypothesis that disparate DN components support processes related to episodic retrieval and mentalizing, participants performed separate episodic memory and mentalizing tasks while scanned with fMRI. The episodic memory task activated several regions overlapping with the medial temporal subsystem, including the left pIPL, Rsp, and PHG (Fig. 1A; Inline Supplementary Table S2). The bilateral hippocampal formation was also engaged at a lower threshold that did not survive voxel-wise and cluster-extent thresholds. In addition, the PCC core was engaged, as well as additional regions outside the medial temporal subsystem, including the left lateral frontal pole and the left anterior parietal lobule — both components of a frontoparietal control network (Vincent et al., 2008).

Inline Supplementary Table S2 can be found online at http://dx.doi.org/10.1016/j.neuroimage.2014.01.032.

In contrast, the mentalizing task elicited BOLD activity in a number of regions within the dorsal medial subsystem, including bilateral TPJ, lateral temporal cortex extending to the temporal poles, and dMPFC (Fig. 1B; Inline Supplementary Table S2). The bilateral PCC was also activated, as well as the bilateral vMPFC and anterior temporal cortex/amygdala, regions that comprise a “limbic” network (Yeo et al., 2011).

Part 2: Patterns of task-related dissociation are reflected in the brain’s resting architecture

Results from Part 1 demonstrate patterns of task-related dissociation consistent with a functionally-heterogeneous DN whereby the medial temporal subsystem becomes preferentially engaged when retrieving episodic information and the dorsal medial subsystem becomes preferentially engaged when inferring the mental states of other people. To examine whether the regions recruited during these mental states are organized into stable functional-anatomical brain networks, we measured resting state activity in the same group of participants and used rs-fcMRI to examine the architecture of task-evoked regions in the absence of an overt task. Both the left pIPL defined from the episodic retrieval contrast, and the left TPJ defined from the mentalizing contrast exhibited positive functional correlations with a network of medial and lateral regions throughout the DN (Figs. 2A and B). However, when directly contrasting the two rs-fcMRI maps at the within-subject level using paired t-tests, we observed patterns of divergence and convergence consistent with Part 1. The pIPL was significantly more correlated with regions comprising the medial temporal subsystem that were preferentially activated during the episodic retrieval task, while the TPJ was significantly more correlated with regions comprising the dorsal medial subsystem that were preferentially engaged during the mentalizing task (Fig. 2C; Inline Supplementary Table S3). In contrast, the PCC and dMPFC – which were similarly engaged across tasks – were correlated with the task-defined parietal regions to an equivalent degree.

Inline Supplementary Table S3 can be found online at http://dx.doi.org/10.1016/j.neuroimage.2014.01.032.

Part 3: Large-scale, automated meta-analyses of episodic retrieval, mentalizing, and autobiographical thought

Parts 1 and 2 sought to characterize the functional-anatomic profiles of distinct default network subsystems using task-related and resting-state fMRI in a single group of participants. To examine whether these empirical findings extend to a variety of additional task paradigms used across the wider literature, we next used the Neurosynth framework (Yarkoni et al., 2011) to conduct a quantitative, automated text-based meta-analysis of episodic retrieval and mentalizing across the largest database of neuroimaging studies to date. We also conducted a meta-analysis of autobiographical tasks to test the hypothesis that autobiographical memory and future thought are complex processes that invoke a mixture of retrieval and mentalizing. We predicted that episodic retrieval and mentalizing would be preferentially associated with the distinct subsystems identified in Parts 1 and 2 and our previous work, while autobiographical tasks would be associated with a combination of both subsystems.

Our analyses confirmed both predictions. Similar to the pattern of task-related activity observed in Part 1, episodic retrieval tasks were preferentially linked to a network of regions throughout the medial temporal subsystem, including the HF, PHC, pIPL, and Rsp (Fig. 3A). Regions comprising the midline core of the DN (PCC and dMPFC) as well as the frontoparietal control network were also involved. Conversely, terms related to mentalizing were preferentially associated with a set of brain regions overlapping strongly with the dorsal medial subsystem – including the bilateral dMPFC, TPJ extending into superior temporal sulcus, anterior lateral temporal cortex, and the ventral MPFC (Fig. 3B). Supporting the hypothesis that autobiographical tasks recruit processes related to retrieval and mentalizing, the meta-analysis of autobiographical tasks overlapped with aspects of both subsystems, as well as the PCC and dMPFC core of the DN (Fig. 3C). The spatial pattern of overlap is highlighted in Fig. 4A. Fig. 4B displays the relative loadings on the three identified brain networks (i.e., those associated with episodic retrieval, mentalizing, and autobiographical memory) for the top seven terms in the Neurosynth database associated with each of the meta-analysis (episodic retrieval:...
recollection, retrieval, memory, episodic, remember, old, encoding; mentalizing: theory of mind, mentalizing, person, stories, social, mental, moral; autobiographical tasks: autobiographical, memories, person, retrieval, episodic, self-referential, self). Both the spatial and functional plots demonstrate (a) the relatively high degree of separation between episodic retrieval and mentalizing, and (b) the relative overlap of both functions with autobiographical memory/future thought.

Discussion

In recent years, a growing number of introspective processes have been linked to the brain’s default network, supporting its proposed overarching role in internal mentation (Andrews-Hanna, 2012; Andrews-Hanna et al., in press; Buckner et al., 2008; Smallwood et al., 2012). Among the most established of these relationships is the DN’s involvement in autobiographical memory and autobiographical future thought (Cabeza and St. Jacques, 2007; Gilboa et al., 2004; McDermott et al., 2009; Schacter et al., 2007, 2012; Spreng et al., 2009; Svoboda et al., 2006; Szpunar, 2012). However, recent findings that the DN is anatomically and functionally heterogeneous suggest functional specificity within the network (Andrews-Hanna et al., 2010a; Kim, 2012; Leech et al., 2011; Seghier and Price, 2012). Here we used multiple neuroimaging approaches to explore the nature of this specificity, focusing on the DN’s role in two processes hypothesized to contribute to and/or interact during autobiographical thought: episodic retrieval and mentalizing.

Converging results across task-related fMRI, rs-fcMRI, and fMRI meta-analyses revealed the medial temporal subsystem – a network of brain regions including the medial temporal lobe and its cortical projections, Rsp and pIPL – was preferentially linked to episodic retrieval tasks, while the dorsal medial subsystem – a network of brain regions including the dMPFC, TPJ, LTC and temporal pole – was preferentially linked to mentalizing tasks. Importantly, autobiographical tasks were linked to activity patterns spanning both subsystems, suggesting the
importance of episodic retrieval and mentalizing when individuals spontaneously or deliberately recall their past and imagine their future. Collectively, these results suggest the DN is a heterogeneous brain system whose distinct components have functionally dissociable roles that often interact during more complex forms of thought. It should be noted, however, that although we focus on “subsystems” as an important level of functional specificity within the DN, specificity could also be considered on a finer-grained scale down to neuronal assemblies within brain regions.

The role of the medial temporal subsystem in episodic retrieval and autobiographical thought

Prior reports from rs-fcMRI and task-related fMRI highlight a network of regions comprising a “medial temporal subsystem” as being an important functional-anatomic component of the DN (Andrews-Hanna et al., 2010a,b). Here we provide insight into the adaptive functions of the medial temporal subsystem by observing, across multiple studies, a pattern of activity overlapping strongly with this subsystem when individuals recollected prior information. These results are consistent with prior reports in humans and animals that regions within the medial temporal subsystem functionally correlate at rest (Kahn et al., 2008; Libby et al., 2012; Nelson et al., 2010; Vincent et al., 2006, 2010; Yeo et al., 2011) and are connected by long-range white matter tracts (Cavada and Goldman-Rakic, 1989; Greicius et al., 2009; Kobayashi and Amaral, 2003; Uddin et al., 2010; van den Heuvel et al., 2008).

Our findings also converge with laboratory-based fMRI retrieval studies, particularly those that distinguish between the episodic process of recollection and the arguably non-episodic process of familiarity (see Rugg and Yonelinas, 2003; Vilberg and Rugg, 2008; Yonelinas, 2002 for a discussion of the dual processes account of recognition memory). Meta-analyses comparing recollection to familiarity localize recollection to hippocampal and cortical regions overlapping with the medial temporal subsystem, and familiarity to a distinct set of frontoparietal regions (Cabeza et al., 2008; Hutchinson et al., 2012; Kim, 2010; Rugg and Vilberg, 2013; Spaniol et al., 2009; Vilberg and Rugg, 2008). However,
Whether the medial temporal subsystem is specific for episodic retrieval also remains a matter of debate. Regions throughout the medial temporal subsystem become engaged when participants retrieve contextual associations (Bar, 2007, 2009) and semantic knowledge (Binder and Desai, 2011; Binder et al., 2009; Seghier and Price, 2012), as well as when individuals simulate novel scenes (Hassabis et al., 2007) and imagine their autobiographical future (Addis et al., 2007; Andrews-Hanna et al., 2010a; D’Argembeau et al., 2009; Schacter et al., 2008, 2012; Spreng and Grady, 2010; Spreng et al., 2009; Szpunar et al., 2007). Collectively, these findings suggest the medial temporal subsystem plays an overarching role in memory-based construction by integrating past information into coherent spatial and temporal contexts (Andrews-Hanna, 2012; Andrews-Hanna et al., in press; Hassabis and Maguire, 2007; Ranganath and Ritchey, 2012; Schacter et al., 2007, 2012).

The role of the dorsal medial subsystem in mentalizing and autobiographical thought

Converging results from task-related fMRI, resting-state functional connectivity, and fMRI meta-analyses reveal an important role of the dorsal medial subsystem in reflecting on the mental states of other people. These findings are supported by a large body of literature linking the dMPFC, TPJ and lateral temporal cortex to spontaneous and volitional forms of mentalizing, including reflecting on one’s own mental states (reviewed in Beer and Ochsner, 2006; Denny et al., 2012; Frith and Frith, 2003; Gilbert et al., 2007; Lieberman, 2007; Mar, 2011; Mars et al., 2012; Ochsner et al., 2004; Olson et al., 2013; Saxe, 2006; Schilbach et al., 2008; Spreng et al., 2009; van der Meer et al., 2010; Van Overwalle, 2009). Recent studies suggest that different regions within the dMPFC subsystem support distinct aspects of social cognition. For example, the right TPJ has been suggested to play a particular role in reflecting on another person’s true and false beliefs (Dönhel et al., 2012; Saxe and Powell, 2006; Saxe and Kanwisher, 2003), especially when these processes require a participant to integrate present information with prior knowledge about that person to build a coherent model of another’s mind (Cloutier et al., 2011; Saxe and Wexler, 2005; Young et al., 2010). In contrast, the dMPFC activates more broadly when participants appraise or assess social information (i.e. personality traits and preferences; Beer and Ochsner, 2006; Hassabis et al., in press; Mitchell et al., 2006; Moran et al., 2011; Ochsner et al., 2004; Van Overwalle, 2009) and generate high-level construals pertaining to social and non-social information (Baetens et al., in press). Furthermore, a recent meta-analysis showed that while the dMPFC and left TPJ activate during self-reflection tasks, the right TPJ does not (Denny et al., 2012).

Regions throughout the dorsal medial subsystem also become engaged during conceptual processing and semantic retrieval, as well as when individuals read complex narratives (Andrews-Hanna et al., in press; Binder and Desai, 2011; Binder et al., 2009; Mar, 2011; Olson et al., 2013; Patterson et al., 2007; Yarkoni et al., 2008). However, many of the stimuli used in such studies are social in nature, and recent studies suggest the dorsal medial subsystem plays a particular role in processing social conceptual knowledge (Contreras et al., 2012; Ross and Olson, 2011; Skipper et al., 2011). Furthermore, narratives that encourage participants to reflect on others’ internal mental states or emotions engage the dorsal medial subsystem while narratives describing others’ external (i.e. physical) characteristics engage a distinct set of regions (Bruneau et al., 2012). Collectively, these findings suggest a broader role of the dorsal medial subsystem in mentalizing, perhaps by retrieving stored conceptual knowledge about one’s self and/or other people (Andrews-Hanna et al., in press).

The overlap between mentalizing and autobiographical tasks in Part 3 also raises the possibility that mentalizing represents an important aspect of autobiographical thought (Giambra, 1979; Mitchell, 2006; Schilbach et al., 2008; Singer, 1966; Spreng et al., 2009). Indeed, prior studies suggest that one’s social relationships contribute heavily to

Fig. 3. Meta-analyses of episodic retrieval, mentalizing, and autobiographical thought. The Neurosynth framework (Yarkoni et al., 2011) was used to conduct automated meta-analyses of A. episodic recollection tasks, B. mentalizing tasks, and C. autobiographical tasks across a large database of published neuroimaging studies. Shown are whole-brain corrected reverse inference maps reflecting the specificity of the particular pattern of activation for the given task, as compared to a broad corpus of other tasks in the database.

continued debate surrounds the precise functional role of these regions in memory retrieval (Cabeza et al., 2008; Ranganath and Ritchey, 2012; Rugg and Vilberg, 2013; Vann et al., 2009; Vilberg and Rugg, 2008; Wagner et al., 2005). In contrast, laboratory-based recognition tasks that do not distinguish between recollection and familiarity reveal activity patterns suggestive of both processes at play or a predominance of non-episodic contributions to retrieval (Burianova et al., 2010; McDermott et al., 2009; Nelson et al., 2010).

The role of the dorsal medial subsystem in mentalizing and autobiographical thought

Converging results from task-related fMRI, resting-state functional connectivity, and fMRI meta-analyses reveal an important role of the dorsal medial subsystem in reflecting on the mental states of other people. These findings are supported by a large body of literature linking the dMPFC, TPJ and lateral temporal cortex to spontaneous and volitional forms of mentalizing, including reflecting on one’s own mental states (reviewed in Beer and Ochsner, 2006; Denny et al., 2012; Frith and Frith, 2003; Gilbert et al., 2007; Lieberman, 2007; Mar, 2011; Mars et al., 2012; Ochsner et al., 2004; Olson et al., 2013; Saxe, 2006; Schilbach et al., 2008; Spreng et al., 2009; van der Meer et al., 2010; Van Overwalle, 2009). Recent studies suggest that different regions within the dMPFC subsystem support distinct aspects of social cognition. For example, the right TPJ has been suggested to play a particular role in reflecting on another person’s true and false beliefs (Dönhel et al., 2012; Saxe and Powell, 2006; Saxe and Kanwisher, 2003), especially when these processes require a participant to integrate present information with prior knowledge about that person to build a coherent model of another’s mind (Cloutier et al., 2011; Saxe and Wexler, 2005; Young et al., 2010). In contrast, the dMPFC activates more broadly when participants appraise or assess social information (i.e. personality traits and preferences; Beer and Ochsner, 2006; Hassabis et al., in press; Mitchell et al., 2006; Moran et al., 2011; Ochsner et al., 2004; Van Overwalle, 2009) and generate high-level construals pertaining to social and non-social information (Baetens et al., in press). Furthermore, a recent meta-analysis showed that while the dMPFC and left TPJ activate during self-reflection tasks, the right TPJ does not (Denny et al., 2012).

Regions throughout the dorsal medial subsystem also become engaged during conceptual processing and semantic retrieval, as well as when individuals read complex narratives (Andrews-Hanna et al., in press; Binder and Desai, 2011; Binder et al., 2009; Mar, 2011; Olson et al., 2013; Patterson et al., 2007; Yarkoni et al., 2008). However, many of the stimuli used in such studies are social in nature, and recent studies suggest the dorsal medial subsystem plays a particular role in processing social conceptual knowledge (Contreras et al., 2012; Ross and Olson, 2011; Skipper et al., 2011). Furthermore, narratives that encourage participants to reflect on others’ internal mental states or emotions engage the dorsal medial subsystem while narratives describing others’ external (i.e. physical) characteristics engage a distinct set of regions (Bruneau et al., 2012). Collectively, these findings suggest a broader role of the dorsal medial subsystem in mentalizing, perhaps by retrieving stored conceptual knowledge about one’s self and/or other people (Andrews-Hanna et al., in press).

The overlap between mentalizing and autobiographical tasks in Part 3 also raises the possibility that mentalizing represents an important aspect of autobiographical thought (Giambra, 1979; Mitchell, 2006; Schilbach et al., 2008; Singer, 1966; Spreng et al., 2009). Indeed, prior studies suggest that one’s social relationships contribute heavily to
autobiographical memories (Lardi et al., 2010; Singer et al., 2007; Thorne et al., 2004) and spontaneous thoughts individuals experience in the laboratory and daily life (Andrews-Hanna et al., 2013; Mar et al., 2012; Ruby et al., 2013; see also Immordino-Yang et al., 2012). A recent study found that 40% of participants’ self-defining memories involved other people, and that reflecting on the broader meaning of autobiographical memories (as compared to retrieving specific items from the past) engaged regions throughout the dorsal medial subsystem (D’Argembeau et al., in press). Social information also represents a major topic of interpersonal conversation (Dunbar, 1997).

---

Fig. 4. Autobiographical thought comprises multiple component processes. A. In this overlap analysis, whole-brain reverse inference maps from the meta-analyses of episodic retrieval, mentalizing, and autobiographical tasks were corrected for multiple comparisons and overlaid on a surface projection (Van Essen, 2005). B. The association between the top seven functional terms linked to the episodic recollection and mentalizing meta-analyses are shown in relationship to each other (green = episodic recollection; blue = mentalizing) and the autobiographical thought meta-analysis (red). Both panels illustrate the functional differences between episodic recollection and mentalizing and the overlap of both functions with autobiographical memory/future thought. The distance from the origin in the polar plot in Panel B reflects the Pearson correlation across all voxels between each meta-analysis map for the concept as a whole and individual terms in the Neurosynth database. For example, the meta-analysis map for the concept of mentalizing (defined by several related terms) correlated 0.45 with the meta-analysis map for the individual term “social.” If a network map or mask loaded highly on multiple terms that shared a morphological root (e.g., retrieval and retrieved; story and stories), the term with the highest loading was included in the figure.
Functional characteristics of the midline core of the DMN

In our studies, the PCC was engaged to a common degree across tasks and functionally correlated with both task-defined parietal seeds to an equivalent degree at rest. Consistent with these findings, the PCC comprises dorsal and ventral components that exhibit widespread connectivity patterns throughout the DN and the rest of the brain (Andrews-Hanna et al., 2010a; Buckner et al., 2008, 2009; Dasterjedi et al., 2011; Hagmann et al., 2008; Leech and Sharp, 2014; Leech et al., 2011; Margulies et al., 2009; Vogt et al., 2006). The PCC is also engaged across a wide range of introspective tasks (Andrews-Hanna, 2012; Buckner et al., 2008), suggesting it may integrate relevant information from a variety of perceptual, attentional and mnemonic sources (see also Leech et al., 2011; Pearson et al., 2011).

Though the aMPFC shares similar properties with the PCC, the aMPFC is particularly responsive to tasks that involve reflecting upon, or subjectively evaluating, personally-significant information (Andrews-Hanna, 2012; Roy et al., 2012). Self-referential strategies explain a large portion of the variance in aMPFC activity during introspective tasks (Andrews-Hanna et al., 2010a), and aMPFC activity increases with perceived self-relevance (Benoit et al., 2010; Denny et al., 2012; Krienen et al., 2010; Macrae et al., 2004; Mitchell et al., 2006; Moran et al., 2006, 2009). The extensive activation of the aMPFC during autobiographical tasks compared to laboratory-based retrieval tasks may reflect strategies involving evaluation of the personal meaning, or overarching significance, of autobiographical thoughts (see also Cabeza and St Jacques, 2007; Cabeza et al., 2004; D’Argembeau et al., in press; Gilboa, 2004; Kim, 2012; Levine, 2004; Roy et al., 2012).

Summary, limitations, and future directions

In summary, three studies employing task-related fMRI, resting state fMRI, and fMRI meta-analyses revealed largely distinct patterns of activity associated with episodic recollection and mentalizing, and overlapping patterns of activity with autobiographical thought. These findings speak to an ongoing debate regarding the function of the DN (Andrews-Hanna, 2012; Bar, 2007, 2009; Binder et al., 2009; Buckner and Carroll, 2007; Buckner et al., 2008; Hassabis and Maguire, 2007; Schacter et al., 2007; Smallwood et al., 2012; Spreng et al., 2009) by suggesting that different DN components contribute differently to autobiographical thought. Synthesizing these findings, we suggest that recalling a personal event from the past or imagining a possible future event is likely to comprise several interacting component processes including retrieval of specific items or sources of knowledge, integration of such items into a coherent spatio-temporal context, and reflecting on our own and/or others’ feelings, emotions, and perceived significance surrounding the autobiographical event.

However, there is considerable within and between-subject variability in the content characterizing autobiographical thoughts (Andrews-Hanna et al., 2013; Ruby et al., 2013; Stawarczyk et al., 2013), and the precise nature of this content is likely to be a key factor underlying the involvement of the DN subsystems in autobiographical experiences. When interpreted in this context, our results are largely consistent with prior studies observing patterns of overlap and non-overlap between autobiographical and theory of mind tasks (Rabin et al., 2010; Spreng and Grady, 2010; Spreng and Mar, 2012; Spreng et al., 2009).

One limitation of our findings is that autobiographical tasks were not administered to the same group of participants as those in Part 1 and 2. However, relationships between episodic retrieval, mentalizing, and autobiographical memory/future thought were assessed in Part 3 using large-scale meta-analyses which are less influenced by intricacies of specific task paradigms or sample selection criteria than a single study. Additionally, we observed strong convergence between the tasks employed in Part 1 and the meta-analysis of episodic retrieval and mentalizing in Part 3, leading us to anticipate a similar degree of convergence for autobiographical tasks.