

Using state-trace analysis to dissociate the functions of the human hippocampus and perirhinal cortex in recognition memory

Bernhard P. Staresina^{a,b}, Juergen Fell^b, John C. Dunn^c, Nikolai Axmacher^{b,d}, and Richard N. Henson^{a,1}

^aCognition and Brain Sciences Unit, Medical Research Council, Cambridge CB2 7EF, United Kingdom; ^bDepartment of Epileptology, University of Bonn, D-53105 Bonn, Germany; ^cSchool of Psychology, University of Adelaide, Adelaide, SA 5005, Australia; and ^dGerman Center of Neurodegenerative Diseases (DZNE), 53175 Bonn, Germany

Edited* by Mortimer Mishkin, National Institute for Mental Health, Bethesda, MD, and approved January 8, 2013 (received for review September 10, 2012)

A recurring issue in neuroscience concerns evidence as to whether two or more brain regions implement qualitatively different functions. Here we introduce the application of state-trace analysis to measures of neural activity, illustrating how this analysis can furnish compelling evidence for qualitatively different functions, even when the precise “neurometric” mapping between function and brain measure is unknown. In doing so, we address a long-standing debate about the brain systems supporting human memory: whether the hippocampus and the perirhinal cortex, two key components of the medial temporal lobe memory system, provide qualitatively different contributions to recognition memory. An alternative account has been that both regions support a single shared function, such as memory strength, with the apparent dissociations obtained by previous neuroimaging studies merely reflecting different, nonlinear neurometric mappings across regions. To adjudicate between these scenarios, we analyze intracranial electroencephalographic data obtained directly from human hippocampus and perirhinal cortex during a recognition paradigm and apply state-trace analysis to responses evoked by the retrieval cue as a function of different types of memory judgment. Assuming only that the neurometric mapping in each region is monotonic, any unidimensional theory (such as the memory-strength account) will produce a monotonic state trace. Critically, results showed a nonmonotonic state trace; that is, activity levels in the two regions did not show the same relative ordering across memory conditions. This nonmonotonic state trace demonstrates that there are at least two different functions implemented across the hippocampus and perirhinal cortex, allowing formal rejection of a single-process account of medial temporal lobe contributions to recognition memory.

A fundamental problem in neuroscience concerns the mapping between a neural measure (dependent variable) and the hypothetical quantity (latent variable) assumed to vary with an experimental manipulation. In most cases, it seems unlikely that this mapping is linear, for example that a doubling in working memory load would always double the magnitude of the hemodynamic response measured by functional magnetic resonance imaging (fMRI). The unknown nature of this “neurometric” mapping is particularly problematic for attempts to dissociate the function of two or more brain regions (1).

One research domain in which this measurement problem has recently been highlighted is the domain of human recognition memory. Although there is consensus that intact memory relies on the medial temporal lobe (MTL) (e.g., 2, 3), there has been intense debate about the relative contributions of different MTL structures, particularly the hippocampus and the adjacent perirhinal cortex, to different expressions of memory. For example, one set of theories, referred to as dual-process models and largely building on functional dissociations obtained via fMRI, argues for a role of the hippocampus in recollection/associative memory and for a role of the perirhinal cortex in familiarity/item memory (e.g., 4–7). However, others have argued that the existing fMRI data are not sufficient to support the claim that these MTL

regions implement qualitatively different types of memory processes (8–10) (see also ref. 1).

The argument made by Squire et al. (10), emphasizing the above problem of unknown neurometric mappings, is illustrated in Fig. 1A. Assume that there are three experimental conditions (condition 1, condition 2, and condition 3). In the test phase of a recognition memory paradigm, these conditions might correspond to three trial types: (i) correct rejection of stimuli not seen previously; (ii) correct recognition of previously seen stimuli without recall of contextual details; and (iii) correct recognition of previously seen stimuli with recall of contextual details. The comparisons of condition 1 versus condition 2, and of condition 2 versus condition 3, are typical of those used to isolate the qualitatively different processes of familiarity and recollection, and are often attributed to the perirhinal cortex and hippocampus, respectively (see above). This attribution is based on conventional statistical tests within each of these regions, in which activity in the perirhinal cortex may differ significantly between conditions 1 and 2 but not between conditions 2 and 3, whereas activity in the hippocampus may differ significantly between conditions 2 and 3 but not between conditions 1 and 2. Indeed, there may even be a significant interaction between the two regions and the three conditions. However, as Squire et al. observed, the three experimental conditions are also likely to differ along a single dimension of memory strength (lowest in condition 1 and highest in condition 3 in Fig. 1A). Critically, the same pattern of significant results could be observed if the perirhinal cortex and hippocampus had different neurometric mappings between the level of this hypothetical memory strength and the measured fMRI signal. First, perirhinal fMRI signal might decrease with increasing memory strength, whereas hippocampal fMRI signal might increase with increasing memory strength. Second, fMRI signal in both regions may be nonlinearly related to memory strength, such that perirhinal fMRI signal approaches a minimal level at high memory strengths whereas hippocampal fMRI signal approaches a minimal level at low memory strengths (Fig. 1A). With such potentially different, nonlinear neurometric mappings across different MTL regions, standard analyses of fMRI data are therefore inconclusive as to whether qualitatively different functions are implemented by different MTL regions. In the absence of such evidence, it is arguably more parsimonious to assume that activity in both regions reflects the same, shared process such as memory strength.

Author contributions: B.P.S. and R.N.H. designed research; B.P.S., J.F., and N.A. performed research; J.C.D. and R.N.H. contributed new reagents/analytic tools; B.P.S. and R.N.H. analyzed data; and B.P.S., J.F., J.C.D., N.A., and R.N.H. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

¹To whom correspondence should be addressed. E-mail: rik.henson@mrc-cbu.cam.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1215710110/-DCSupplemental.

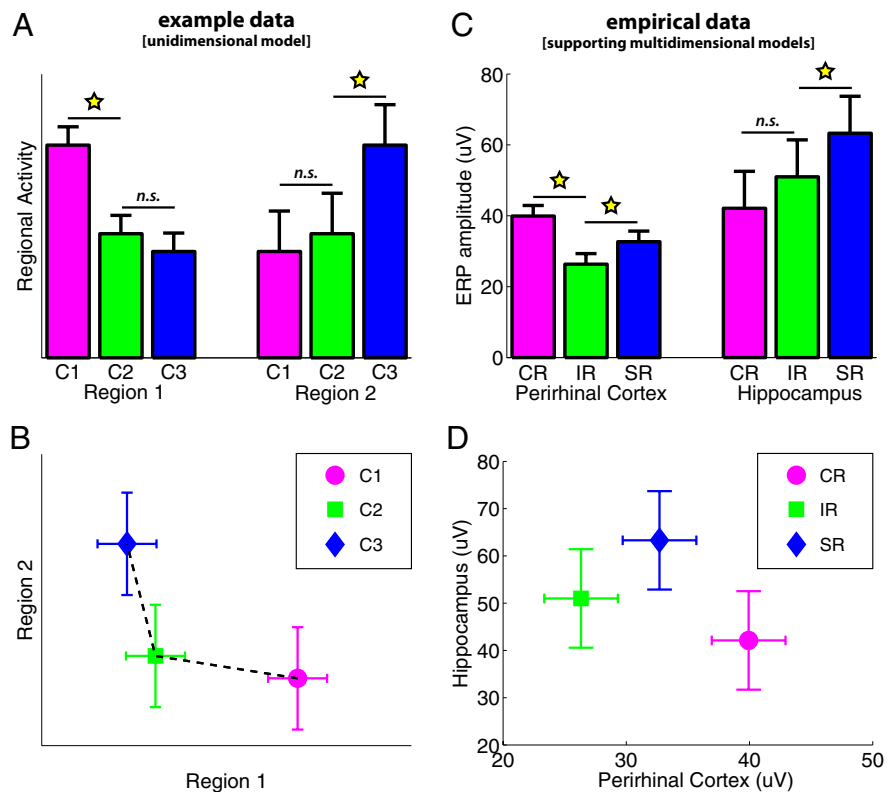


Fig. 1. (A and B) Synthetic data, modeled after refs. 15–18 and exemplifying activity in two brain regions across three experimental conditions (C1–C3). (C and D) Real data from the five participants in the present iEEG experiment. (A and C) Conventional bar graphs, with SE of within-participant pooled variance. (B and D) State-trace plots for the same data, where activity in one region is plotted against that in the other for the three conditions. The synthetic data in A might appear to show dissociable responses across the two regions, but can actually be explained by sensitivity to the same underlying variable (e.g., memory strength) but with different neurometric functions, in which activity is sensitive to either low or high levels of memory strength in regions 1 and 2, respectively. This corresponds to the state trace in B that can be fit by a nonlinear, but nonetheless monotonic, function. The data in C reflect the absolute mean evoked response across a 100-ms window centered on the participant-specific peak time point from electrodes in the perirhinal cortex and hippocampus, for correct rejection, item recognition, and source recognition trial types. Importantly, when replotted these data in D, there is reliable evidence that the mean state trace across participants cannot be fit by a monotonic function (*Results*), thus refuting the notion of a single shared function of memory strength across both regions. Stars, $P < 0.05$, two-tailed; n.s., not significant.

A solution to the dilemma that neuroimaging dissociations across regions might simply reflect different, nonlinear neurometric mappings is afforded by a method called “state-trace” analysis. Developed in the psychological literature (e.g., 11–14), this method—if applied to neural measures—allows one to infer that more than one function is implemented across two or more brain regions, with the minimal assumption that the neurometric mapping within each region is at least monotonic, that is, that an increase in functional engagement always results in an increase (or always results in a decrease) in neural response (*Methods*). Briefly, state-trace analysis entails plotting the data for each experimental condition (of which there must be at least three) as a point in a “state” space, whose axes are defined by the dependent variables (of which there must be at least two). According to the null hypothesis that only one hypothetical function is implemented by those regions, a “trace” curve can be drawn between these points that is monotonic (a monotonically decreasing example is shown in Fig. 1B, indicated by the dotted line). Here we apply this analysis to measures of neural activity, namely to the amplitude of event-related potentials (ERPs) from intracranial electroencephalographic (iEEG) recordings obtained directly from the hippocampus and perirhinal cortex of human epilepsy patients. The reliable, non-monotonic state trace that results provides compelling evidence that these two regions implement more than one function during recognition memory tasks.

Results

The key dependent variable of interest for the current analysis is the mean ERP amplitude across a 100-ms time window centered on the time point showing the peak response for each region and each participant after averaging across conditions (Fig. 1C). For further details, such as behavioral accuracy or mean reaction times, see ref. 19. The three conditions of interest were: (i) correct classification of a previously unseen item as new (correct rejection; CR); (ii) correct identification of a previously seen item as old, without additionally remembering the associated source detail (item recognition; IR); and (iii) correct identification of a previously seen item as old plus correct memory for the associated source detail (source recognition; SR) (Fig. S1). Pairwise comparisons of conditions showed significant differences for CR vs. IR and SR vs. IR in the perirhinal cortex [both $t(4) > 3.29$, $P < 0.05$] and a significant difference for SR vs. IR [$t(4) = 2.81$, $P < 0.05$] although not for CR vs. IR [$t(4) = 0.96$, $P = 0.39$] in the hippocampus (for individual peak latencies and condition amplitudes, see Fig. S2).

The state-trace plot for these data is shown in Fig. 1D. To reject the null hypothesis of a monotonic state trace, we fit a common monotonic regression model to the mean data averaged over participants, and compared the goodness of fit of either a monotonically increasing fit or a monotonically decreasing fit against an empirical sampling distribution (using 10,000 Monte Carlo samples; *Methods*).

First, we directly examined the fit of the memory-strength model described in the introduction, which predicts a specific ordering of the three conditions in which memory strength increases from CR to IR to SR. A monotonically increasing fit, in which the magnitude of both hippocampus and perirhinal ERP components increases with increasing memory strength (i.e., $CR < IR < SR$), was rejected with $P < 0.0001$. The relationship proposed by Squire and colleagues (10), however, at least for retrieval, is a monotonically decreasing relationship, in which perirhinal activity decreases with memory strength (whereas hippocampal activity increases). In other words, this proposal requires the conditions be ordered $CR > IR > SR$ in the perirhinal cortex and $CR < IR < SR$ in the hippocampus. Whereas the hippocampus data were so ordered ($CR = 42$, $IR = 51$, $SR = 63$), the perirhinal data were not ($CR = 40$, $IR = 26$, $SR = 33$), such that together, a monotonically decreasing relationship could be rejected too, $P = 0.0098$.

In a second analysis, we tested whether the conditions could be ordered the same way in both regions along *any* single dimension (not necessarily according to the memory-strength hypothesis). For a monotonically increasing relationship between perirhinal and hippocampal data, all of the six possible orderings of the three conditions could be rejected, $P = 0.0125$. For a monotonically decreasing relationship between perirhinal and hippocampal data, the best-fitting order was $CR < SR < IR$, which had a corresponding $P = 0.0538$. Together, the results of these two analyses refute not only the specific memory-strength model outlined in the introduction but also render unlikely any theory that would order the three conditions in any fashion along a single dimension.

Discussion

The present intracranial ERP data provide compelling evidence that neural activity in the hippocampus and perirhinal cortex must reflect at least two qualitatively different functions engaged during a recognition memory task. This refutes the possibility that activity in these regions can be explained by a single dimension such as memory strength, and provides the crucial evidence that has so far been missing from functional neuroimaging of such memory tasks (10). In doing so, we have demonstrated the potential of state-trace analysis to answer such neuroscientific questions.

How do our results go beyond previous neuroimaging studies of recognition memory? A number of fMRI studies have previously claimed functional dissociations between the perirhinal cortex and hippocampus during recognition memory (for recent reviews, see refs. 5 and 6, although see also refs. 20–22 for counterclaims). Across those studies, one consistent finding has been that perirhinal cortex activity decreases from correct rejection of new items to familiarity- and recollection-based recognition of old items. This pattern has been observed using the remember/know/new procedure (16), confidence ratings ranging from 1 (definitely new) to 6 (definitely old) (15), or a hybrid procedure comparing correct rejection of new items and three levels of familiarity plus a “recollection” response for old items (17). Conversely, the hippocampus has typically been found to show a recollection effect, without further differentiating between familiar old items and new items (15, 17, 18). This is the pattern schematized in Fig. 1A. However, as pointed out by Squire and colleagues (10) and explained in the introduction, such patterns across the perirhinal cortex and hippocampus are still compatible with a single-process, memory-strength account if the perirhinal cortex and hippocampus have different nonlinear neurometric functions. Indeed, when applying state-trace analysis to a simulated pattern that mimics the fMRI results mentioned above (Fig. 1A), the resulting state-trace plot can still be fit by a monotonically decreasing function (Fig. 1B), and hence the single memory-strength account cannot be refuted ($P = 1$). The points in the state-trace plot for our intracranial recordings, on the other

hand, cannot be fit by a monotonic trace (Fig. 1D), indicating that they cannot be explained by this memory-strength model.

Note that we do not claim that iEEG data are necessary to reject such models; the fMRI data from some of the above studies claiming functional dissociations may also conform to a nonmonotonic state trace, at retrieval, like here (e.g., 23), or at encoding (e.g., 24, 25). In many of these studies, however, the data were extracted from voxels within MTL regions that were first identified by one or more statistical contrasts across conditions, which biases any subsequent statistical tests (such as state-trace analysis) across regions [i.e., renders these analyses circular (1, 26)]. In our iEEG study, the contacts and time windows analyzed were selected in an unbiased fashion (Methods and Fig. S2). For the few fMRI studies that selected data from MTL regions in an unbiased manner (e.g., 24), the question is then whether those data would be able to formally reject a monotonic model; that is, it would be necessary to apply the state-trace analysis of the type demonstrated here.

In addition to rejecting the specific memory-strength model, regardless of whether memory strength is positively or negatively monotonically related to neural activity in the hippocampus and perirhinal cortex, our data are also difficult to explain by any single-process model, that is, any one-dimensional model that orders our three conditions in the same way in both MTL regions. Indeed, our data reject any such model that assumes a positive monotonic relationship in which the hippocampus and perirhinal cortex respond in the same direction to changes in the underlying dimension (i.e., both increase or both decrease their response, resulting in a monotonically increasing state trace). Our data do not quite reject a one-dimensional model in which the hippocampus and perirhinal cortex show a negative monotonic relationship (i.e., in which the hippocampal response increases but the perirhinal response decreases, or vice versa, to the same change in underlying process, resulting in a monotonically decreasing state trace), although the P value (0.0538) was very close to conventional significance levels (0.05). We are not aware that any such model, in which our conditions would need to be ordered $CR < SR < IR$ along the underlying dimension, has yet been proposed.

That said, it is important to note that whereas our observation of a nonmonotonic state trace supports multidimensional memory theories that postulate more than one hypothetical function across regions (such as the dual-process models of recognition mentioned in the introduction), it does not on its own tell us what those functions (processes) are. Nor does it imply that each brain region implements mutually exclusive processes in our task. For example, our own preferred interpretation is that both regions conjointly support source memory, whereas only the perirhinal cortex additionally supports novelty detection (see ref. 19 for further discussion). Thus, although our data might be used to support previous claims that the perirhinal cortex supports familiarity and that the hippocampus supports recollection (e.g., 4, 5, 27), there may be yet other multidimensional explanations of our data, such as activity in both of these MTL regions being related to a single memory-strength signal, plus a second, possibly nonmnemonic function, that also happens to differ across our three conditions. Indeed, Squire et al. did not claim that the hippocampus and perirhinal cortex never perform different processes, only that these processes have not yet been properly characterized and tested in neuroimaging experiments, and that the anatomical inputs to these regions have not been taken sufficiently into account (28).

The nonmonotonic state trace in our study was obtained from data that captured the initial peak response in each region, individually defined for each participant (Fig. S2). One might wonder whether the implication of at least two different processes during the first second of an evoked response is of little significance, because the two brain regions simply need time to “converge”

on activity that represents a single memory signal such as memory strength. Indeed, in our previous analysis (19), we provided evidence that the perirhinal cortex and hippocampus become increasingly functionally coupled during the first second after stimulus onset during SR trials. However, although this coupling might indicate a shared process, it does not preclude other processes simultaneously co-occurring within one or both brain regions (e.g., novelty detection in the perirhinal cortex).

State-trace analysis is an elegant method that is beginning to be applied to behavioral data to address a number of theoretical models (e.g., 14, 29–31). The only assumption needed for using state-trace analysis in the context of neural measures is that the neurometric mappings between the dependent variable in each region (here, ERP magnitude) and the hypothetical latent variable (here, memory strength) are monotonic. Note that monotonic mappings can include nonlinear mappings, such as the floor/ceiling effects illustrated in Fig. 14. On that note, although the possibility of nonlinear neurometric mappings seems particularly appealing for fMRI [given that blood oxygenation level-dependent (BOLD) signal is known to be a complex and nonlinear function of several hemodynamic variables (e.g., 32) and the vasculature is known to differ across different parts of the brain (33)], we see no reason why the possibility of nonlinear neurometric mappings should not apply equally to other measures in neuroscience, such as neuronal firing rates or local field potentials. A major strength of state-trace analysis is that it does not require neurometric mappings to be linear, merely monotonic (i.e., increased expression of the latent variable always leads to increases, or always leads to decreases, in the neural measure). One might argue that some neurometric mappings might be nonmonotonic, for example, the U-shaped relationship that has been observed between BOLD signal in auditory cortex and rate of auditory word presentation (34). However, without a priori theoretical reasons for proposing such nonmonotonic mappings, the default assumption for monotonic mappings seems necessary to make progress in neuroscience. Otherwise, if the neurometric mapping could take any form, then it would be possible to explain any pattern of data with any theory.

Note that a “reversed association”—a pattern of data that has previously been claimed necessary to imply qualitatively different processes in behavioral (35) and neuroimaging (1) data—is subsumed by state-trace analysis as a special case. Importantly, the pattern of significant pairwise differences claimed necessary for a reversed association is sufficient, but not necessary, for a nonmonotonic state trace (36) (and state-trace analysis, by implementing a single omnibus test rather than a series of pairwise tests, can therefore be more powerful). Statistical inference for state-trace analysis is an active area of investigation, and here we have taken a standard bootstrapping approach to classical null hypothesis significance testing, although Bayesian model comparison methods have been developed more recently (14). As with any methodology, the application of state-trace analysis to neuroscience data implicitly makes further “bridging” assumptions, for example that the neural measure is a sufficient summary of a brain region’s current activity, in terms of spatial scale, for example, and that it is causally related to the hypothetical processes being manipulated experimentally. A further issue related to state-trace analysis, at least when applied to complex, interactive systems such as the brain, is whether a nonmonotonic state trace across two or more regions actually arises from functional differences within those regions, or is a consequence of two or more different functions occurring upstream of those regions (which differentially and nonlinearly transform the input to the regions of interest; see ref. 36 for further discussion). Methodological questions like these clearly deserve further consideration in the neuroscience literature.

In sum, by introducing state-trace analysis to measures of neural activity, the current results provide compelling evidence for the existence of at least two qualitatively different functions

across the human hippocampus and perirhinal cortex during recognition memory.

Methods

Participants. Intracranial EEG was recorded from five patients (three female) suffering from pharmacoresistant epilepsy. Depth electrodes comprising 10 platinum contacts were implanted stereotactically along the longitudinal axis of each MTL (Fig. S1A) during presurgical evaluation. In four of the five participants, iEEG recordings identified a unilateral seizure-onset zone in the MTL, and only data from the contralateral hemisphere were used for analyses. For the remaining male participant, no seizure-onset zone was identified within the MTL of either hemisphere, and data from the left hemisphere were used according to the selection criteria described below. Thus, right-hemisphere data were used from three participants and left-hemisphere data from the remaining two participants. Participants ranged in age from 19 to 51 y (mean 34 y) and in duration of their epilepsy from 8 to 46 y (mean 22 y). At the time of the recordings, all participants received anticonvulsive medication (plasma levels within the therapeutic range). Informed consent for the iEEG recordings and the use of the data for research purposes was obtained from all participants. The study was approved by the Ethics Committee of the Medical Faculty of the University of Bonn.

Procedure. The experiment was conducted in a sound-attenuated room, with the participant sitting upright in a comfortable chair. A laptop computer, used for stimulus presentation, was positioned on a table at an ~50-cm distance. The stimulus material consisted of 375 German nouns and four different associated source details: the colors blue and red, and the scenes “office” and “nature” (Fig. S1B). There were six study-test cycles, with each study or test block using only one category (colors or scenes), and color and scene cycles were alternated, with the assignment of the first cycle counterbalanced across participants. As all experimental parameters were identical for color and scene conditions, and because we observed no performance differences (e.g., 19), data from color and scene conditions were collapsed for the current analysis to increase statistical power. The assignment of nouns to the list of study items or lures for the test phase and to color vs. scene conditions was randomized across participants.

During each study block, participants saw 50 nouns together with one of two possible sources. Nouns were presented in white uppercase letters and centered on a black background. The associated source (a color or scene, depending on the current block) was presented in a 250 × 350-pixel frame positioned 150 pixels underneath the noun. A given noun/source combination was presented on the screen for 3 s and participants indicated whether the given combination was plausible or implausible. Each trial was preceded by a jittered intertrial interval (ITI) (700–1,300 ms, mean 1,000 ms) during which a fixation cross was shown in the center of the screen. Participants were encouraged to give their responses as fast as possible. A study block was followed by a 60-s distraction period during which the participant conversed with the experimenter.

Test blocks contained all 50 previously seen (studied) nouns as well as 25 experimentally novel (unstudied) nouns (lures). The lures were pseudorandomly intermixed, holding the average delay between study and test constant across nouns. Upon being presented with a noun, participants could give one of four possible answers: (i) new (item not seen during the study phase); (ii) old, seen with source 1 (blue or office for color and scene runs, respectively); (iii) old, seen with source 2 (red or nature for color and scene runs, respectively); or (iv) old, but cannot remember the source (“don’t know” response). Thus, participants indicated, with one button press, whether they thought the noun was old or new and whether they also remembered the studied source detail. Responses were given in a self-paced manner, with an upper time limit of 5 s (only three responses were given after 5 s across all participants). Like in the study block, each trial was preceded by a jittered ITI (700–1,300 ms, mean 1,000 ms) showing a fixation cross.

One study-test cycle lasted 9 min. One participant did all six cycles in one session, three participants did the first four cycles in one session and the remaining two cycles in another session the same day, and one participant did the first four cycles in one session and the remaining two cycles in another session the next day. Stimuli were presented using Presentation (Neurobehavioral Systems). The whole iEEG experiment lasted ~1 h.

The experimental paradigm yielded three conditions of interest: (i) unstudied items correctly identified as new (correct rejection); (ii) studied items correctly identified as old, without remembering the correct source from the study episode (item recognition); and (iii) studied items correctly identified as old, additionally remembering the correct source (source recognition). For IR, to increase statistical power, we collapsed trials for which participants gave “don’t know” responses and trials for which the wrong source was

indicated. Misses (studied items incorrectly classified as new) and false alarms (unstudied items incorrectly classified as old) were not included in the analysis.

Electrode Selection and Artifact Rejection. Depth electroencephalograms were referenced to linked mastoids and recorded with a sampling rate of 1 kHz. Perirhinal cortex and hippocampal contacts were selected based on anatomical and functional properties. First, only contacts located in the perirhinal cortex and anterior hippocampus were considered. To this end, the postimplantation MRI was coregistered to the preimplantation MRI, assessing correspondence of individual electrode contacts with anatomical landmarks of the perirhinal cortex and hippocampus. In three of the five participants, the selected contact was clearly located in the perirhinal cortex, whereas for the remaining two participants, the contact was located between the peri- and entorhinal cortex. However, there was no qualitative difference in the overall response profiles, and we refer to the ento-/perirhinal contact as perirhinal for brevity. Given the prototypical negative component (“N400”) characterizing functional ento-/perirhinal recordings (e.g., 37), we further required the perirhinal cortex response profile to contain a clear negative peak within the first second post stimulus onset. A peak was defined as any time point whose negative amplitude exceeded two SDs of all negative values from 0 to 2 s, based on the average of all conditions of interest (CR, IR, SR) to avoid any selection bias. If more than one contact fulfilled the criteria for perirhinal cortex and hippocampus selection, we chose the contact with the highest absolute amplitude (baseline-corrected) summed across the first second. For further analysis and validation that the electrodes were not picking up a common source, see ref. 19.

Artifact rejection was performed on trial epochs from -1 to 3 s time-locked to stimulus onset. Before manual artifact rejection, an automated procedure was implemented in MATLAB (MathWorks) to reject trials in which at least one time point in the perirhinal cortex or hippocampus exceeded three interquartile ranges of all trial-specific values in both amplitude and gradient (difference from previous time point). Across participants, an average of 26% (range 16–39%) of all trials from the recognition memory phase was excluded.

Event-Related Potential Analysis. On average, there were 87 CR trials (range 32–115), 104 IR trials (range 50–155), and 85 SR trials (range 34–162) across participants after artifact rejection. Statistical analyses were performed on the unfiltered raw ERPs after baseline correction (subtracting the average 100-ms prestimulus interval). For statistical analyses, a 100-ms time window was centered on each participant’s peak value in each region within the first 1,000 ms after stimulus onset (where the peak was identified after collapsing all three conditions and applying a 10-Hz low-pass filter) (Fig. S2) (note that the same results were obtained when using the peak values themselves,

rather than a 100-ms time window centered on the peaks). All t tests were two-tailed. Note that, although the original perirhinal voltage was negative relative to baseline, the data in Fig. 1 represent the absolute value of the voltage deviation from baseline, given that the polarity of the evoked response is somewhat arbitrary (depending, for example, on the cytoarchitecture and orientation of the layers of the anatomical structure relative to the recording and reference electrodes)—that is, it is the magnitude of the initial evoked component that is likely to be related to memory, even though this component was positive in the hippocampus and negative in the perirhinal cortex. Nonetheless, the sign of the perirhinal data did not affect the significance of results, given that testing a positive versus negative monotonic relationship between two dependent variables (*Results*) is equivalent to flipping the sign of one of the dependent variables; that is, the tests of a negative monotonic relationship between hippocampus and perirhinal cortex in *Results* is equivalent to keeping the original negative polarity of the perirhinal voltage.

The mean values of the simulated data in Fig. 1 *A* and *B* were based on figure 5 of Squire et al. (10), reflecting a commonly reported pattern in the fMRI literature (15–18), and the noise levels (SEs) were chosen to approximately match those in the real data in Fig. 1 *C* and *D*.

Testing Monotonicity of State Traces. To fit a monotonic model to 2D data means finding an order on the x and y variables (here, data from two brain regions) that maximizes a goodness-of-fit measure. For each possible order of the three experimental conditions (e.g., condition $1 < 2 < 3$), monotonic regression is used to find the one that best fits both dependent variables simultaneously in a least-squares sense. To establish the significance of the fit, a modified parametric bootstrap method was used (38) in which the data were bootstrap-sampled a large number of times (10^4 in our analyses). For each bootstrap, the monotonic model was fit to the data, and the best-fitting parameters (here corresponding to the condition means) were then used to randomly generate new data, and these new data were refit by the model. The obtained goodness of fit therefore represents a sample from the distribution of that fit statistic given that the model is true, taking into account uncertainty in the estimation of its parameters. The fit statistic for the observed data is then compared with the obtained empirical distribution to give a P value that the observed data come from that distribution (see ref. 31 for further details). If fewer than 5% of bootstraps give a fit statistic worse than that for the observed data, the monotonic model is rejected.

ACKNOWLEDGMENTS. This work was supported by a Sir Henry Wellcome Postdoctoral Fellowship (to B.P.S.), Deutsche Forschungsgemeinschaft Grant FE 366/5-1 (to J.F.) and Grant AX82/2 (to N.A.), and UK Medical Research Council Program MC_A060_SPR10 (to R.N.H.).

- Henson R (2006) Forward inference using functional neuroimaging: Dissociations versus associations. *Trends Cogn Sci* 10(2):64–69.
- Scoville WB, Milner B (1957) Loss of recent memory after bilateral hippocampal lesions. *J Neurol Neurosurg Psychiatry* 20(1):11–21.
- Squire LR, Stark CEL, Clark RE (2004) The medial temporal lobe. *Annu Rev Neurosci* 27:279–306.
- Davachi L (2006) Item, context and relational episodic encoding in humans. *Curr Opin Neurobiol* 16(6):693–700.
- Diana RA, Yonelinas AP, Ranganath C (2007) Imaging recollection and familiarity in the medial temporal lobe: A three-component model. *Trends Cogn Sci* 11(9):379–386.
- Eichenbaum H, Yonelinas AP, Ranganath C (2007) The medial temporal lobe and recognition memory. *Annu Rev Neurosci* 30:123–152.
- Mayes A, Montaldi D, Migo E (2007) Associative memory and the medial temporal lobes. *Trends Cogn Sci* 11(3):126–135.
- Song Z, Wixted JT, Hopkins RO, Squire LR (2011) Impaired capacity for familiarity after hippocampal damage. *Proc Natl Acad Sci USA* 108(23):9655–9660.
- Song Z, Wixted JT, Smith CN, Squire LR (2011) Different nonlinear functions in hippocampus and perirhinal cortex relating functional MRI activity to memory strength. *Proc Natl Acad Sci USA* 108(14):5783–5788.
- Squire LR, Wixted JT, Clark RE (2007) Recognition memory and the medial temporal lobe: A new perspective. *Nat Rev Neurosci* 8(11):872–883.
- Bamber D (1979) State-trace analysis: A method of testing simple theories of causation. *J Math Psychol* 19(2):137–181.
- Loftus GR, Oberg MA, Dillon AM (2004) Linear theory, dimensional theory, and the face-inversion effect. *Psychol Rev* 111(4):835–863.
- Newell BR, Dunn JC (2008) Dimensions in data: Testing psychological models using state-trace analysis. *Trends Cogn Sci* 12(8):285–290.
- Prince M, Brown S, Heathcote A (2012) The design and analysis of state-trace experiments. *Psychol Methods* 17(1):78–99.
- Daselaar SM, Fleck MS, Cabeza R (2006) Triple dissociation in the medial temporal lobes: Recollection, familiarity, and novelty. *J Neurophysiol* 96(4):1902–1911.
- Gonsalves BD, Kahn I, Curran T, Norman KA, Wagner AD (2005) Memory strength and repetition suppression: Multimodal imaging of medial temporal cortical contributions to recognition. *Neuron* 47(5):751–761.
- Montaldi D, Spencer TJ, Roberts N, Mayes AR (2006) The neural system that mediates familiarity memory. *Hippocampus* 16(5):504–520.
- Yonelinas AP (2002) The nature of recollection and familiarity: A review of 30 years of research. *J Mem Lang* 46(3):441–517.
- Staresina BP, Fell J, Do Lam ATA, Axmacher N, Henson RN (2012) Memory signals are temporally dissociated in and across human hippocampus and perirhinal cortex. *Nat Neurosci* 15(8):1167–1173.
- Eldridge LL, Engel SA, Zeineh MM, Bookheimer SY, Knowlton BJ (2005) A dissociation of encoding and retrieval processes in the human hippocampus. *J Neurosci* 25(13):3280–3286.
- Fenker DB, Schott BH, Richardson-Klavehn A, Heinze HJ, Düzel E (2005) Recapitulating emotional context: Activity of amygdala, hippocampus and fusiform cortex during recollection and familiarity. *Eur J Neurosci* 21(7):1993–1999.
- Kirwan CB, Stark CEL (2004) Medial temporal lobe activation during encoding and retrieval of novel face-name pairs. *Hippocampus* 14(7):919–930.
- Weis S, et al. (2004) Process dissociation between contextual retrieval and item recognition. *Neuroreport* 15(18):2729–2733.
- Davachi L, Mitchell JP, Wagner AD (2003) Multiple routes to memory: Distinct medial temporal lobe processes build item and source memories. *Proc Natl Acad Sci USA* 100(4):2157–2162.
- Ranganath C, et al. (2004) Dissociable correlates of recollection and familiarity within the medial temporal lobes. *Neuropsychologia* 42(1):2–13.
- Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI (2009) Circular analysis in systems neuroscience: The dangers of double dipping. *Nat Neurosci* 12(5):535–540.
- Aggleton JP, Brown MW (1999) Episodic memory, amnesia, and the hippocampal-anterior thalamic axis. *Behav Brain Sci* 22(3):425–444, discussion 444–489.
- Wixted JT, Squire LR (2011) The medial temporal lobe and the attributes of memory. *Trends Cogn Sci* 15(5):210–217.

29. Brainerd CJ, Reyna VF, Howe ML (2009) Trichotomous processes in early memory development, aging, and neurocognitive impairment: A unified theory. *Psychol Rev* 116(4):783–832.
30. Dunn JC (2008) The dimensionality of the remember-know task: A state-trace analysis. *Psychol Rev* 115(2):426–446.
31. Newell BR, Dunn JC, Kalish M (2010) The dimensionality of perceptual category learning: A state-trace analysis. *Mem Cognit* 38(5):563–581.
32. Friston KJ, Mechelli A, Turner R, Price CJ (2000) Nonlinear responses in fMRI: The Balloon model, Volterra kernels, and other hemodynamics. *Neuroimage* 12(4):466–477.
33. Logothetis NK, Wandell BA (2004) Interpreting the BOLD signal. *Annu Rev Physiol* 66:735–769.
34. Büchel C, Holmes AP, Rees G, Friston KJ (1998) Characterizing stimulus-response functions using nonlinear regressors in parametric fMRI experiments. *Neuroimage* 8(2):140–148.
35. Dunn JC, Kirsner K (1988) Discovering functionally independent mental processes: The principle of reversed association. *Psychol Rev* 95(1):91–101.
36. Henson RN (2011) How to discover modules in mind and brain: The curse of non-linearity, and blessing of neuroimaging. A comment on Sternberg (2011). *Cogn Neuropsychol* 28(3-4):209–223.
37. Smith ME, Stapleton JM, Halgren E (1986) Human medial temporal lobe potentials evoked in memory and language tasks. *Electroencephalogr Clin Neurophysiol* 63(2):145–159.
38. Wagenmakers EJ, Ratcliff R, Gomez P, Iverson GJ (2004) Assessing model mimicry using the parametric bootstrap. *J Math Psychol* 48(1):28–50.