Information theory, novelty and hippocampal responses: unpredicted or unpredictable?

Bryan A. Strange[a,b,*], Andrew Duggins[a], William Penny[a], Raymond J. Dolan[a], Karl J. Friston[a]

[a]Wellcome Department of Imaging Neuroscience, Functional Imaging Laboratory, Institute of Neurology, 12 Queen Square, London WC1N 3BG, UK
[b]Institute of Cognitive Neuroscience, 17 Queen Square, London WC1N 3AR, UK

Abstract

Shannon's information theory provides a principled framework for the quantitative analysis of brain responses during the encoding and representation of event streams. In particular, entropy measures the expected uncertainty of events in a given context. This contextual uncertainty or unpredictability may, itself, be important for balancing [bottom-up] sensory information and [top-down] prior expectations during perceptual synthesis. Using event-related functional magnetic resonance imaging (fMRI), we found that the anterior hippocampus is sensitive to the entropy of a visual stimulus stream. In contrast, activity in an extensive bilateral cortico-thalamic network was dictated by the surprise or information associated with each particular stimulus. In short, we show that the probabilistic structure or context in which events occur is an important predictor of hippocampal activity.
© 2005 Elsevier Ltd. All rights reserved.

Current notions of representational learning and inference in the brain rest on an interplay between bottom-up sensory information and prior expectations, mediated by lateral and top-down influences (Friston, 2002; Hinton, Dayan, Frey, & Neal, 1995; Kawato, Hayakawa, & Inui, 1993; Mumford, 1992). Irrespective of the precise mechanisms employed by the brain, the relative weight afforded these two sources of information is a generic and important issue. For example, a specific neuronal mechanism has been proposed for balancing the evidence from sensory inputs and prior expectations according to their predictability (Yu & Dayan, 2002). We therefore addressed the hypothesis that the hippocampus is sensitive to the probabilistic context established by event streams. This sensitivity would enable the hippocampus, or related systems, to regulate the balance between top-down and bottom-up effects in sensory cortical hierarchies.

To formulate our hypothesis in a quantitative way we used information to measure the stimulus-bound 'surprise' of a particular event, and 'entropy' to measure the context in terms of the average predictability of a sequence (Jones, 1979; Shannon, 1948). We hoped to show that entropy could explain variations in hippocampal responses, even after accounting for responses induced by event-bound surprise. The distinction between entropy and surprise is critical. Surprise is unique to a particular event and measures its improbability (e.g. a small *p*-value is informative in classical inference in rejecting the null hypothesis). Conversely, entropy measures the expected or average surprise over all events, reflecting the predictability of an outcome before it occurs.

$$I(x_i) = -\ln p(x_i); \quad H(X) = \sum_i -p(x_i)\ln p(x_i) = \langle I(x_i) \rangle$$

Surprise $I(x_i)$ quantifies the information conveyed by the occurrence of event x_i, whereas entropy $H(X)$ quantifies the expected information of events sampled from X.

Hippocampal damage impairs episodic memory (Scoville & Milner, 1957) and it has been observed that hippocampal

* Corresponding author. Address: Functional Imaging Laboratory, Wellcome Department of Imaging Neuroscience, Institute of Neurology, 12 Queen Square, London WC1N 3BG, UK. Tel.: +44 207 833 7472; fax +44 207 813 1420.
E-mail address: bstrange@fil.ion.ucl.ac.uk (B.A. Strange).

activity reflects regularities embedded in an environment or task (Eichenbaum, Dudchenko, Wood, Shapiro, & Tanila, 1999). Information theory posits that these regularities depend upon the entropy. Using event-related functional MRI and carefully constructed sequences of visual stimuli, we tested the hypothesis that (i) hippocampal responses are sensitive to the entropy established by sequential events whereas (ii) responses in sensory areas reflect the surprise of each event. Areas that respond more vigorously to surprising stimuli can be regarded as exhibiting more prediction error from a predictive coding perspective (Rao & Ballard, 1999).

As entropy is effectively the running average surprise (see below), we were able to manipulate entropy and surprise independently and dissociate their neurophysiological correlates. Briefly, we varied the entropy of the stimuli over blocks of trials. This allowed us to look for brain responses that covaried with entropy over blocks. Because each block comprised likely and unlikely events we were also able to assess correlates of surprise within-block. The analysis proceeded by modeling entropy and surprise-related responses in a subject-specific first-level analysis. The results of this analysis were taken to a second-level analysis to implement random effects inferences about these responses over subjects.

During fMRI, 12 subjects were presented with 12 blocks of 40 trials. Each trial comprised a brief presentation of a colored shape (stimulus duration: 500 ms; stimulus onset asynchrony: 2.2 s). In all trials within a block, two colors and two shapes were combined to form four possible outcomes, with different stimuli presented in the different blocks. The stimuli appeared in miniature for 5 s before the beginning of a block and remained in a row at the bottom of the screen throughout the block. Subjects were required to respond to the sampled item by pressing a key to identify the target's position in the row. A schematic of a trial is shown in Fig. 1a.

Each trial used an independent sample from a distribution that remained constant within a block, but that varied over blocks. Note that there was no underlying sequence governing stimulus presentation, only the relative proportions of stimuli were varied from block to block. Subjects were asked to consider each block as a 'hat' containing a large number of objects of four distinct types with each hat containing a different set. They were also told that each trial would be equivalent to sampling an object and then returning it to the hat. Subjects were informed that the proportion of objects in a particular hat was completely unpredictable, and independent of the other hats sampled.

From the point of view of the subject, the suprise of each trial depended on the history of previous trials within a block. In these circumstances allocation of neuronal resources is based on the subjective probabilities inherent in the modeling process, not on the objective frequencies of the events (Sinkkonnen, 2000). The information or 'surprise' inherent in an event is based on the probability of that event. Clearly for a system to encode events efficiently it must know these probabilities or infer them on the basis of experience (i.e. how frequently they occur). The brain may embody these probabilistic regularities through plastic changes in development (e.g. language) or in a context-specific form (e.g. the 'oddball' paradigm). Our study represents an example of the latter. To compute surprise and entropy we treated each subject as an ideal observer and used the Bayesian posterior probability of an event, given the history of trials within a block. Entropy and surprise were calculated from posterior or conditional probabilities on a trial-by-trial basis and used to predict neuronal responses. These metrics are 'inferred' on the basis
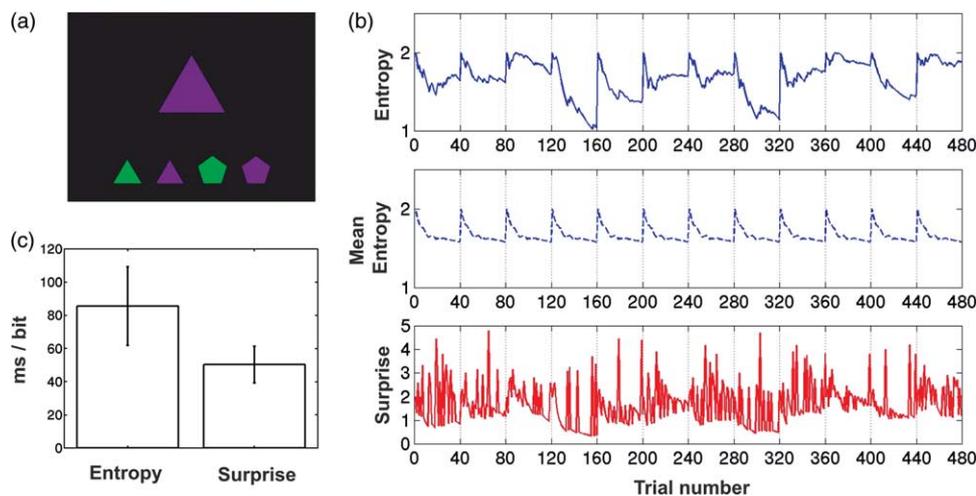
Fig. 1. Task design. (a) The choice reaction time paradigm. Subjects were required to respond to the sampled item (in this case a purple triangle) by pressing the key to indicate the position of that item in the row of alternative coloured shapes (below). (b) Information theoretic quantities for a typical scanning session. Entropy, mean entropy (averaged across all blocks) and surprise are plotted (units: bits). Dotted vertical lines divide successive blocks. A fixation cross was presented for 20 s between blocks. (c) Behavioral data. Increase in reaction time per bit of entropy and surprise ($\pm$SE of the mean of 12 subjects). Mean RT across all trials: 550.2 ($\pm$15.0) ms.

of exposure to stimulus streams and change systematically as more samples are seen (a typical experimental session is plotted in Fig. 1b).

1. Materials and methods

1.1. Subjects

Informed consent was obtained from 12 right-handed subjects (7 male; age range 19–39 years; mean age 27.2). Ethics approval was obtained from the local Ethics Committee.

1.2. Bayesian learning

For the four outcomes, where $j = 1$–4, we denote the total number of occurrences of outcome j up to observation i by n_j^i. Our estimate of the probability of outcome j in the next observation is then

$$p(x_i) = \frac{n_j^i + 1}{\sum_k n_k^i + 1}$$

where k is for all four outcomes, and hence the lower term of the equation is equivalent to the number of completed (past) trials in the block plus one. This estimate is the mean value of a 'Dirichlet' distribution whose parameters have been updated using Bayesian learning (Bernardo & Smith, 1994). The prior distribution, i.e. before any outcomes have been observed, is a Dirichlet with parameters identically equal to 1 (hence the 1 s in the above equation). These prior values act as 'pseudo-counts' such that before seeing any outcomes we imagine having seen 1 of each (the value 1 is chosen to give the weakest influence of the prior). Hence, the Dirichlet parameters characterise the knowledge of an ideal observer of the sequence.

1.3. Imaging

A 2T Siemens VISION system (Siemens, Erlangen) was used to acquire T1-weighted anatomical images and gradient-echo echo-planar T2*-weighted MRI image volumes with blood oxygenation level dependent (BOLD) contrast. For each subject, 552 volumes were acquired plus 6 'dummy' volumes (to allow for T1 equilibration effects). Volumes were acquired continuously every 2506 ms. Each volume comprised 33 3.3 mm axial slices, with an in-plane resolution of 3×3 mm, positioned to cover the cerebrum. The imaging time series was realigned, slice-time corrected, normalized into standard Talaraich and Tournoux anatomical space, and smoothed with a Gaussian kernel of 6 mm full width half maximum as described previously (Friston et al., 1995a).

Imaging data were analyzed using Statistical Parametric Mapping (SPM99) employing an event-related model with a two-stage random effects procedure. Trial-specific responses were modeled by convolving a 'stick' function, modeling trial onsets, with a canonical hemodynamic response function (HRF). To characterize responses in terms of information theoretic measures, we specified two parametric modulators of the trial-specific response: entropy and surprise. The surprise covariate modeled the event-related response components explained by the information of the event that could not be explained by the entropy, and *vice versa*, by virtue of the orthogonalisation implicit to the general linear model. For each subject, the mean entropy for the *i*th trial over blocks was included as a nuisance covariate to model non-specific time effects within blocks (Fig. 1b). Other nuisance variables included residual reaction times (that could not be explained by the information and entropy), response errors, low frequency drifts in signal (cut-off 62 s) and movement parameters, determined during realignment. A 5-second cue period preceding each block was also modeled. Subject-specific parameter estimates pertaining to the height of the HRF for each regressor were calculated for each voxel (Friston et al., 1995b). In the second stage of the random effects analysis, each subject's parameter estimate image for surprise and entropy were entered into two one-sample *t*-tests across the 12 subjects. We report medial temporal responses to entropy at a threshold of $p < 0.05$, corrected for the volume of hippocampus searched (a sphere of 16 mm diameter centred on x, y, z co-ordinates -26, -10, -14). The term hippocampus is used here to refer to dentate gyrus, CA subfields and subiculum. Surprise effects are reported at $p < 0.05$ corrected for search volumes (thalamus: two spheres of 30 mm diameter centred on -8, -16, 10 and 10, -12, 8; posterior fusiform: two spheres of 30 mm diameter centred on -36, -72, -22 and 28, -76, -16; see Table 1). A similar two-stage random effects procedure was also employed in the analysis of reaction time data (in the reaction time General Linear Model the only nuisance variable modeled response errors).

Table 1
Functional double dissociation between cortico-thalamic coding of surprise and hippocampal sensitivity to entropy

Brain region	x, y, z co-ordinates	Z score surprise	Z score entropy
L hippocampus	-26, -10, -14	<1.6 ns	3.32*
L thalamus	-8, -16, 10	4.74*	<1.6 ns
R thalamus	10, -12, 8	4.41*	<1.6 ns
L posterior fusiform cortex	-36, -72, -22	4.87*	<1.6 ns
R posterior fusiform cortex	28, -76, -16	4.18*	<1.6 ns

Posterior fusiform activity is presented as an example of cortical responses.
*$p < 0.05$ corrected for search volume (see Section 1). ns, not significant.

2. Results

Consistent with previous observations (Hick, 1952; Hyman, 1953), entropy and surprise both modulated reaction times significantly (Fig. 1c). Fig. 2 demonstrates that activation in left anterior hippocampus was modulated by the entropy of the sampling distribution. Greater activation was observed for unpredictable stimulus streams, i.e. blocks with more even sampling distributions, evident in the plot of response estimates. The parameter estimates also demonstrate that the anterior hippocampus did not show a significant mean response to all stimuli, nor was the response modulated by surprise (see Table 1). Activation in right anterior hippocampus was also observed in relation to entropy, but failed to reach corrected significance. Similar results have been observed during a sequence learning task where greater anterior hippocampal activation was evoked by random (high entropy) than by ordered (low entropy) sequences (Katz Sakai and Richard Passingham, personal communication).

In contrast to hippocampal sensitivity to entropy, surprise modulated responses in an extensive bilateral network (Fig. 3). Surprise-related activations were observed in bilateral fusiform, parietal, premotor and inferior frontal cortices, as well as in bilateral thalamus (Fig. 3). Activation in these regions conforms to efficient coding of both visual stimuli and subsequent motor responses (i.e. less activation with more frequent events, or more prediction error with less frequent events). This network is co-extensive with the classical visual attention network (Mesulam, 1990; Rees & Lavie, 2001). Thus, our findings demonstrate a functional double dissociation between regions encoding entropy and surprise (Table 1). Hippocampal activity was determined by the predictability of events before they occurred, whereas activation in a cortico-thalamic network was dictated by the probability of each event.

Fig. 2. Anterior hippocampus is sensitive to the entropy of a sequence. (a) The SPM (threshold $p < 0.005$ for illustration; extent threshold 5 voxels) is overlaid on sagittal ($x = -26$), coronal ($y = -10$) and transverse ($z = -14$) sections of the T1 MNI reference brain to demonstrate activation in left anterior hippocampus (x, y, z co-ordinates $-26, -10, -14$; $Z = 3.32$; $p < 0.05$ corrected). The parameter estimates ($\pm$SE) for the extent to which the height of the modeled HRF in left anterior hippocampus is modulated by entropy (E) and surprise (S) are plotted, as well as the parameter estimate for mean response to all trials (Epoch). Weaker activation in right anterior hippocampus ($36, -14, -20$; $Z = 2.72$; $p < 0.005$ uncorrected) is evident in transverse section. The colored bar indicates the T statistic of the activation. (b) Anterior hippocampal activation in two subjects (s1, s2; single-subject SPM threshold $p < 0.001$) is overlaid on sagittal and coronal sections of that subject's mean functional T2* image (colour contrast of T2* images inverted for illustration; position of sections indicated by the white rectangles in (a). These sections indicate that entropy-evoked activation is in anterior hippocampus and not surrounding medial temporal cortices.

Fig. 3. Efficient encoding of visual stimuli in cortical and subcortical structures. Surface rendering of regional cortical responses that covaried with surprise (SPM threshold $p < 0.005$ uncorrected; extent threshold 5 voxels). Responses in left (L) and right (R) posterior fusiform (Fus) and premotor (PreM) cortices and intraparietal sulcus (IPS) are indicated ($\rightarrow$). Below, the same SPM is overlaid on a coronal section of the T1 reference brain ($y = -12$) to demonstrate bilateral activation in the thalamus (Th; see Table 1). The extent to which responses in these regions are modulated by surprise are plotted ($\pm$SE) along with the response to surprise in the left anterior hippocampal (H) region shown in Fig. 2a.

3. Discussion

Our study addresses the role of the hippocampus from a particular (information theoretic) perspective that is closely related to novelty detection. Functional neuroimaging studies have consistently demonstrated novelty-dependent activation of anterior hippocampus (Dolan & Fletcher, 1997; Strange, Fletcher, Henson, Friston, & Dolan, 1999; Tulving, Markowitsch, Craik, Habib, & Houle, 1996). More recent evidence, however, suggests that the hippocampal role in novelty detection reflects a more general role in detecting mismatches between expectation and experience (Hasselmo, Wyble, & Wallenstein, 1996; Ploghaus et al., 2000; Strange & Dolan, 2001). Hippocampal sensitivity to improbable events evoking 'mismatch' (i.e. oddballs (Halgren et al., 1980)) would suggest augmented responses to surprising events. This was not observed (Table 1), nor have fMRI studies of oddball responses demonstrated hippocampal activation (Downar, Crawley, Mikulis, & Davis, 2000; McCarthy, Luby, Gore, & Goldman-Rakic, 1997; but see Strange & Dolan, 2001). Hippocampal sensitivity to entropy, however, suggests that a hippocampal role in processing improbable events may reflect *learning that oddballs can occur*, as opposed to the encoding of their actual occurrence. This proposal is consistent with our previous finding (Strange & Dolan, 2001) that oddball-evoked activation in anterior hippocampus *adapts* with repeated presentations of oddballs. In a previous fMRI study (Bischoff-Grethe, Proper, Mao, Daniels, & Berns, 2000) that manipulated entropy in a blocked design, the anterior medial temporal region was not included in the field of view. This study demonstrated responses in Wernicke's area that covaried with entropy, a finding not reproduced in our study.

Our current results suggest that hippocampal processing may reflect a more generic context-sensitivity to the stochastic or probabilistic structure of observed events. The key point is that the hippocampus represents the *expected* information or novelty of any event before it occurs. This is in contradistinction to other brain areas, usually at lower hierarchical levels, that encode the novelty per se (Brown & Xiang, 1998; Miller, Li, & Desimone, 1991). Representing the entropy or predictability of sensory input may play a fundamental role in perceptual synthesis. If one accepts that neuronal representations are constructed on the basis of bottom-up evidence from early sensory cortices and top-down influences that mediate prior expectations (c.f. predictive coding (Rao & Ballard, 1999) and related schemes using generative models), then the relative balance of bottom-up and top-down effects depends upon the veracity of the priors (Courtney & Ungerleider, 1997). In situations with high entropy, in which inputs are unpredictable, more weight should be given to the bottom-up forward connections. It is possible that this balance is mediated by modulation of synaptic connections by systems that are sensitive to the entropy. The hippocampal formation has the necessary supraordinate position in cortical sensory hierarchies and the requisite reciprocal connectivity to play this role.

The precise relationship between this putative hippocampal role in perceptual synthesis and its critical role in episodic memory remains to be determined. However, a deficit in balancing top-down priors and bottom-up effects may explain hippocampal lesion-induced deficits in any task in which the probabilistic structure of the environment (i.e. entropy) is manipulated experimentally (e.g. the Morris water maze task (Morris, Garrud, Rawlins, & O'Keefe, 1982) or transitivity (Eichenbaum et al., 1999)). Importantly, several forms of non-declarative learning (e.g. sequence learning, delay classical conditioning, probabilistic learning (Stadler & Frensch, 1998)), which are intact following hippocampal lesions, involve learning regularities in the context of low entropy (e.g. the US–CS relationship in conditioning is, of course, highly predictable).

The model used to explain our neurophysiological observations was based upon an ideal Bayesian observer. This enabled some interesting inferences about how probability and uncertainty are encoded in the brain. However, such a simple model precludes any comment about whether the probability learning was explicit or implicit and, furthermore, does not address motivational factors or the influence of cognitive set. These sorts of issues may be usefully addressed using paradigms similar to the one employed here in the context of reward learning.

Functional neuroimaging studies have generally used psychological manipulations, such as encoding and retrieval tasks, to explore the hippocampal role in episodic memory (Lepage, Habib, & Tulving, 1998; Schacter & Wagner, 1999). The approach adopted here differs in that we employed functional imaging to test quantitative, information theoretic predictions that speak to a computational role of the hippocampus. Our data suggest that hippocampal neuronal activity is dictated by the probabilistic structure of the environment with activity in this region representing the *expected* information or novelty of an event before it occurs.

References

Bernardo, J. M., & Smith, A. F. M. (1994). *Bayesian theory*. New York: Wiley.

Bischoff-Grethe, A., Proper, S. M., Mao, H., Daniels, K. A., & Berns, G. S. (2000). Conscious and unconscious processing of nonverbal predictability in Wernicke's area. *Journal of Neuroscience*, 20, 1975–1981.

Brown, M. W., & Xiang, J.-Z. (1998). Recognition memory: Neuronal substrates of the judgement of prior occurrence. *Progress in Neurobiology*, *55*, 149–189.

Courtney, S. M., & Ungerleider, L. G. (1997). What has fMRI taught us about human vision? *Current Opinion in Neurobiology*, *7*, 554–561.

Dolan, R. J., & Fletcher, P. C. (1997). Dissociating prefrontal and hippocampal function in episodic memory encoding. *Nature*, *388*, 582–585.

Downar, J., Crawley, A. P., Mikulis, D. J., & Davis, K. D. (2000). A multimodal cortical network for detecting changes in the sensory environment. *Nature Neuroscience*, *3*, 277–283.

Eichenbaum, H., Dudchenko, P., Wood, E., Shapiro, M., & Tanila, H. (1999). The hippocampus, memory, and place cells: Is it spatial memory or a memory space? *Neuron*, *23*, 209–226.

Friston, K. J. (2002). Functional integration and inference in the brain. *Progress in Neurobiology*, *68*, 113–143.

Friston, K. J., Ashburner, J., Frith, C. D., Poline, J.-B., Heather, J. D., & Frackowiak, R. S. J. (1995a). Spatial registration and normalisation of images. *Human Brain Mapping*, *2*, 165–189.

Friston, K. J., Homes, A. P., Worsely, K. J., Poline, J.-B., Frith, C. D., & Frackowiak, R. S. J. (1995b). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, *2*, 189–210.

Halgren, E., Squires, N. K., Wilson, C. L., Rohrbaugh, J. W., Babb, T. L., & Crandall, P. H. (1980). Endogenous potentials generated in the human hippocampal formation and amygdala by infrequent events. *Science*, *210*, 803–805.

Hasselmo, M. E., Wyble, B. P., & Wallenstein, G. V. (1996). Encoding and retrieval of episodic memories: Role of cholinergic and GABAergic modulation in the hippocampus. *Hippocampus*, *6*, 693–708.

Hick, W. E. (1952). On the rate of gain of information. *Quarternary Journal of Experimental Psychology*, *4*, 11–26.

Hinton, G. E., Dayan, P., Frey, B. J., & Neal, R. M. (1995). The "wake-sleep" algorithm for unsupervised neural networks. *Science*, *268*, 1158–1161.

Hyman, R. (1953). Stimulus information as a determinant of reaction time. *Journal of Experimental Psychology*, *45*, 175–182.

Jones, D. S. (1979). *Elementary information theory*. Oxford: Oxford University Press.

Kawato, M., Hayakawa, H., & Inui, T. (1993). A forward-inverse optics model of reciprocal connections between visual areas. *Network: Computation in Neural Systems*, *4*, 415–422.

Lepage, M., Habib, R., & Tulving, E. (1998). Hippocampal PET activations of memory encoding and retrieval: The HIPER model. *Hippocampus*, *8*, 313–322.

McCarthy, G., Luby, M., Gore, J., & Goldman-Rakic, P. (1997). Infrequent events transiently activate human prefrontal and parietal cortex as measured by functional MRI. *Journal of Neurophysiology*, *77*, 1630–1634.

Mesulam, M. M. (1990). Large-scale neurocognitive networks and distributed processing for attention, language, and memory. *Annals of Neurology*, *28*, 597–613.

Miller, E. K., Li, L., & Desimone, R. (1991). A neural mechanism for working and recognition memory in inferior temporal cortex. *Science*, *254*, 1377–1379.

Morris, R. G. M., Garrud, P., Rawlins, J. P., & O'Keefe, J. (1982). Place navigation impaired in rats with hippocampal lesions. *Nature*, *297*, 681–683.

Mumford, D. (1992). On the computational architecture of the neocortex. II. The role of cortico-cortical loops. *Biological Cybernetics*, *66*, 241–251.

Ploghaus, A., Tracey, I., Clare, S., Gati, J. S., Rawlins, J. N. P., & Matthews, P. M. (2000). Learning about pain: the neural substrate of the prediction error for aversive events. *Proceedings of the National Academy of Science USA*, *97*, 9281–9286.

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*, 79–87.

Rees, G., & Lavie, N. (2001). What can functional imaging reveal about the role of attention in visual awareness? *Neuropsychologia*, *39*, 1343–1353.

Schacter, D. L., & Wagner, A. D. (1999). Medial temporal lobe activations in fMRI and PET studies of episodic encoding and retrieval. *Hippocampus*, *9*, 7–24.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurosurgery and Psychiatry*, *20*, 11–21.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*, 379–423.

Sinkkonnen, J. (2000). Information and resource allocation. In R. Baddeley, P. Hancock, & P. Földiak (Eds.), *Information theory and the brain* (pp. 241–253). Cambridge, UK: Cambridge University Press, 241–253.

Stadler, M. A., & Frensch, P. A. (Eds.). (1998). *Handbook of implicit learning*. Thousand Oaks, CA: Sage.

Strange, B. A., & Dolan, R. J. (2001). Adaptive anterior hippocampal responses to oddball stimuli. *Hippocampus*, *11*, 690–698.

Strange, B. A., Fletcher, P. C., Henson, R. N. A., Friston, K. J., & Dolan, R. J. (1999). Segregating the functions of human hippocampus. *Proceedings of the National Academy of Science USA*, *96*, 4034–4039.

Tulving, E., Markowitsch, M. J., Craik, F. I. M., Habib, R., & Houle, S. (1996). Novelty and familiarity activations in PET studies of memory encoding and retrieval. *Cerebral Cortex*, *6*, 71–79.

Yu, A. J., & Dayan, P. (2002). Acetylcholine in cortical inference. *Neural Networks*, *15*, 719–730.