The Context of Uncertainty Modulates the Subcortical Response to Predictability

Amanda Bischoff-Grethe[1], Megan Martin[2], Hui Mao[3], and Gregory S. Berns[2,4]

Abstract

■ Implicit motor learning tasks typically involve comparisons of subject responses during a sequence versus a random condition. In neuroimaging, brain regions that are correlated with a sequence are described, but the temporal relationship of sequence versus nonsequence conditions is often not explored. We present a functional magnetic resonance imaging (fMRI) study describing activation related to sequential predictability in an implicit sensorimotor learning task and the history (context) dependence of these effects. Participants regarded four squares displayed horizontally across a screen and pressed a button when any one of the four targets was illuminated in a particular color. A repeating spatial sequence with varying levels of predictability was embedded within a random color presentation. Both the right dorsolateral prefrontal cortex (R DLPFC) and right caudate displayed a positive correlation to increasing predictability, whereas the left posterior parietal cortex (L PPC) displayed a negative correlation. However, the activation changes within the caudate were significant when transitioning from high predictability to low predictability but not for the reverse case, suggesting a sensitivity not only to predictability but to order effects as well. These results support the hypothesized relationship between basal ganglia and visuomotor sequential learning, but demonstrate the importance of context upon sequence learning. ■

INTRODUCTION

The prediction of future events is a problem faced by nearly every organism and occurs across a wide variety of contexts. Some processes clearly are more predictable than others, and an assessment of the inherent predictability of a given process may have significant relevance to an individual. However, the assessment of predictability may be significantly influenced by the context in which it occurs. For example, in a highly predictable environment the appearance of randomness might be quite noticeable, but in a noisy environment the appearance of order may not be as obvious. In this article, we describe the human neural circuitry associated with monitoring temporal predictability on an implicit visuomotor sequence task and the effect of the context of uncertainty.

Predictability, or its converse, entropy, can be quantified by global measures of information transmission (Cover & Thomas, 1991; Shannon & Weaver, 1949). Fundamentally, these rely on the probabilities with which events occur. A simple way of controlling temporal predictability is to design an artificial grammar, which is simply a set of rules governing the probability of transition from one state to another (Cleeremans & McClelland, 1991; Cohen, Ivry, & Keele, 1990; Stadler, 1989; Reber, 1967, 1993). In linguistics, a state might represent a word, with complex chains of transition probabilities capturing the statistics of the underlying grammatical rules. This type of statistical description is not limited to linguistics but can be applied to any system with discrete states, e.g., chemical reactions. More generally, these grammars, or Markov chains, can be used to generate temporal sequences with precisely defined statistics. By varying only the transition probabilities between states, the statistics can be changed without altering the underlying rules. More importantly, an overall measure of statistical uncertainty, and therefore predictability, can be defined by the entropy of the sequence.

Sequential, or temporal, predictability has been shown to appear in a variety of contexts, including language complexity (Just, Carpenter, Keller, Eddy, & Thulborn, 1996) and implicit motor learning (Hazeltine, Grafton, & Ivry, 1997; Grafton, Hazeltine, & Ivry, 1995). In a previous functional magnetic resonance imaging (fMRI) study, we measured the effect of temporal predictability in a nonverbal, nonmotor task (Bischoff-Grethe, Proper, Mao, Daniels, & Berns, 2000). Subjects were presented with four targets placed horizontally across a screen and which individually lit in one of three randomly selected colors. The subjects were instructed to silently count the occurrence of blue squares. Embedded within the paradigm was a visuospatial sequence

[1] Dartmouth College, [2] Dept. of Psychiatry and Behavioral Sciences, Emory University School of Medicine, [3] Dept. of Radiology, Emory University School of Medicine, [4] Georgia Institute of Technology

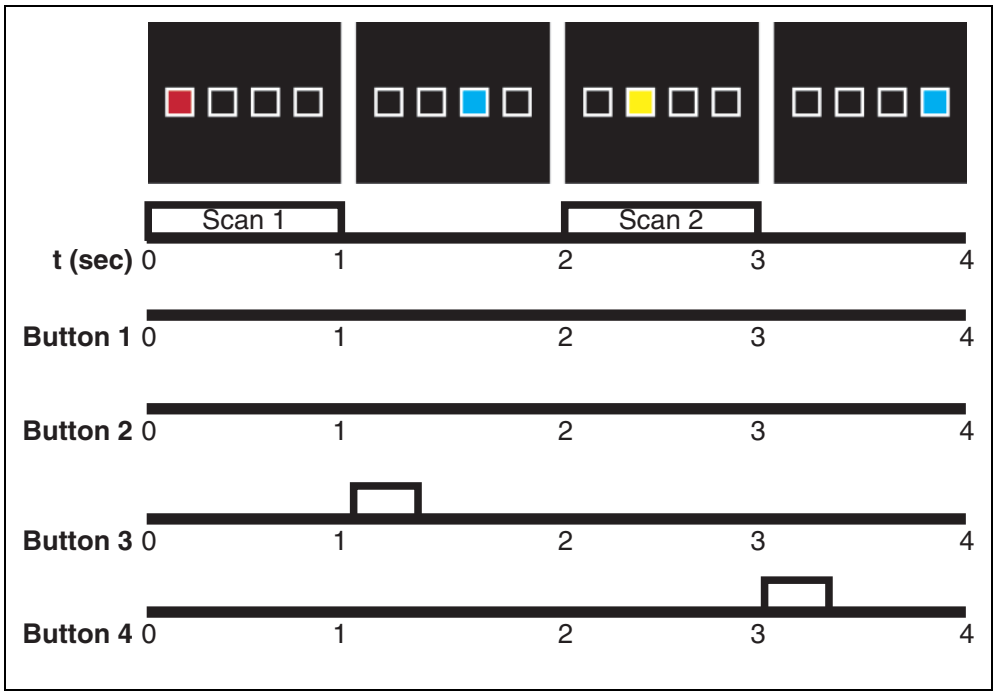

Figure 1. Experimental design of the presented task. Four squares were presented horizontally across a screen. The squares were individually illuminated in one of three randomly selected colors (red, blue, or yellow) for 1 sec. Participants were given a button box and, using the fingers of the right hand, pressed the key corresponding to the position of the blue box when it appeared on the display. The sequence of spatial positions (regardless of square color) was determined by an artificial grammar with varying predictability between condition blocks. The given example shows 4 sec of a scan where the spatial sequence was presented (1–3–2–4). The occurrence of the functional scans and the subject's performance with the button box are also shown.

that appeared within several of the experimental blocks. We found that Wernicke's area was highly correlated with the inherent predictability of the visuospatial sequence, independently of whether participants were aware of the sequence or not.

This study raised two important questions: (1) Are the effects of predictability specific to Wernicke's area? Or, are they modality dependent? (2) Are the effects of predictability symmetric with respect to the order in which it is changed? To answer these questions, we performed a motor version of this fMRI experiment with particular attention to the effect of the order in which the conditions were presented. Because this was a motor task, we expected predictability to modulate activity in the basal ganglia. The participants regarded a screen upon which four squares were displayed horizontally (Figure 1). The squares individually illuminated for 1 sec in one of three randomly selected colors (red, blue, or yellow). The participants were given a button box and were instructed to press the button corresponding to the position of a box only if the box were illuminated in blue. The spatial order in which the boxes were illuminated was determined by one of three conditions: (1) a four-element repeating spatial sequence, 1–3–2–4 (high predictability condition); (2) a probabilistic version of the sequence, based upon a Markov chain; and (3) a randomly ordered presentation (low predictability condition). Participants were either presented with the high predictability condition followed by the low predictability condition during the first functional imaging run (the HI–LO group), or they were first presented with the low predictability condition followed by the high predictability condition (the LO–HI group). During a second imaging run, the groups were presented with the respective converse order. This allowed us to examine order effects both within and across the two subject groups. The participants were not told of the existence of a spatial sequence, nor did they report awareness of one when debriefed after the second imaging session. Given the similarity to our previous task, we expected the change of modality for the response would still produce similar results. Further,

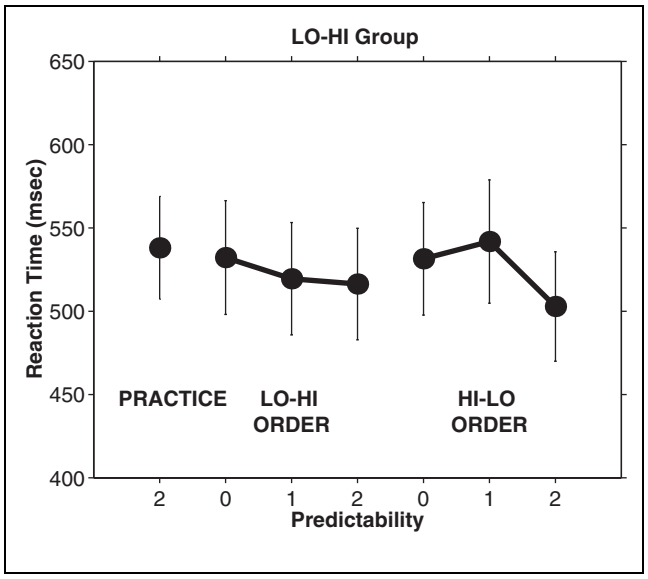

Figure 2. The RTs for the LO–HI group across predictability levels and session blocks. Participants performed the high predictability condition (predictability = 2) during the practice block, and one block of each predictability level during each functional scanning session. In the first session the subjects were presented with the LO–HI order (1–0–2); the converse (HI–LO, 2–0–1) was presented in the second session.

the current study had the advantage of collecting whole brain images, allowing us to examine activation patterns within subcortical regions.

RESULTS

Reaction times (RTs) were collected only for participants in the LO–HI group due to technical difficulties (Figure 2). During the practice session using the high predictability condition (predictability = 2), LO–HI participants performed with an average RT of 538 msec (*SE* = 30.89, 90.6% accuracy). During the first imaging session (when order LO–HI was presented) participants averaged an RT of 516.3 msec (*SE* = 33.51) during the high predictability condition, and 532.2 msec (*SE* = 34.17) during the low predictability condition (98.2% overall accuracy). During the second imaging session (order HI–LO), participants averaged an RT of 502.8 msec (*SE* = 32.86) during the

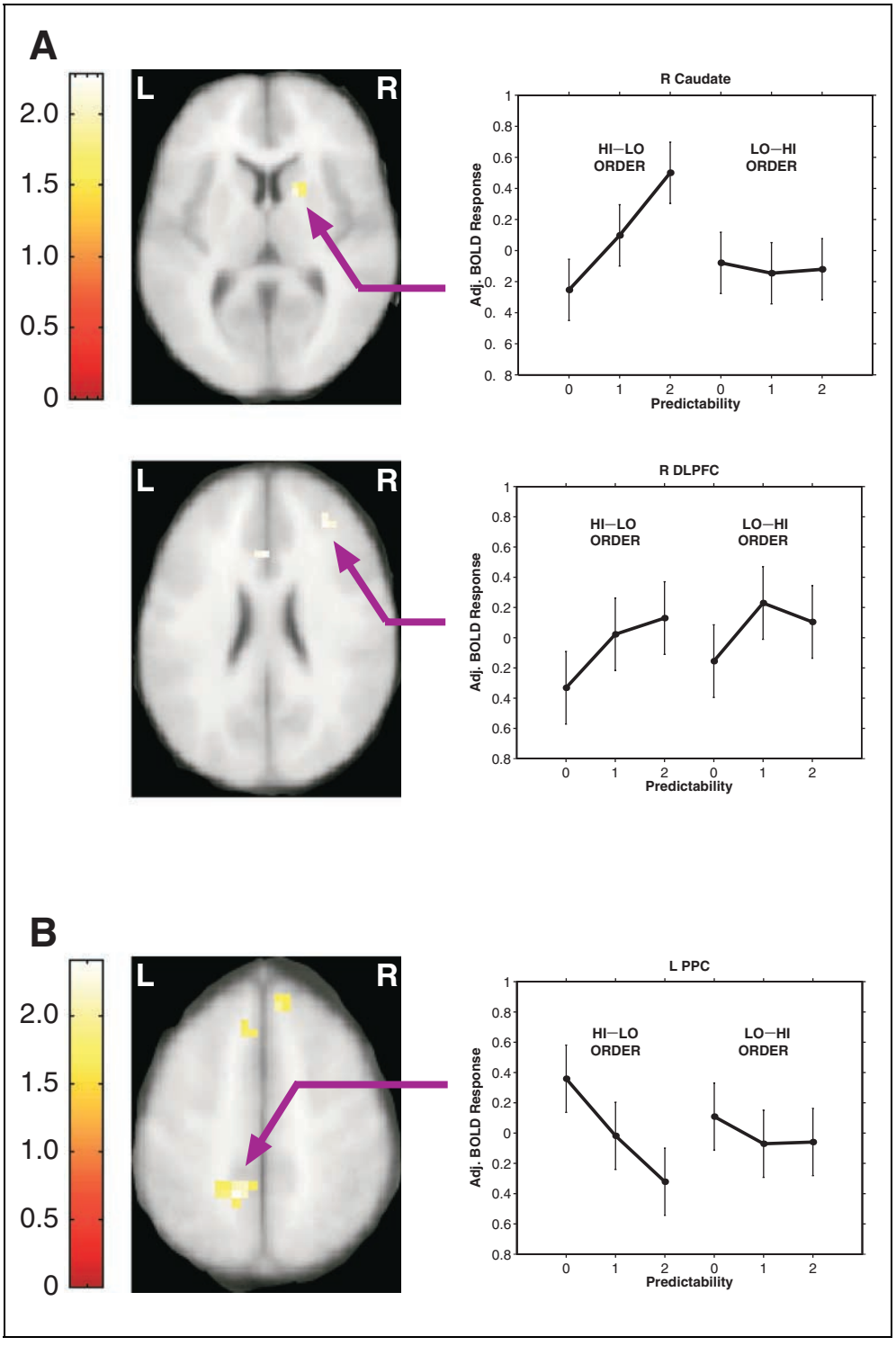

Figure 3. The statistical parametric maps (SPMs) illustrating regions correlated with predictability. Transverse planes (Talairach, $Z = +4$, $Z = +24$, and $Z = +52$) are shown. (A) Correlation of predictability in right caudate (top row) and R DLPFC (middle row) with a decreasing contrast [1 0 −1 1 0 −1] applied. (B) Correlation of predictability in L PPC with an increasing contrast [−1 0 1 −1 0 1] applied. Graphs to right of SPMs illustrate the relationship between the given brain region and predictability transitions, averaged across the two experiment groups. Both the caudate and prefrontal cortex increased activation with increasing predictability, whereas the parietal cortex increased activation with decreasing predictability.

Table 1. Regional Activations Associated with Comparisons of Task Performance in a Random Effects Design

Region	*Cluster (voxels)*	*MNI Coordinates (x, y, z)*			*Z Score*
Positive Correlation to Predictability					
Right caudate	9*	16	8	4	2.28
Left anterior cingulate (BA 24)	5	0	24	20	1.91
Right dorsolateral prefrontal cortex (BA 46)	5*	32	40	24	1.90
Negative Correlation to Predictability					
Left posterior parietal cortex (BA 5/7)	20*	−12	−48	52	2.41
Right superior frontal cortex (BA 8)	10	8	36	52	2.05
Left superior frontal cortex (BA 8)	9	−4	24	56	1.88

Coordinates are from the MNI template and use the same orientation and origin as found in the Talairach atlas (Talairach & Tournoux, 1988). The table shows local maxima > 8.00 mm apart per cluster ($k \geq 5$).

*$p < .05$ corrected for multiple comparisons in a sphere of 10–14 mm diameter.

high predictability condition and 531.5 msec (*SE* = 33.80) during the low predictability condition (97.8% overall accuracy). While the RT for the low predictability condition was greater than that for the high predictability condition for both imaging sessions, this difference was more apparent in the HI–LO order.

The imaging results identified a restricted number of regions that were correlated with sequential predictability (Figure 3, Table 1). Regions with a positive correlation to predictability included the right caudate, left anterior cingulate, and right dorsolateral prefrontal cortex (R DLPFC). Of these regions, only the right caudate showed a relationship to predictability that was dependent on the block order. Transitioning from a high- to low-predictability condition resulted in a stronger correlation. This indicated a temporal hysteresis effect: The order in which the conditions were presented had a nonlinear relationship with the BOLD response, such that the order of transitions between conditions also contributed to the changes in activation. Regions with a negative correlation to predictability included the left posterior parietal cortex (L PPC) and bilateral superior frontal cortices, with the L PPC displaying a slightly stronger correlation in the HI–LO order (Figure 3B).

DISCUSSION

The results of our study indicate that information itself, as embodied by the temporal predictability of sequential stimuli, is a parameter that can be used to probe specific neural circuits. Because the nature of fMRI necessitates relative measurements, arbitrary subtractions of cognitive events can be misleading. By making a simple yet subtle manipulation in sequential predictability, we were able to isolate activity in a restricted cortical–subcortical network. Our results suggest that information itself is a form of neural currency, and by manipulating the information content of a particular modality stream, the function of this circuit can be better understood.

In general, we find further evidence for the role of the right prefrontal–caudate circuit in spatial sequence learning. Previous imaging studies have identified these regions as being involved in implicit learning on serial reaction time tasks (SRT) (Berns, Cohen, & Mintun, 1997; Rauch et al., 1997; Grafton et al., 1995). Because our paradigm was not strictly an SRT task, our results suggest both a more generic and specific function of the right prefrontal–caudate circuit. Previous imaging studies of the SRT task have compared the activation of sequence against random conditions (Rauch et al., 1997; Grafton et al., 1995). While this is suggestive of implicit learning because subjects get faster, it does not preclude the possibility that these regions are simply responsive to the presence or absence of information. Our results extend these findings by showing a generally monotonic relationship of activity to information content. Furthermore, for the caudate in particular, the context in which this information occurs may have a strong modulating effect. This result echoes our previous finding of increased right caudate activation in an SRT task when the sequence was changed, but maintaining the same overall predictability (Berns et al., 1997). It is also consistent with recent findings of context-dependent activity in the globus pallidus (Gdowski, Miller, Parrish, Nenonene, & Houk, 2001).

The head of the caudate is part of a cognitive circuit that includes both DLPFC and PPC (Goldman-Rakic, 1996; Passingham, 1993). Both DLPFC and PPC have been closely linked to conscious spatial attention (Courtney, Petit, Maisog, Ungerleider, & Haxby, 1998; Goldman-Rakic, 1996; McCarthy et al., 1994), but none of the participants in our study were given any explicit

information that a spatial sequence existed, nor did they show evidence of explicit knowledge when debriefed afterward. There is strong evidence that DLPFC is responsive during spatial working memory tasks. Previously, PET has shown learning-related blood flow increases in R DLPFC (BA 46) with exposure to a spatial sequence (Grafton et al., 1995). This activation was not seen when the sequence learning task relied upon color-coded stimuli presented at a central location (Hazeltine et al., 1997), nor when a distractor (tone counting) was employed (Hazeltine et al., 1997; Grafton et al., 1995). However, DLPFC has been shown to modulate its activation in implicit sequencing tasks (Berns et al., 1997). It is possible that the modulation of R DLPFC activity seen in our task was related to working memory when the sequence was presented intact.

It has been suggested that unambiguous sequences can be learned without attention, whereas those sequences that are context dependent require attention to be learned implicitly (Curran & Keele, 1993). Our results lend biological credence to this hypothesis. Although our study involved implicit learning of an unambiguous sequence (1–3–2–4), it was rendered ambiguous by virtue of being probabilistic at times. When the sequence was changed to either a probabilistic one or an entirely random one, activity increased in parietal cortex. The role of the superior parietal cortex in spatial attention is well-known (Ungerleider & Haxby, 1994; Corbetta, Miezin, Shulman, & Petersen, 1993), especially during visual conditional task performance (Deiber et al., 1997; Deiber, Ibanez, Sadato, & Hallett, 1996; Kalaska, 1996). Our observed activation with low predictability sequences, irrespective of order, supports the role of attention in general and the superior parietal cortex in particular in resolving ambiguous spatial sequences.

Although the motor task itself can be considered a ''standard'' sensorimotor mapping (as the spatial location of the blue square directs the subject which button to press), our results did not indicate supplementary motor area (SMA) activation. SMA is commonly known as a motor sequencing region, and has been shown to be active in a variety of motor sequence learning studies (Grafton et al., 1995; Jenkins, Brooks, Nixon, Frackowiak, & Passingham, 1994). We believe that the lack of SMA activation in our study was due to the disparity between the visuospatial sequence and the motor performance. Unlike typical sequence learning studies (both explicit and implicit) in which participants perform a motor sequence, our participants merely observed the sequence and pressed a button only when a blue target was observed. The appearance of blue targets itself was not a spatial sequence. Thus, there was no motor sequence being performed by the participants, effectively obviating the need for learning a specific motor sequence. A lack of activation within putamen similarly suggests a lack of motor learning, as this area is typically active during motor sequence learning (Jueptner, Frith, Brooks, Frackowiak, & Passingham, 1997; Grafton et al., 1995; Jenkins et al., 1994).

When we talk about sequences or sequential behavior, we are really examining the issue of temporal target expectation; that is, the knowledge (whether implicit or explicit) that a particular order of stimuli or movements will occur in a predictable temporal pattern. It has long been proposed that the basal ganglia play a role in sequential learning (Marsden & Obeso, 1994). In particular, the phenomenon of behavioral learning may be traced to the release of dopamine and its hypothesized role as an event predictor (Schultz & Dickinson, 2000). Dopaminergic neurons within the substantia nigra pars compacta have been clearly shown to alter their response with regard to unexpected reward or appetitive events (Schultz, 1998), but whether a similar phenomenon occurs for ''neutral'' stimuli is unknown. In the context of the reward system, dopaminergic activity may represent expectancy about future stimuli or events. It appears that these changes in dopaminergic response propagate to other regions of the basal ganglia (Gdowski et al., 2001; Aosaki, Graybiel, & Kimura, 1994; Apicella, Scarnati, Ljungberg, & Schultz, 1992). Recent models of basal ganglia function support the concept that dopamine release is involved in sequential learning (Berns & Sejnowski, 1998); one model in particular demonstrates that the loss of the reinforcement signal (dopamine) can extinguish conditioned responses (Suri & Schultz, 1999). In our experiment, the high predictability condition elicited an increased activation in caudate, suggesting an expectation of temporal events that may project to the dopamine system. In essence, these neurons may be computing an expectation of future events, which the dopamine system could use to estimate prediction errors.

In summary, our findings both solidify the role of the right prefrontal–caudate circuit in learning visuomotor sequences and extends this to demonstrate that there is a parametric relationship to the informational content of the sequence itself. The context, namely the longer scale temporal order, apparently modulates the response in the caudate but not the prefrontal cortex, suggesting that the computation of sequential information may be computed on several timescales, all of which may converge in the basal ganglia. We also find further biological support for the hypothesis that attention is engaged when resolving ambiguous sequences.

METHODS

Experimental Design

Twenty right-handed participants aged 19 to 46 (average age = 28.5 years) gave written informed consent. Four squares were displayed horizontally across the screen (Figure 1). The squares were individually illuminated in one of three randomly selected colors (red, blue, or

yellow), and each square was illuminated for 1 sec. Participants were given a button box and were instructed to watch for blue squares. When a square was illuminated blue, they were instructed to press the button corresponding to the position of the blue square on the screen. Participants were instructed to use the four fingers of the right hand. The spatial order in which the squares were illuminated was determined by one of three conditions: (A) a four-element repeating spatial sequence, 1–3–2–4–. . ., where the number designates the square position beginning from the left (predictability = 2); (B) a probabilistic version of the sequence, based upon a Markov chain (predictability ≈ 1); and (C) a randomly ordered presentation (predictability = 0). Participants were given a 2-min practice session on the high predictability condition. Participants were not informed of the existence of a spatial sequence, nor did they report knowledge of a spatial sequence when questioned upon completion of the second imaging session.

During a functional imaging session, each condition was maintained for a block of 90 stimulus presentations (45 scans) before switching to the next condition block. All three conditions appeared once during each imaging session, and the task proceeded continuously without a break between the conditions during each session. There was a 2-min break between the scan sessions. The possible orders of the conditions were HI–LO (2–0–1) or LO–HI (1–0–2), indicating the order of the transition between the low (0) and high (2) predictability conditions. We divided the participants into two groups according to which condition they performed first: the HI–LO group (n = 12; 2 male, 10 female) and the LO–HI group (n = 8; 1 male, 7 female).

Entropy

We used a Markov chain to determine the conditional entropy (Cover & Thomas, 1991; Shannon & Weaver, 1949). We define the conditional entropy as $H(X_{i+1}|X_i) = -\sum p(x_i) \sum p(x_{i+1}|x_i) \log_2 p(x_{i+1}|x_i)$, where $H(X_{i+1}|X_i)$ is the first-order conditional entropy, $p(x_i)$ is the probability of event x_i occurring (e.g., which spatial position is chosen), and $p(x_{i+1}|x_i)$ is the probability of x_{i+1}, given that x_i occurs previously. Predictability, I, is defined by the mutual information theorem: $I(X_{i+1}|X_i) = H(X) - H(X_{i+1}|X_i)$, and represents the decrement in uncertainty provided by the preceding stimulus. Thus, the predictability measures of the three conditions were (A) the full sequences, 1–3–2–4. . . (I = 2.0); (B) a Markov sequence with the conditional probability matrix (I = 1.08) found in Table 2; and (C) a fully random sequence (I = 0).

Magnetic Resonance Imaging (MRI)

Two fMRI sessions of 270 scans each held 2 min apart were obtained during the continuous performance of the task. fMRI was performed with gradient-recalled echo planar imaging (TR = 2000 msec, TE = 40 msec, flip angle = 90°, 64 × 64 matrix, 24 5-mm contiguous axial slices) on a Philips 1.5 T scanner (Kwong et al., 1992; Ogawa et al., 1992). Structural, T1-weighted, MRIs were obtained for subsequent spatial normalization.

Table 2. Markov Sequence Conditional Probability Matrix

	x_{i+1}			
x_i	1	2	3	4
1	0	.15	.7	.15
2	.15	0	.15	.7
3	.15	.7	0	.15
4	.7	.15	.15	0

The matrix defines the probability relationships among the different elements of the conditional sequence for a Markov chain with entropy = .92 (where X_i represents the spatial position at trial i, and X_{i+1} the spatial position at the next trial).

Statistical Analysis

The data were analyzed using Statistical Parametric Mapping (SPM99, Wellcome Department of Cognitive Neurology, London, UK) (Friston et al., 1995). Motion-correction to the first functional scan was performed within subject using a six-parameter rigid-body transformation. The mean of the motion-corrected images was then coregistered to the individual's 24-slice structural MRI using a 12-parameter affine transformation. The images were then spatially normalized to the Montreal Neurologic Institute (MNI) template (Talairach & Tournoux, 1988) by applying a 12-parameter affine transformation followed by a nonlinear warping using basis functions (Ashburner & Friston, 1999). The spatially normalized scans were then smoothed with an 8-mm isotropic Gaussian kernel to accommodate anatomical differences across participants.

The data were analyzed on a voxel-by-voxel basis using an ANOVA with conditional predictability (0, 1, 2) as the main effect. A random-effects model was used to make statistical inferences (Friston, Holmes, & Worsley, 1999). A high-pass filter (cut-off = 180 sec) was applied to each time series, followed by the computing of six adjusted mean images for each participant, one per entropy level per imaging session. A subjectwise ANCOVA was used to remove any global signal intensity differences. A multigroup design matrix was specified with six adjusted mean images per subject using the subject groups LO–HI and HI–LO. Linear contrasts between the two extreme predictability levels (0 and 2) were examined for each group, using a threshold for significance of $p < .05$ with small-volume correction in regions of a priori hypotheses (caudate: 14-mm sphere; DLPFC 10-mm sphere; PPC 14-mm sphere). Using SPM99, a decreasing contrast vector [1 0 −1] was applied to the ordered predictability

levels (0, 1, 2), once per imaging session for each of the two groups. Additionally, an increasing contrast vector [−1 0 1] was applied in a similar fashion to each imaging session of the two groups. This permitted us to examine activations exhibiting either a negative or positive correlation to predictability and the interaction with the order of presentation, for example, LO–HI or HI–LO.

Acknowledgments

Supported by NIDA (K08 DA00367 to G.S.B.), NIMH (RO1 MH61010 to G.S.B.), the Stanley Foundation (G.S.B.), and a National Science Foundation Markey fellowship (A.B.G.).

Reprint request should be sent to Gregory S. Berns, Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine, 1639 Pierce Drive Suite 4000, Atlanta, GA 30322, USA, or via e-mail: gberns@emory.edu.

The data reported in this experiment have been deposited in the National fMRI Data Center (http://www.fmridc.org). The accession number is 2-2001-111PT.

REFERENCES

Aosaki, T., Graybiel, A. M., & Kimura, M. (1994). Effect of the nigrostriatal dopamine system on acquired neural responses in the striatum of behaving monkeys. *Science, 265,* 412–415.

Apicella, P., Scarnati, E., Ljungberg, T., & Schultz, W. (1992). Neuronal activity in monkey striatum related to the expectation of predictable environmental events. *Journal of Neurophysiology, 68,* 945–960.

Ashburner, J., & Friston, K. J. (1999). Nonlinear spatial normalization using basis functions. *Human Brain Mapping, 7,* 254–266.

Berns, G. S., Cohen, J. D., & Mintun, M. A. (1997). Brain regions responsive to novelty in the absence of awareness. *Science, 276,* 1272–1275.

Berns, G. S., & Sejnowski, T. J. (1998). A computational model of how the basal ganglia produce sequences. *Journal of Cognitive Neuroscience, 10,* 108–121.

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.

Cleeremans, A., & McClelland, J. L. (1991). Learning the structure of event sequences. *Journal of Experimental Psychology: General, 120,* 235–253.

Cohen, A., Ivry, R. I., & Keele, S. W. (1990). Attention and structure in sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 17–30.

Corbetta, M., Miezin, F. M., Shulman, G. L., & Petersen, S. E. (1993). A PET study of visuospatial attention. *Journal of Neuroscience, 13,* 1202–1226.

Courtney, S. M., Petit, L., Maisog, J. M., Ungerleider, L. G., & Haxby, J. V. (1998). An area specialized for spatial working memory in human frontal cortex. *Science, 279,* 1347–1351.

Cover, T. M., & Thomas, J. A. (1991). *Elements of information theory.* New York: Wiley.

Curran, T., & Keele, S. W. (1993). Attentional and nonattentional forms of sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 189–202.

Deiber, M.-P., Ibanez, V., Sadato, N., & Hallett, M. (1996). Cerebral structures participating in motor preparation in humans: A positron emission tomography study. *Journal of Neurophysiology, 75,* 233–247.

Deiber, M. P., Wise, S. P., Honda, M., Catalan, M. J., Grafman, J., & Hallett, M. (1997). Frontal and parietal networks for conditional motor learning: A positron emission tomography study. *Journal of Neurophysiology, 78,* 977–991.

Friston, K. J., Holmes, A. P., & Worsley, K. J. (1999). How many subjects constitute a study? *Neuroimage, 10,* 1–5.

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

Gdowski, M. J., Miller, L. E., Parrish, T., Nenonene, E. K., & Houk, J. C. (2001). Context dependency in the globus pallidus internal segment during targeted arm movements. *Journal of Neurophysiology, 85,* 998–1004.

Goldman-Rakic, P. S. (1996). The prefrontal landscape: Implications of functional architecture for understanding human mentation and the central executive. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences, 351,* 1445–1453.

Grafton, S. T., Hazeltine, E., & Ivry, R. (1995). Functional mapping of sequence learning in normal humans. *Journal of Cognitive Neuroscience, 7,* 497–510.

Hazeltine, E., Grafton, S. T., & Ivry, R. (1997). Attention and stimulus characteristics determine the locus of motor-sequence encoding. A PET study. *Brain, 120,* 123–140.

Jenkins, I. H., Brooks, D. J., Nixon, P. D., Frackowiak, R. S. J., & Passingham, R. E. (1994). Motor sequence learning: A study with positron emission tomography. *Journal of Neuroscience, 14,* 3775–3790.

Jueptner, M., Frith, C. D., Brooks, D. J., Frackowiak, R. S., & Passingham, R. E. (1997). Anatomy of motor learning: II. Subcortical structures and learning by trial and error. *Journal of Neurophysiology, 77,* 1325–1337.

Just, M. A., Carpenter, P. A., Keller, T. A., Eddy, W. F., & Thulborn, K. R. (1996). Brain activation modulated by sentence comprehension. *Science, 274,* 114–116.

Kalaska, J. F. (1996). Parietal cortex area 5 and visuomotor behavior. *Canadian Journal of Physiology and Pharmacology, 74,* 483–498.

Kwong, K. K., Belliveau, J. W., Chesler, D. A., Goldberg, I. E., Weisskoff, R. M., Poncelet, B. P., Kennedy, D. N., Hoppel, B. E., Cohen, M. S., Turner, R., & et al. (1992). Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proceedings of the National Academy of Sciences, U.S.A., 89,* 5675–5679.

Marsden, C. D., & Obeso, J. A. (1994). The functions of the basal ganglia and the paradox of stereotaxic surgery in Parkinson's disease. *Brain, 117,* 877–897.

McCarthy, G., Blamire, A. M., Puce, A., Nobre, A. C., Bloch, G., Hyder, F., Goldman-Rakic, P., & Shulman, R. G. (1994). Functional magnetic resonance imaging of human prefrontal cortex activation during a spatial working memory task. *Proceedings of the National Academy of Sciences, U.S.A., 91,* 8690–8694.

Ogawa, S., Tank, D. W., Menon, R., Ellermann, J. M., Kim, S. G., Merkle, H., & Ugurbil, K. (1992). Intrinsic signal changes accompanying sensory stimulation: Functional brain mapping with magnetic resonance imaging. *Proceedings of the National Academy of Sciences, U.S.A., 89,* 5951–5955.

Passingham, R. E. (1993). *The frontal lobes and voluntary action* (vol. 21). New York: Oxford University Press.

Rauch, S. L., Whalen, P. J., Savage, C. R., Curran, T., Kendrick, A., Brown, H. D., Bush, G., Breiter, H. C., & Rosen, B. R.

(1997). Striatal recruitment during an implicit sequence learning task as measured by functional magnetic resonance imaging. *Human Brain Mapping, 5,* 124–132.
Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior, 6,* 317–327.
Reber, A. S. (1993). *Implicit learning and tacit knowledge. an essay on the cognitive unconscious.* Oxford: Oxford University Press.
Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology, 80,* 1–27.
Schultz, W., & Dickinson, A. (2000). Neuronal coding of prediction errors. *Annual Review of Neuroscience, 23,* 473–500.
Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication.* Chicago: University of Illinois Press.
Stadler, M. A. (1989). On learning complex procedural knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 1061–1069.
Suri, R. E., & Schultz, W. (1999). A neural network model with dopamine-like reinforcement signal that learns a spatial delayed response task. *Neuroscience, 91,* 871–890.
Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the brain.* New York: Thieme.
Ungerleider, L. G., & Haxby, J. V. (1994). 'What' and 'where' in the human brain. *Current Opinion in Neurobiology, 4,* 157–165.