Neural Correlates of Sequence Learning with Stochastic Feedback

Bruno B. Averbeck[1], James Kilner[1], and Christopher D. Frith[1,2]

Abstract

■ Although much is known about decision making under uncertainty when only a single step is required in the decision process, less is known about sequential decision making. We carried out a stochastic sequence learning task in which subjects had to use noisy feedback to learn sequences of button presses. We compared flat and hierarchical behavioral models and found that although both models predicted the choices of the group of subjects equally well, only the hierarchical model correlated significantly with learning-related changes in the magneto-encephalographic response. The significant modulations in the magneto-encephalographic signal occurred 83 msec before button press and 67 msec after button press. We also localized the sources of these effects and found that the early effect localized to the insula, whereas the late effect localized to the premotor cortex. ■

Downloaded from http://direct.mit.edu/jocn/article-pdf/23/6/1346/1775571/jocn.2010.21436.pdf by guest on 18 June 2021

INTRODUCTION

Coherent sequences of decisions are an important feature of organized behavior, and real-world decisions are often made with uncertainty. Given the importance of coherent behavior and the fact that it is often disrupted by brain disorders (McKenna & Oh, 2005; Schwartz, Reed, Montgomery, Palmer, & Mayer, 1991; Andreasen, 1979a, 1979b; Penfield & Evans, 1935), there is a need to understand the neural processes involved in the learning and orchestration of sequences of decisions. However, most work on decision making has focused on single-step decisions (Wittmann, Daw, Seymour, & Dolan, 2008; Fellows & Farah, 2007; Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006; Hsu, Bhatt, Adolphs, Tranel, & Camerer, 2005; Samejima, Ueda, Doya, & Kimura, 2005; Glimcher & Rustichini, 2004).

In the current study, we addressed several aspects of sequential decision making. The first questions we addressed were when and where do the cognitive processes relevant to learning sequences from stochastic feedback take place? We used magneto-encephalographic (MEG) imaging, which has high temporal precision, to record brain activity during the task. This allowed us to define in time when cognitive processes became active and allowed us to analyze data on a movement-by-movement basis, without making movements artificially slow. Furthermore, we used source localization tools to estimate where in space these processes were taking place. As such, we examined how the signals related to learning in the task evolve over time and space.

In addition, we were interested in the extent to which subjects used optimal strategies to learn in our task. In an interesting contradiction, two dominant fields of enquiry separately describe behavioral processes as either optimal (Knill & Saunders, 2003) or heuristic, which implies suboptimality (Kahneman, Slovic, & Tversky, 1982). To a large extent, these perspectives come from studying different classes of behaviors (Trommershauser, Maloney, & Landy, 2008). Many who study sensory–motor integration have found that subject performance can have features of optimality (Kording & Wolpert, 2006; Knill & Saunders, 2003; Trommershauser, Maloney, & Landy, 2003a, 2003b; Ernst & Banks, 2002; Todorov & Jordan, 2002; Jacobs, 1999; Kersten, 1999; Knill & Richards, 1996; Poggio, Torre, & Koch, 1985), whereas those who study decision under risk have consistently shown that subject performance is not optimal and instead heuristic (Gilovich, Griffin, & Kahneman, 2002; Payne, Bettman, & Johnson, 1992; Tversky & Kahneman, 1986; Kahneman et al., 1982), the neural correlates of which are currently being explored (Tobler, Christopoulos, O'Doherty, Dolan, & Schultz, 2008). Thus, the difference may have to do with the class of behavioral process being studied as well as other features, including whether models are updated with learning (Trommershauser et al., 2008).

The stochastic sequence learning task we used has elements of both sensory–motor integration and decision making, as the uncertainty in our task is external and not due to noise in the sensory–motor system. Furthermore, by bringing decision making into a sequential framework, we examined aspects of the cognitive processes that underlie decision-making behaviors that do not exist in single-step decision-making paradigms. Specifically, we asked whether subjects were able to learn optimal

[1]University College London, [2]Århus University Hospital, Denmark

© 2011 Massachusetts Institute of Technology

sequences of decisions when information at other points in the sequence affected the current choice. This relationship among parts of a sequence can be modeled using a hierarchical model. Many important human behaviors, including speech, contain hierarchical structure between sequence elements, but it is not clear if these processes are solved by a domain general system or by a domain-specific language system (Fiebach & Schubotz, 2006).

METHODS

Subjects and Task

Fourteen subjects (7 men) carried out a stochastic sequence learning task while being scanned. Informed consent was obtained from each subject in accordance with procedures approved by the Joint Ethics Committee of the National Hospital for Neurology and Neurosurgery and the Institute of Neurology, London. Each trial began with the presentation of a green outline circle at the center of the screen, which cued the subjects to execute a movement (Figure 1A). They responded by executing a button press with either their left or right thumb. After each response, they were given feedback about whether they had pressed the correct button for that movement of the sequence. If they were correct, the outline circle filled green (positive feedback), and if they were incorrect, the outline circle filled red (negative feedback). After the feedback had been given for 200 msec, they were again presented with a green outline circle that cued the next movement of the sequence. They again pressed one of two buttons and were given feedback. This was repeated four times, such that each trial was composed of four button presses, with each button press followed by feedback. In 15% of the cases, the wrong feedback was given. In other words, if they had pressed the correct button, they were given red (negative) feedback, and if they had pressed the incorrect button, they were given green (positive) feedback. Thus, the feedback from any individual button press did not necessarily allow the subjects to correct their mistakes and execute the correct sequence in the subsequent trial. Integrated over trials, however, the subjects could infer the correct button press sequence.

Exact task timing depended on the RTs of the subjects and in the case of visual feedback, on where the screen refresh cycle was when we initiated draw commands. The mean RT following the pacing cue was 385 msec (SD = 284 msec). Visual feedback followed the button press by 40 msec (SD = 6 msec). The mean time between the feedback and the presentation of the subsequent pacing cue was 226 msec (SD = 9 msec), and the mean time between button presses was 602 msec (SD = 275 msec). The break between trials was indicated with a blank screen that lasted 1 sec.

We used 6 of the 16 possible sequences (i.e., 4 button presses with either the left or the right thumb gives 16 possible sequences), which were balanced for first-order button press probabilities and contained at least two left and two right button presses. The sequences used were LLRR, RRLL, LRLR, RLRL, LRRL, and RLLR. Each subject executed eight sets of six blocks, where each block was one sequence (Figure 1B). A single set consisted of all six sequence blocks where the order of the blocks was chosen pseudorandomly. For example, a block of Sequence 4 followed by a block of

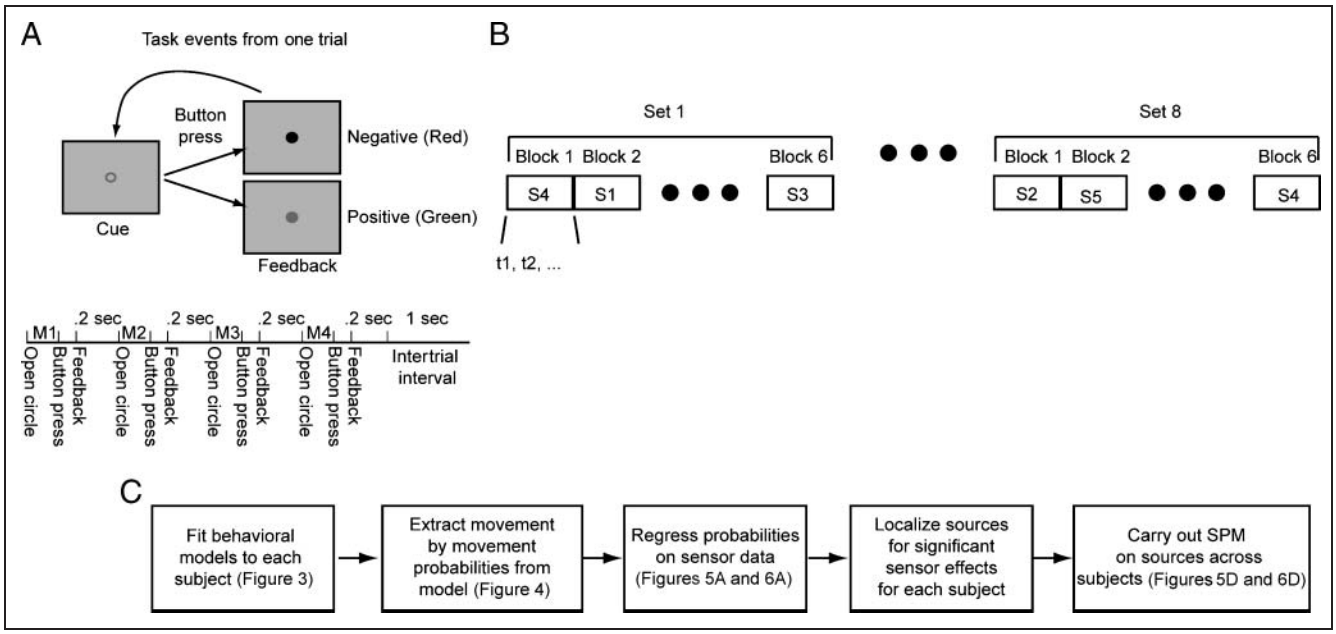

Figure 1. Task. (A) Top shows sequence of images presented in a single trial. Bottom shows timing of events. Each cue, button press, feedback series is presented four times in a single trial. Cue and positive feedback were green; negative feedback was red. (B) Sequence of sets, blocks, and trials. Each set is composed of six sequence blocks. Subjects progress through a single sequence block by learning and executing the sequence correctly eight times. (C) Sequence of analyses presented in Results.

Downloaded from http://direct.mit.edu/jocn/article-pdf/23/6/1346/1775571/jocn.2010.21436.pdf by guest on 18 June 2021

Sequence 1, and so forth. In each block, subjects had to determine the correct sequence using the stochastic feedback and execute it correctly eight times before they advanced to the next block. Subjects were informed when the block switched, and thus they knew when they had to start learning a new sequence. If they failed to complete the sequence correctly eight times by 20 trials, they were advanced to the next sequence. If the subjects completed all blocks of trials, each sequence was executed correctly 64 times within a session, resulting in a total of 384 correct sequences and 1,536 correct button presses for each subject. In addition, subjects received at least one set of training on all six sequences before entering the MEG and were instructed that they only had to determine which of the six sequences was correct in the current block. The training familiarized them with the six sequences, the stochastic feedback, and the other aspects of the task. The subjects were not told specifically that all sequences would require two left and two right button presses. Thus, within the experiment, the subjects were familiar with the mechanics of the task and the sequences, and their task was to use the stochastic feedback for each sequence to estimate which sequence was correct and then execute that sequence.

The task was challenging, and we found post hoc that in about 9% of the blocks, the 12 subjects retained for analysis failed to learn the sequence. Specifically, they went through 20 trials without learning to criterion. With respect to the MEG analysis described below, we carried out the analysis both with and without the blocks of data in which they failed to reach criterion. Although including or not including these data affected the exact F values of the MEG analysis, the rank order of the fit of the three models was the same (see Results), and the plots showing the spatial distribution of the effects were similar. We report results for the analyses in which we excluded the blocks in which subjects did not learn.

Data Analysis

Analyses proceeded in a series of steps (Figure 1C). First, we fit behavioral models to each subject's learning data. Then, we extracted the movement-by-movement estimates of learning from each subject's behavioral model. This learning estimate (quantified as the probability that the subject knew which button they should press at each point in time) was then regressed on the data in sensor space to see if the button-press-related activity was modulated by learning. We then used source localization on a subject-by-subject basis to localize the significant effects that we found in sensor space. After localization was done for each subject, the results were used to carry out SPM statistics on whether the localizations were significant across subjects.

Behavioral Model

We fit Bayesian statistical models to the subjects' behavior. The models allowed us to quantify trial by trial how much the subjects had learned about which sequence or which button was correct in the current block. The subjects could press either the left or the right button at each point in the sequence, and therefore they had a binary decision. The model assumes that the subjects were trying to learn the sequence and therefore that they were trying to optimize the number of times green feedback was received. Statistically, this can be accomplished by remembering how often green feedback was given for the left (or right) button at each point in the sequence. For example, if green feedback was given more often for the left button, then the left button should be pressed. Thus, the model integrates information about red versus green feedback given for left and right button presses individually for each of the four button presses in the sequence.

The model began with a binomial likelihood function for each movement of the sequence, given by

$$p(D_T|\theta_{i,j}) = \theta_{i,j}^{r_{i,j}}(1-\theta_{i,j})^{N_j - r_{i,j}} \tag{1}$$

where $\theta_{i,j}$ is the probability that pressing button i ($i \in$ {left, right}) on movement j ($j \in \{1,\ldots,4\}$) would be followed by green feedback. The variable $r_{i,j}$, defined below, is the number of times reward (green feedback) was given when button i was pressed on movement j (or red feedback was given when the other button was pressed), and N_j is the number of trials. The vector D_T represents all the data collected up to trial T for the current block, which in this case are the values of r and N. This was the only data relevant to inferring the correct sequence of button presses. Importantly, the model does not contain any information about previous sequences from the current set. Subjects were not told that all sequences were given in each set, and therefore it was unlikely that they would be able to infer set boundaries and use this information to improve learning.

The probability that the left button should be pressed for movement j after T trials (i.e., that it is more likely to be the correct button) is given by

$$\begin{aligned}\Delta_{j,T}^F \equiv P_{j,T}^F(B=L) &= p_T(p(\theta_{\text{left},j}|D_T) > p(\theta_{\text{right},j}|D_T)) \\ &= \int_0^1 p(\theta_{\text{left},j}|D_T)\int_0^{\theta_{\text{left},j}} p(\theta_{\text{right},j}|D_T)d\theta_{\text{right},j}d\theta_{\text{left},j}.\end{aligned} \tag{2}$$

We have written the posterior here (i.e., $p(\theta_{\text{right},j}|D_T)$). Button probabilities were equally likely in the experiment so the prior was flat and the posterior is just the normalized likelihood for this estimate. The superscript F on Δ indicates that this is the probability used for button presses in the flat model.

For the hierarchical model, a few additional computational steps were necessary to compute the button probabilities. In effect, the hierarchical model uses feedback about all four button presses in the sequence to estimate the probability that either the left or the right button is

Downloaded from http://direct.mit.edu/jocn/article-pdf/23/6/1346/1775571/jocn.2010.21436.pdf by guest on 18 June 2021

correct at each individual point in the sequence. This is the feature that distinguishes it from the flat model, and this is possible because not all button press combinations were used in the experiment and therefore certain possibilities could be eliminated. We eliminated these possibilities computationally as follows. Given that we had an estimate of the probability that either the left or the right button should be pressed at each point in the sequence, we could calculate the probability that each of the sequences was correct. This probability was given by

$$P(D_T|s_k) = \prod_{j=1}^{4} \left(\Delta^F{}_{j,T} B_{j,k} + \left(1 - \Delta^F{}_{j,T}\right)(1 - B_{j,k}) \right), \quad (3)$$

where $B_{j,k}$ is 1 if the left button was correct for movement j, sequence k (s_k), and 0 if the right button was correct. Formally, B is the conditional probability of the left button being correct given the sequence, but because it is a delta function, we have not used probability notation. The posterior probability of each sequence was then given by Bayes rule

$$P(S) \equiv P(s_k|D_T) = \frac{P(D_T|s_k)p(s_k)}{P(D_T)}. \quad (4)$$

Again, we have used a flat prior on sequences, as this was in fact the prior in the experiment, so $p(s_k) = 1/6$ for all sequences.

In the hierarchical model, these sequence probabilities can then be used to infer the button probabilities. As mentioned above, the hierarchical model uses information about all four button presses to infer the probabilities of each individual button press (Equation 3). In Equation 4, the individual button probabilities were used to infer the sequence probabilities. These sequence probabilities can then be used to infer the button probabilities under the hierarchical model. This was calculated as

$$\Delta^H_{j,T} \equiv P^H_{j,T}(B = L) = \sum_{k=1}^{6} B_{j,k} p(s_k|D_T), \quad (5)$$

Where again $B_{j,k}$ was 1 if the left button was correct for movement j, sequence k, and 0 if the right button was correct. Incidentally, although the second-order transitions between buttons were not balanced (i.e., L → R vs. L → L), the hierarchical model subsumes and models this fact by modeling all the correlations between button probabilities.

Equations 1–5 describe an ideal observer model, but it is likely that the subjects' behavior will deviate from this model. To better predict their behavior, we added two parameters to the basic model that allowed for differential weighting of positive and negative feedback. These parameters affect both the flat and the hierarchical model. When the models were fit to the subject behavior, different parameters were fit under each model to individual subjects, as shown in the Results section.

The differential weighting was implemented by using the following equation for the feedback:

$$\begin{aligned} fb_{\text{positive}} &= 0.5 + \alpha \\ fb_{\text{negative}} &= 0.5 + \beta \end{aligned} . \quad (6)$$

The subscripts positive and negative indicate whether the feedback was positive (green) or negative (red). The total reward (feedback) in Equation 1 was then given by

$$r_{i,j} = \sum_{t=1}^{T} fb_{\text{positive}}\, u(t) + fb_{\text{negative}}\,(1 - u(t)). \quad (7)$$

The parameter $u(t)$ is 1 if green feedback was given and 0 if red feedback was given on trial t. Thus, α and β scale the amount that is learned from positive and negative feedback. For an ideal observer, both parameters would be .5. The parameters α and β were fit to individual-subject decision data by maximizing the likelihood of the subject's sequence of decisions, given the model parameters. Thus, we maximized:

$$p^M(D^*|\alpha, \beta) = \prod_{t=1}^{N} \left(\Delta^M_t C_t + \left(1 - \Delta^M_t\right)(1 - C_t) \right), \quad (8)$$

where C is the choice that the subject made for each movement ($C = 1$ for left, $C = 0$ for right), and the superscript on Δ and p indicates the probability calculated under either the flat or the hierarchical model (i.e., $M = F$ or $M = H$) and D^* indicates the vector of decision data, C_t, or the sequence of button presses. This function was maximized using nonlinear function maximization techniques in Matlab. Examination of the likelihood function around the maximum likelihood estimates showed that likelihood values dropped off smoothly and in a relatively Gaussian way as the parameter values changed. We also tried multiple starting values for the parameters (0, 0.1, 0.3, and 0.5) and found that the algorithm always converged to the same answer. Thus, our estimates were not likely local minima. In part, this could be due to the large amount of data that was collected for each subject.

Two of the subjects performed very poorly in the task (Subjects 5 and 8), and their parameter values reflected this as they had values near zero for both α and β, implying no integration of the feedback. Once the parameters were found that maximized the likelihood, either Equation 2 ($P^F(B)$) or Equation 5 ($P^H(B)$) was used to generate the probability, on a trial by trial basis, that the subject knew which button to press at each point in the sequence, and Equation 4 was used to estimate the sequence probability ($P(S)$), given the feedback from the current block.

Downloaded from http://direct.mit.edu/jocn/article-pdf/23/6/1346/1775571/jocn.2010.21436.pdf by guest on 18 June 2021

Finally, for model comparison on the behavioral data the log-likelihood ratio for the two models was calculated as:

$$llr = \log p^F(D^*|\alpha, \beta) - \log p^H(D^*|\alpha, \beta). \quad (9)$$

Because the models had the same number of parameters, no corrections for degrees of freedom had to be done. We could simply compare the fit of the two models on a subject-by-subject basis and look at the distribution of the *llr* across subjects to see if it was significantly positive or negative.

We also examined three other models that were extensions of the basic model. None of them fit the subject behavior better than the flat model, and therefore we will not go into the formal description of the models. However, it is worth describing them briefly as we did explore various possibilities, beyond the flat and hierarchical models, for how subjects might have been performing the task. We were trying to model the possibility that subjects were integrating information over the first few trials of the block, but at some point they decided which sequence they thought was correct in the block. In that case, the subjects' probability estimate would jump from its current value to 1. Effectively, we were trying to model the possibility that when the probability estimate crossed a threshold the subjects were sufficiently confident of which sequence they were executing and their belief estimate went to 1 or some very high value. We did this in three ways. First, we passed the button press probabilities (Equation 2) generated by the basic model through a soft-max function (Bishop, 1995). This function is used to convert value estimates in reinforcement learning models into probabilities. It contains a temperature parameter that controls how much probabilities are amplified. In effect, it causes the probabilities during learning to go much more quickly to 1. We estimated the temperature as a free parameter but found that it did not improve the fit. The second approach was to implement a threshold. In this case, when the belief value crossed that threshold it was set to 1. In the third case, we switched from the button probability generated by Equation 2 to the sequence button probability generated by Equation 5 when the probability passed a threshold. The threshold was always allowed to vary as a free parameter. As stated, however, none of these three approaches worked better than the basic button model.

MEG Data Acquisition and Preprocessing

MEG data were recorded using 275 third-order axial gradiometers with the Omega275 CTF MEG system (VSM MedTech, Vancouver, Canada) located in a magnetically shielded room. The signals were recorded at a sampling rate of 480 Hz. Visual stimulus lag was estimated using a photodiode to measure the onset time at the screen relative to the signal sent from the task control computer to the data acquisition computer and was found to be 25 msec. This delay was used in calculating all of our timing values. Data analysis was carried out using SPM5 (Wellcome Department of Imaging Neuroscience, London) and custom written Matlab routines to implement the behavioral model.

We began the analyses by low-pass filtering the MEG signal at 50 Hz and downsampling to 120 Hz. The data were then epoched into a 400-msec window centered on the button press. Trials with a response that had an absolute value greater than 3000 fT were discarded as outliers. Statistical analysis and source localization was carried out using SPM. Details of the general linear model (GLM) that was fit are given below. Gaussian random field theory was used to control for multiple comparisons in either 2-D space × 1-D time (sensor space) or 3-D space (source space) (Kiebel, Kilner, & Friston, 2007; Kilner, Kiebel, & Friston, 2005). Sensors were converted into a 2-D space using Gaussian interpolation. Smoothing was never done across both time and space. Thus, the signals were first filtered in time and then filtered in space at a single time point.

When relevant sensor space effects were identified, we estimated the sources of these effects using source reconstruction algorithms in SPM5. For each subject, we constructed a forward model describing the transformation between distributed dipole sources and the magnetic field distribution measured by the MEG sensors. Sources were modeled using the 7204 vertex template cortical mesh available in SPM5, defined in Talairach and Tournoux coordinates. It was coregistered to the sensor locations via three fiducial marker positions (Mattout, Henson, & Friston, 2007), and the gain matrix of the lead-field model was then computed using a spherical head model, and source estimates were computed using restricted maximum likelihood estimation to invert the forward model within a parametric empirical Bayes framework (Mattout, Phillips, Daunizeau, & Friston, 2007; Mattout, Phillips, Penny, Rugg, & Friston, 2006). This inversion proceeded using multiple sparse empirical priors for covariance components (Friston, Harrison, et al., 2008). The greedy search algorithm (Friston, Chu, et al., 2008) provided an optimal mixture of sparse prior components. This produced source reconstructions for each experimental condition and for each subject.

We then compared activation levels on the mesh across subjects using a random effects model (i.e., a second level model in SPM). Voxels were accepted at an uncorrected p value of .01, and all significance values are reported at cluster level corrected for whole brain. Sources were estimated in a 2 × 2 design, Probability × Button, where probability was low or high (i.e., early or late learning). Significance of the probability factor was calculated by first computing main effect contrasts subject by subject and then doing univariate t tests.

For the GLM regressions that looked for an effect of the learning-related probabilities on the MEG sensor response, we carried out analyses in two ways. First, we examined the

Downloaded from http://direct.mit.edu/jocn/article-pdf/23/6/1346/1775571/jocn.2010.21436.pdf by guest on 18 June 2021

probability of the button that was actually pressed, under both the flat and hierarchical models. Thus, the first level regressor was given as

$$n(t) = \Delta_t^M C_t + \left(1-\Delta_t^M\right)(1-C_t). \qquad (10)$$

Interestingly, this regressor was not significant under either model (data not shown). We then used a regressor that had only the probability of the correct button press or sequence at the current point in time. This was given by

$$m(t) = \Delta_t^M B_{j,k} + \left(1-\Delta_t^M\right)(1-B_{j,k}). \qquad (11)$$

This regressor is what is plotted in Figure 4. It tracks learning more closely and is the same whether the most probable or the least probable button was pressed. The sequence probability given by Equation 4 was also used, in which case the currently correct sequence was plugged in as the right-hand side of Equation 11.

When carrying out these analyses, we correlated the regressor, $m(t)$, on a movement-by-movement basis with the time-varying MEG responses at each sensor and each time point in the 400 msec window around button press, similar to what is done in fMRI data with voxels. In our case, we treated the 2-D sensor space by 1-D time space as a 3-D volume, as described above. Thus, we fit the following GLM to each point in the 3-D volume

$$q_{x,y,l} = a_0 + bm(t) + cg(t) + dm(t)g(t) \qquad (12)$$

where $q_{x,y,l}$ indicates the response or MEG equivalent voxel in the 3-D volume (i.e., x and y = position and l = time around button press) and $g(t)$ is the button press for that movement, which is a dummy variable with values -1 and 1. The last parameter is an interaction term. This model was fit to individual subjects, and then the parameter estimates for the individual subjects were taken to the second level in an SPM random effects approach.

RESULTS

Behavior

Fourteen subjects carried out the stochastic sequence learning task (Figure 1), in which they had to learn sequences of four left/right button presses (e.g., LLRR, LRLR, etc.). Explicit feedback was given after each button press, but 15% of the time inaccurate feedback was given. Two of the subjects failed to learn the sequences to criterion, so their results will not be discussed further. It took the remaining 12 subjects, on average, 3.2 trials of learning with each new sequence before they executed a complete trial correctly, where a correct trial was defined as pressing all four buttons in the sequence correctly (Figure 2A). The subjects' performance reached a plateau by about three trials correct (about six total, i.e., combined correct and error trials) in each block, and they executed the sequence in the remaining trials with few errors. When the performance was examined as the serial position of the movement, results were similar. It could be seen, however, that the early movements of the sequence were learned slightly faster (Figure 2B). There was also a small bias to perform better on the early and last movements of the sequence between correct Trials 3 and 4 (Figure 2B), consistent with the primacy and recency gradients seen in most sequence tasks (Averbeck, Chafee, Crowe, & Georgopoulos, 2002). However, later in the block, the performance reached a ceiling, and this effect could not be seen.

In the task, the subjects made a sequence of four left or right button presses. Thus, there were 16 possible sequences (2^4). However, we only used 6 of the 16 possible sequences in the experiment. Because we only used a subset of the possible sequences, there were correlations between the correct button presses, and feedback about

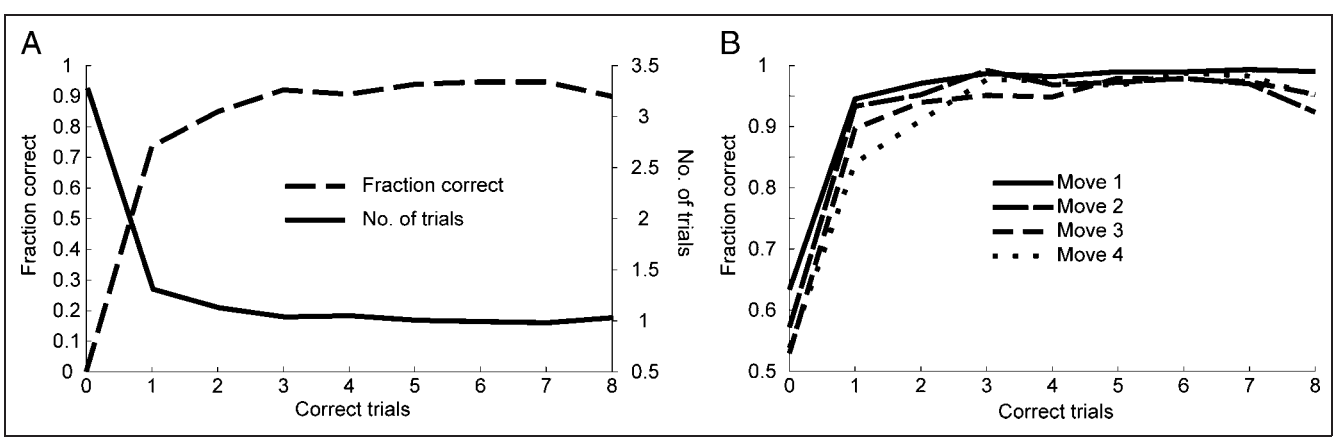

Figure 2. Learning rate. (A) Fraction correct complete sequences (i.e., four correct button presses) and number of trials it takes to get one trial correct. Both are plotted as a function of the number of correct trials in the current block. (B) Fraction correct individual button presses for first though fourth button press in each sequence.

Downloaded from http://direct.mit.edu/jocn/article-pdf/23/6/1346/1775571/jocn.2010.21436.pdf by guest on 18 June 2021

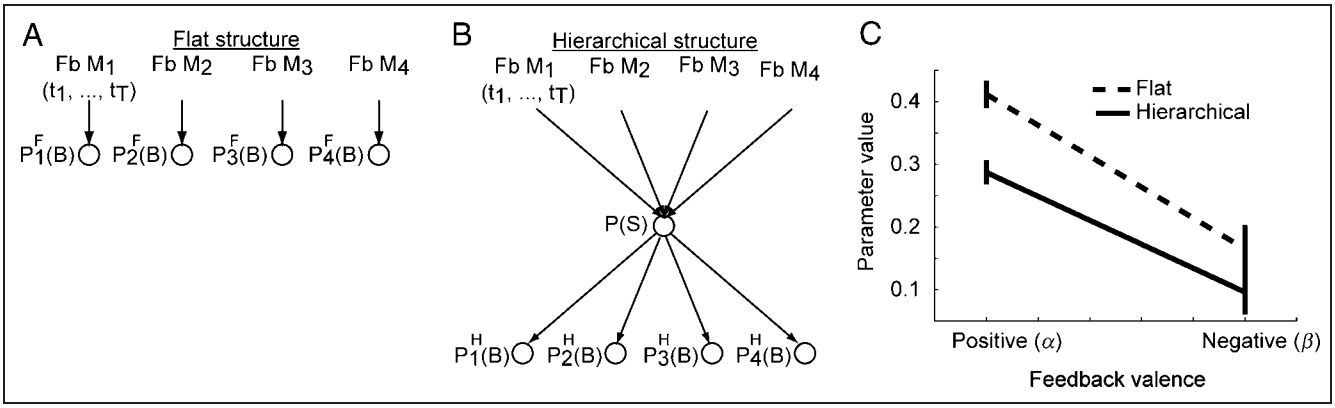

Figure 3. Probability model of learning. (A) Flat architecture, in which feedback about individual movement affects only that movement. (B) Hierarchical architecture, in which feedback about individual movements is used to infer the sequence, which is then used to infer the movements. (C) Mean and *SEM* of parameter values across the 12 subjects for positive (α) and negative (β) feedback under the two models.

button presses at other points in the sequence could be used to better infer the correct button at the current point in the sequence. For example, if in a particular block the subject was certain that the first two button presses were LL, they would predict that the subsequent button presses were RR even if the evidence for RR was equivocal because LLRR was the only sequence we used that started with LL. These correlations between button presses can be represented with a hierarchical structure.

We explored the hypothesis that subjects took this hierarchical or correlational structure into account when they were learning the sequences. To do this, we predicted the subjects' behavior with two different models, one which did not take the hierarchical structure into account (Figure 3A, flat model) and one which did (Figure 3B, hierarchical model). Behavioral parameters were optimized for each subject under each model. We found that subjects learned more from positive feedback than negative feedback under both models (Figure 3C). The parameter values were also lower for the hierarchical model than the flat model. This is due to the fact that the hierarchical model is more efficient with the data because it correctly models the actual stochastic process used in the experiment. Because parameters were optimized under each model, learning rates for the two models were similar, although the hierarchical model tended to learn more smoothly across trials because it integrated information across buttons, as can be seen in a single example block from a single subject (Figure 4).

We next examined whether we could find differences in the ability of the two models to predict the behavior of the subjects. We did this by computing a t test on the log-likelihood ratio of the two models across subjects. A negative log-likelihood ratio for a single subject favored the hierarchical model. There was, however, no significant difference between the models, across our subjects, $t(11) = 1.8, p = .09$. We did find that the behavior of the subjects who learned better was better described by the hierarchical model. Thus, the total number of trials it took subjects to finish the task, which is a measure of how efficiently they learned, was negatively correlated with the log-likelihood ratio ($r = -0.76, p < .01, n = 12$). In other words, the hierarchical model better predicted the behavior of subjects who learned more efficiently.

MEG Responses—Parametric Effect of Probability

We measured MEG responses while subjects learned and executed the sequences. We began by examining effects in sensor space, that is, by looking at task effects on the temporal response in the interpolated scalp map around the time of button press (200 msec before to 200 msec

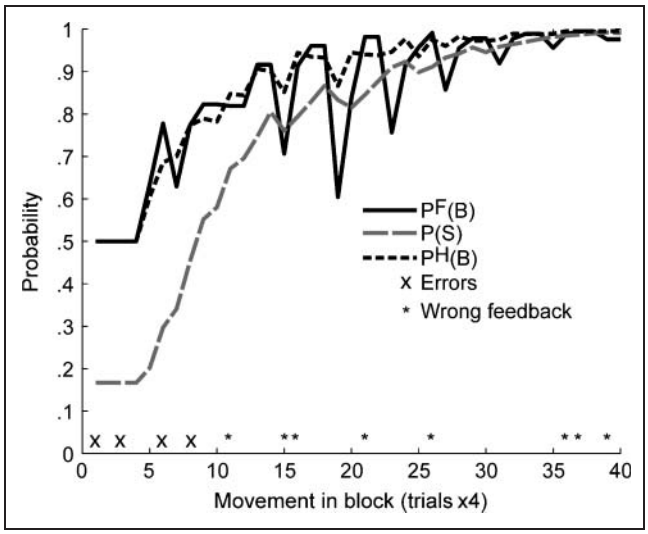

Figure 4. Probabilities for an example block under flat and hierarchical models. Probability that the feedback up to the current trial supported the correct button press under the flat model (P^F), the correct button press under the hierarchical model (P^H), or the correct sequence (P(S)) at each point in the sequence, as it evolves within a single example block of trials. Xs above the x-axis indicate the movements on which the subject pressed the incorrect button, and asterisks indicate the movements in which the wrong feedback was given (i.e., if the subject pressed the correct button for the sequence, they were given red feedback, and if the subjects pressed the incorrect button, they were given green feedback).

Downloaded from http://direct.mit.edu/jocn/article-pdf/23/6/1346/1775571/jocn.2010.21436.pdf by guest on 18 June 2021

after button press). All analyses in sensor space were carried out by converting the 2-D sensor × 1-D time data into a 3-D volume, which allowed us to carry out corrections for multiple comparisons using the tools developed for fMRI data (see Methods). Reported statistical results are based on clusters of samples that exceeded a threshold.

In our analyses of the MEG data, we compared the relative ability of three different learning-related variables, taken directly from the behavioral models discussed above, to predict the change in the MEG response with learning (Figure 4). We examined the button probabilities under the flat ($P^F(B)$) and hierarchical models ($P^H(B)$), and we also considered the probability of the sequence ($P(S)$). In all cases, these probabilities were derived on a movement-by-movement basis with the models fit to the behavioral data of the individual subjects, including separate modeling of positive and negative feedback, as described above. These probabilities represent how much has been learned about the sequence before the current button press. Thus, they represent the subjects' knowledge of which button should be pressed or which sequence is correct in the current block. We also included the button that was pressed and the interaction between the probability and the button that was pressed in the analysis (see Equation 12).

As we were interested in studying the learning effect, we only carried out our analysis on blocks in which the subjects completed eight sequences correctly, although results in sensor space were similar when all blocks were used (data not shown). There were no significant clusters for the button probability under the flat model ($P^F(B)$; $p = .24$, t test, cluster level, $df = 11$), and there was one cluster that just missed significance under the button probability for the hierarchical model ($P^H(B)$; $p = .052$, t test, cluster level, $df = 11$). The sequence effect, however, had two significant clusters ($P(S)$; $p < .05$, t test, cluster level, $df = 11$), one had a maximum at 83 msec before button press (Figure 5) and extended above threshold from 108 to 67 msec before button press. The other had a maximum at 67 msec after button press (Figure 6) and extended above threshold from 58 msec after until 150 msec after button press. There was also a significant interaction between the sequence probability and the button that was pressed in the cluster which followed the button press ($p < .05$, t test, cluster level, $df = 11$, max at 67 msec, above threshold from 58 to 100 msec after button press), but not for the cluster that preceded the button press.

These modulations of the MEG response manifested as changes in the temporal evolution of the signal just before and after button press (Figure 5B); that is, the temporal response was different depending on how well the sequence had been learned in the current block. In all cases, there was an increase in the signal early in the block, when the sequence was not well learned ($P(S) = 0.33$), leading to a decrease later in the block when the sequence was well learned ($P(S) = 1.0$). This could also be seen in sensor space, as the parametric regressor was negative in the region with significant effects (Figure 5C). Thus, as the probability increased across the block (Figure 4), the MEG signal decreased. Interestingly, these learning-related differences occurred before and after the peak response (Figure 5B). This suggests that the peak response reflects a

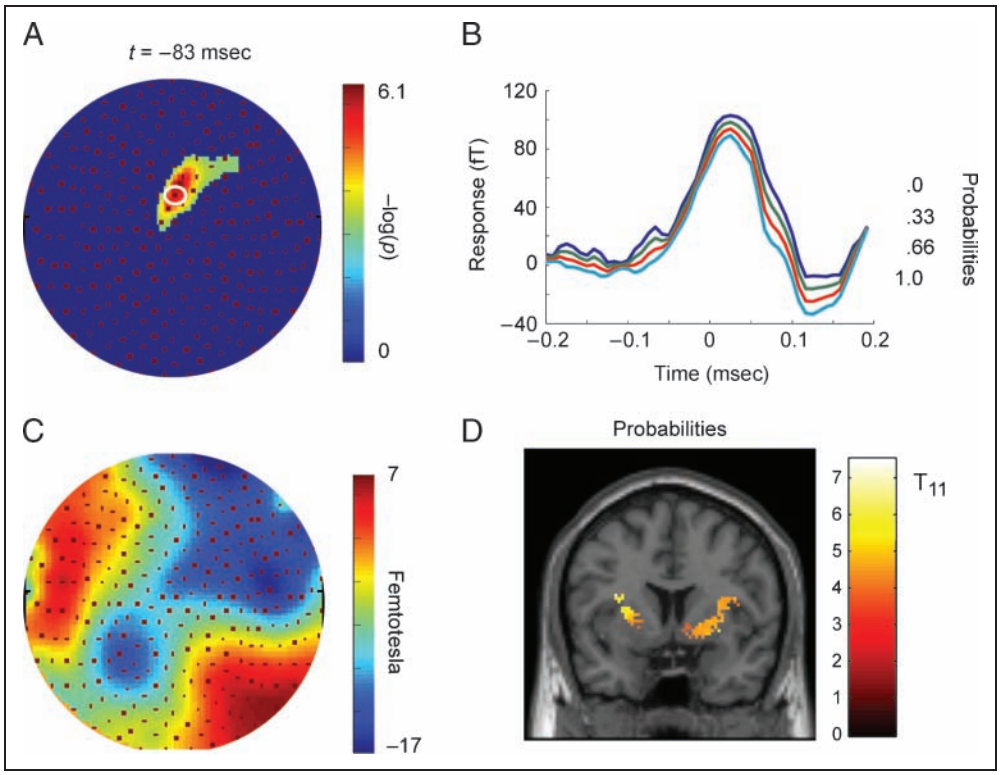

Figure 5. Significant sensor modulation with probability 83 msec before button press based on sequence model. (A) Interpolated map of p values in sensor space. Only p values below .001 are shown, and all values plotted are the negative log of the p value (i.e., a value of 2 is .01). (B) Temporal response at a single sensor (indicated in panel A with white circle). Each line represents predicted response averaged across subjects for one probability, with probabilities indicated on right of panel. (C) Contrast of parametric regressor representing probability main effect in the GLM. (D) Estimated source of effect. Plot shows a significant cluster of voxels (peak activation: $x = -30$, $y = 2$, $z = 10$) in the insula. Clusters were significant bilaterally.

Downloaded from http://direct.mit.edu/jocn/article-pdf/23/6/1346/1775571/jocn.2010.21436.pdf by guest on 18 June 2021

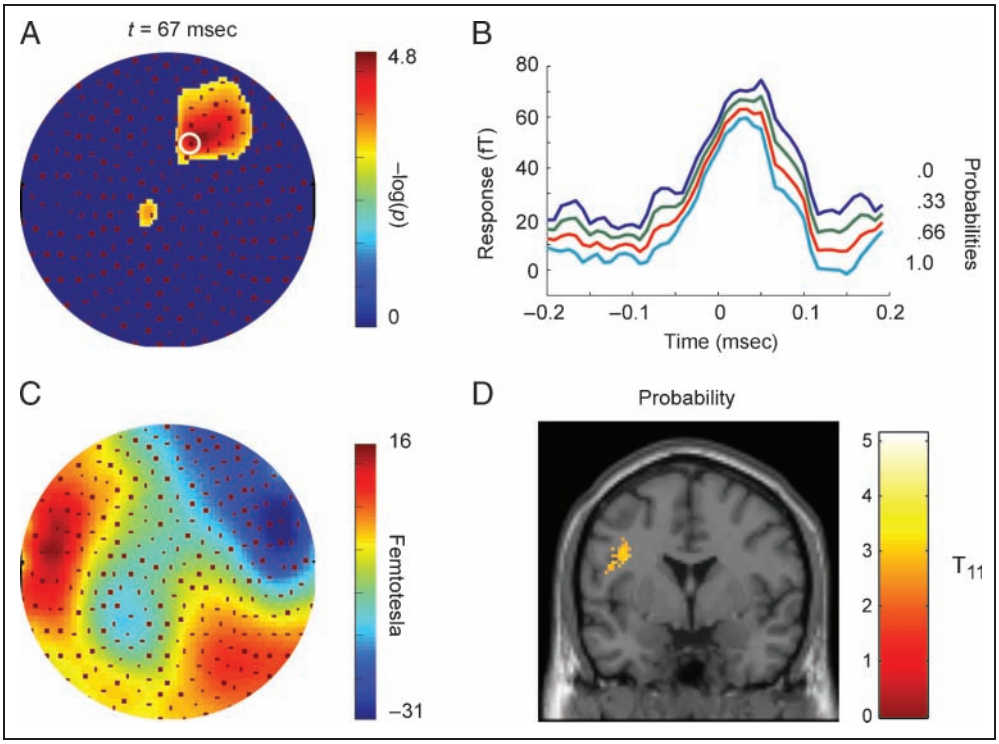

Figure 6. Significant sensor modulation with probability 67 msec after button press. (A) Interpolated map of p values in sensor space. Only p values below .001 are shown, and all values plotted are the negative log of the p value (i.e., a value of 2 is .01). (B) Temporal response at a single sensor (indicated in panel A). (C) Contrast of probability main effect in the GLM. (D) Estimated source of effect. Plot shows a significant cluster of voxels (peak activation: $x = -50$, $y = 6$, $z = 24$) in the premotor cortex.

purely motor effect, which is not modified by learning, whereas the learning effects are related to preparatory and postmotor processing.

As there were differences in how well the hierarchical and nonhierarchical models predicted the behavior of individual subjects, we were interested in whether there would be correlations between the relative fit of the models and how well the sequence model fit the behavior of individual subjects. To test this, we examined the correlation of the relative fit of the models, measured with the log-likelihood ratio, and the contrast estimates for individual subjects. We did not, however, find that there were significant correlations, after corrections for multiple comparisons, in the fit of the models and in the contrast estimates for the sequence model.

Given that these effects were only significant under the sequence model and not the flat model, we next assessed whether the contrast estimates (i.e., the parametric regressors from the GLM) were significantly larger for the sequence model than for the flat model by comparing the distribution of contrast values across subjects between the two models at the peak location, 83 msec before and 67 msec after button press. Neither of these distributions were significantly different ($p = .55$, unequal variance t test, $n = 24$). We did, however, find that there was a significant difference in the variance of the second level contrast distribution at 83 msec before ($p < .05$, F test, $df = 11, 11$) and an almost significant difference at 67 msec after ($p = .056$, F test, $df = 11, 11$) between the sequence model and the flat model. Thus, the increased significance under the sequence model is due to lower variance in the second level contrasts as opposed to a larger mean of the parametric modulator.

Next, we used source localization to identify the possible locations of the sequence probability effects. Significant effects in source space were assessed by estimating distributed activation levels on a cortical mesh for individual subjects and then carrying out second level statistics in SPM on these activation levels. We first assessed the source of the probability effect seen at −83 msec, using a window from −117 to −67 msec. We used a window slightly larger than the window over which the sensor effect was significant as the algorithm rarely converged for small windows. We found a significant source ($p < .05$) bilaterally in the insula, just lateral to the striatum (Figure 5D).

Next, we carried out source localization for the significant effect 67 msec after button press, using a window from 50 to 150 msec after button press. We found a significant source for the probability effect at this time in premotor cortex (Figure 6D; $p < .05$). There was an additional significant source ($p < .05$) in early visual cortex bilaterally (left side $x = -28$, $y = -82$, $z = -14$), perhaps reflecting an attentional effect on the assessment of the feedback and a source at the frontal pole ($x = 12$, $y = 64$, $z = -12$).

DISCUSSION

We examined the behavioral and neural correlates of learning in a stochastic sequence learning paradigm. Comparing flat and hierarchical behavioral models suggested that both predicted the subjects' decisions equally well. However, only the sequence probability from the hierarchical model resulted in significant correlations with the MEG signal, and this significance was due to less

Downloaded from http://direct.mit.edu/jocn/article-pdf/23/6/1346/1775571/jocn.2010.21436.pdf by guest on 18 June 2021

variance in the contrast estimates across subjects. When we localized the learning-related signals, we found a cluster of activity in the insula that preceded the button press and a cluster in the premotor cortex that followed the press.

Behavior

Our sequence task had important features that allowed us to assess how well subjects were learning and whether they were optimal. The learning coefficients show that subjects were not optimal, as positive and negative feedback should have been weighted similarly, whereas in fact subjects relied more on positive feedback as has been seen previously in sequence learning (Averbeck, Sohn, & Lee, 2006). In the task, a hierarchical structure is optimal, whereas a flat structure is not. The behavioral data suggested that subjects that learned better tended to learn in a more hierarchical manner. Perhaps additional training on the sequences would have benefitted the subjects who did not learn as well. Future experiments could clarify this point.

Imaging Results

Two different learning-related signals emerged in the MEG sensor data, one just before button press and one just after button press. The early signal was near the midline, whereas the later signal was lateralized over the right side, although the left side signal may have been just below significance. When we carried out source localization on these two signals, we found a source in the insula for the early signal and a source in premotor cortex for the later signal. Some caution in interpreting these results is necessary, however, as it is difficult to know how precisely the MEG sensor signals can be localized.

Previous fMRI studies have shown activation in the insula during motor learning (Floyer-Lea & Matthews, 2004), and this area has a direct projection to the striatum (Chikama, McFarland, Amaral, & Haber, 1997) and as such it likely takes part in a network of areas related to updating actions on the basis of feedback. Previous work has also shown activity in this area during outcome anticipation that is either negatively or positively valenced (Knutson & Greer, 2008; Volz, Schubotz, & von Cramon, 2004; Critchley, Mathias, & Dolan, 2001). Although the MEG signal that was localized to the insula preceded the button press by 83 msec, it is likely that the press has already been initiated at the cortical level at this time. Therefore, this signal may represent anticipation of either a red or a green outcome, where the anticipation is scaled by how much the subject has learned in the block. Once the sequence is well learned, green feedback is highly likely. It is also interesting that, unlike the signal that follows the button press, this signal was not modulated by the button that was pressed as we did not find an interaction effect in the sensors. This makes it unlikely that this signal was directly involved in learning, as there was no information about the action. This signal may be more related to one's subjective sense of progression through the block, as many studies implicate the anterior insula in subjective interoception (Craig, 2009).

Activation in premotor cortex has been seen in tasks with hierarchical structure (Koechlin & Jubault, 2006; Schubotz & von Cramon, 2003, 2004), and the sequence probability is hierarchical as it represents the entire series of button presses in the order that they unfold. As there is an interaction between sequence probability and the button that was pressed, this signal may have a more direct role in updating the probability information based on the feedback.

A series of studies by Koechlin and colleagues have also suggested that when tasks have explicit hierarchical structure, task factors that map to different levels of the cognitive hierarchy map to different locations in frontal cortex (Badre & D'Esposito, 2007; Koechlin & Summerfield, 2007; Koechlin & Jubault, 2006; Koechlin, Ody, & Kouneiher, 2003). The behavior being studied in their work, however, differs in important ways from the behavior we have studied. Specifically, the studies by Koechlin et al. did not examine decision making in a framework where subjects had to deal with uncertainty about the relationship between actions and outcomes. Rather, the previous studies used rule-based cognitive tasks where the link between stimulus/action/feedback was deterministic given the behavioral rule. Stochasticity in these tasks was implied by the frequency with which the rule changed across blocks. However, all of the information was always provided by the task, and the mapping between actions and outcomes was deterministic. This is very different from the task we have used, which required subjects to deal with uncertainty in an effort to infer the sequence (rule) that was in operation. In our task, even if one knew which sequence was correct in a particular block, one would not be able to predict the feedback on an individual button press. As such, different cognitive processes are likely required to solve our task. The advantage of our approach is that we were able to test directly whether subjects were using hierarchical or flat statistical models when learning the sequences. Thus, we have provided imaging evidence for hierarchical control in a task that could have been solved using a flat model, although the behavioral data were more equivocal. It is not clear what the alternative model would be in the tasks used by Koechlin et al. In their experiments, however, it is less of an issue as they were not studying learning but rather performance of a complex cognitive task.

One important caveat to the model comparison approach we have taken, with respect to the imaging data, is that we have examined significance by linearly correlating the model prediction with the MEG signal (Behrens, Hunt, Woolrich, & Rushworth, 2008; Wittmann et al., 2008; Behrens, Woolrich, Walton, & Rushworth, 2007; Daw et al., 2006). Linear correlations, however, do not allow one to infer conclusively that the underlying neural responses are necessarily favoring one model over the

Downloaded from http://direct.mit.edu/jocn/article-pdf/23/6/1346/1775571/jocn.2010.21436.pdf by guest on 18 June 2021

other. More specifically, the sequence probability is a nonlinear function of the button probability under the flat model. As such, there cannot be more information in the underlying neural response about the sequence variable than there is about the button probability because of the data processing inequality (Cover & Thomas, 1991). Thus, our inference relates only to the specific functional form that we have examined, the linear relationship, and does not tell us about the detailed neural representation of this probability. For this, single-unit studies can be more valuable, partly for practical reasons. It is in many cases possible to examine the relationship between single-unit firing rate responses and various behavioral variables graphically and fit models accordingly. Also, there are often more trials available for fitting more complex models. The high dimensionality of MEG data makes examining the relationship between time varying signals and task variables highly complex. Interestingly, the data from the single-unit studies have consistently shown, in many brain areas and in many similar tasks, that sequence information is explicitly represented in the brain (Averbeck et al., 2006; Tanji, 2001; Nakamura, Sakai, & Hikosaka, 1998).

Comparison with Single-unit Studies

One of the goals of the present study was to approach a question we have already examined in macaques at the single-cell level (Averbeck et al., 2006) in humans using an imaging approach. We had originally intended to use a task that was as similar as possible to the task used in the macaque study. However, unlike macaques that require several trials to learn a three-movement sequence with explicit feedback, human participants given a four-movement sequence learn it in about one trial (unpublished data). This rapid learning makes studying the learning process difficult, and this is why we made the task more difficult by introducing the stochastic feedback.

The second difference has to do with the nature of the information that can be extracted from single-unit data versus MEG imaging data. Specifically, in the macaque study, we were able to track learning by following the emergence of a signal in single neurons that explicitly represented the sequence that was correct in the current block. However, we were not able to extract sequence-specific information from the MEG signal (unpublished data). Thus, we had to use a different approach to examine the learning-related changes in neural activity. Given this difference, however, the premotor signal that follows the button press is in many respects comparable with the location we studied in the macaque, as the activity in the macaque was just anterior to the FEFs, and we used eye movements as our behavioral output in the macaque. Thus, premotor cortex and caudal area 46 may have similar functions for different effectors.

Conclusion

In conclusion, we found that when subjects learned efficiently, they learned hierarchically. Furthermore, the learning-related variable that most strongly correlated with the imaging data was the probability of the sequence, a parameter that is present in the hierarchical model but not in the flat model. Thus, the imaging data and to some extent the behavioral data suggest that when efficient subjects were faced with learning a sequence that had hierarchical structure, they were able to take advantage of that structure.

Reprint requests should be sent to Dr. Bruno B. Averbeck, UCL Institute of Neurology, Sobell Department, Box 28, Queen Square, London WC1N 3BG, UK, or via e-mail: b.averbeck@ion.ucl.ac.uk.

REFERENCES

Andreasen, N. C. (1979a). Thought, language, and communication disorders: I. Clinical assessment, definition of terms, and evaluation of their reliability. *Archives of General Psychiatry, 36,* 1315–1321.

Andreasen, N. C. (1979b). Thought, language, and communication disorders: II. Diagnostic significance. *Archives of General Psychiatry, 36,* 1325–1330.

Averbeck, B. B., Chafee, M. V., Crowe, D. A., & Georgopoulos, A. P. (2002). Parallel processing of serial movements in prefrontal cortex. *Proceedings of the National Academy of Sciences, U.S.A., 99,* 13172–13177.

Averbeck, B. B., Sohn, J. W., & Lee, D. (2006). Activity in prefrontal cortex during dynamic selection of action sequences. *Nature Neuroscience, 9,* 276–282.

Badre, D., & D'Esposito, M. (2007). Functional magnetic resonance imaging evidence for a hierarchical organization of the prefrontal cortex. *Journal of Cognitive Neuroscience, 19,* 2082–2099.

Behrens, T. E., Hunt, L. T., Woolrich, M. W., & Rushworth, M. F. (2008). Associative learning of social value. *Nature, 456,* 245–249.

Behrens, T. E., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience, 10,* 1214–1221.

Bishop, C. M. (1995). *Neural networks for pattern recognition* (1st ed.). Oxford: Oxford University Press.

Chikama, M., McFarland, N. R., Amaral, D. G., & Haber, S. N. (1997). Insular cortical projections to functional regions of the striatum correlate with cortical cytoarchitectonic organization in the primate. *Journal of Neuroscience, 17,* 9686–9705.

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

Craig, A. D. (2009). How do you feel—Now? The anterior insula and human awareness. *Nature Reviews Neuroscience, 10,* 59–70.

Critchley, H. D., Mathias, C. J., & Dolan, R. J. (2001). Neural activity in the human brain relating to uncertainty and arousal during anticipation. *Neuron, 29,* 537–545.

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature, 441,* 876–879.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature, 415,* 429–433.

Downloaded from http://direct.mit.edu/jocn/article-pdf/23/6/1346/1775571/jocn.2010.21436.pdf by guest on 18 June 2021

Fellows, L. K., & Farah, M. J. (2007). The role of ventromedial prefrontal cortex in decision making: Judgment under uncertainty or judgment per se? *Cerebral Cortex, 17,* 2669–2674.

Fiebach, C. J., & Schubotz, R. I. (2006). Dynamic anticipatory processing of hierarchical sequential events: A common role for Broca's area and ventral premotor cortex across domains? *Cortex, 42,* 499–502.

Floyer-Lea, A., & Matthews, P. M. (2004). Changing brain networks for visuomotor control with increased movement automaticity. *Journal of Neurophysiology, 92,* 2405–2412.

Friston, K., Chu, C., Mourao-Miranda, J., Hulme, O., Rees, G., Penny, W., et al. (2008). Bayesian decoding of brain images. *Neuroimage, 39,* 181–205.

Friston, K., Harrison, L., Daunizeau, J., Kiebel, S., Phillips, C., Trujillo-Barreto, N., et al. (2008). Multiple sparse priors for the M/EEG inverse problem. *Neuroimage, 39,* 1104–1120.

Gilovich, T., Griffin, D., & Kahneman, D. (Eds.) (2002). *Heuristics and biases: The psychology of intuitive judgment*. Cambridge, UK: Cambridge University Press.

Glimcher, P. W., & Rustichini, A. (2004). Neuroeconomics: The consilience of brain and decision. *Science, 306,* 447–452.

Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., & Camerer, C. F. (2005). Neural systems responding to degrees of uncertainty in human decision-making. *Science, 310,* 1680–1683.

Jacobs, R. A. (1999). Optimal integration of texture and motion cues to depth. *Vision Research, 39,* 3621–3629.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, UK: Cambridge University Press.

Kersten, D. (1999). High level vision as statistical inference. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences*. Cambridge, MA: MIT Press.

Kiebel, S. J., Kilner, J. M., & Friston, K. J. (2007). Hierarchical models for EEG and MEG. In K. J. Friston, J. T. Ashburner, S. J. Kiebel, T. E. Nischols, & W. D. Penny (Eds.), *Statistical parametric mapping: The analysis of functional brain images* (pp. 211–222). London: Academic Press.

Kilner, J. M., Kiebel, S. J., & Friston, K. J. (2005). Applications of random field theory to electrophysiology. *Neuroscience Letters, 374,* 174–178.

Knill, D. C., & Richards, W. (1996). *Perception as Bayesian inference*. Cambridge, UK: Cambridge University Press.

Knill, D. C., & Saunders, J. A. (2003). Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vision Research, 43,* 2539–2558.

Knutson, B., & Greer, S. M. (2008). Anticipatory affect: Neural correlates and consequences for choice. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 363,* 3771–3786.

Koechlin, E., & Jubault, T. (2006). Broca's area and the hierarchical organization of human behavior. *Neuron, 50,* 963–974.

Koechlin, E., Ody, C., & Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. *Science, 302,* 1181–1185.

Koechlin, E., & Summerfield, C. (2007). An information theoretical approach to prefrontal executive function. *Trends in Cognitive Sciences, 11,* 229–235.

Kording, K. P., & Wolpert, D. M. (2006). Bayesian decision theory in sensorimotor control. *Trends in Cognitive Sciences, 10,* 319–326.

Mattout, J., Henson, R. N., & Friston, K. J. (2007). Canonical source reconstruction for MEG. *Computational Intelligence and Neuroscience*. www.hindawi.com/journals/cln/2007/067613.abs.html.

Mattout, J., Phillips, C., Daunizeau, J., & Friston, K. J. (2007). Bayesian inversion of EEG models. In K. J. Friston, J. T. Ashburner, S. J. Kiebel, T. E. Nichols, & W. D. Penny (Eds.), *Statistical parametric mapping: The analysis of functional brain images*. London: Academic Press.

Mattout, J., Phillips, C., Penny, W. D., Rugg, M. D., & Friston, K. J. (2006). MEG source localization under multiple constraints: An extended Bayesian framework. *Neuroimage, 30,* 753–767.

McKenna, P., & Oh, T. (2005). *Schizophrenic speech: Making sense of bathroots and ponds that fall in doorways*. Cambridge, UK: Cambridge University Press.

Nakamura, K., Sakai, K., & Hikosaka, O. (1998). Neuronal activity in medial frontal cortex during learning of sequential procedures. *Journal of Neurophysiology, 80,* 2671–2687.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1992). Behavioral decision research: A constructive processing perspective. *Annual Review of Psychology, 43,* 87–131.

Penfield, W., & Evans, J. (1935). The frontal lobe in man: A clinical study of maximum removals. *Brain, 58,* 115–133.

Poggio, T., Torre, V., & Koch, C. (1985). Computational vision and regularization theory. *Nature, 317,* 314–319.

Samejima, K., Ueda, Y., Doya, K., & Kimura, M. (2005). Representation of action-specific reward values in the striatum. *Science, 310,* 1337–1340.

Schubotz, R. I., & von Cramon, D. Y. (2003). Functional-anatomical concepts of human premotor cortex: Evidence from fMRI and PET studies. *Neuroimage, 20(Suppl. 1),* S120–S131.

Schubotz, R. I., & von Cramon, D. Y. (2004). Sequences of abstract nonbiological stimuli share ventral premotor cortex with action observation and imagery. *Journal of Neuroscience, 24,* 5467–5474.

Schwartz, M. F., Reed, E. S., Montgomery, M., Palmer, C., & Mayer, N. H. (1991). The quantitative description of action disorganization after brain damage: A case study. *Cognitive Neuropsychology, 8,* 381–414.

Tanji, J. (2001). Sequential organization of multiple movements: Involvement of cortical motor areas. *Annual Review of Neuroscience, 24,* 631–651.

Tobler, P. N., Christopoulos, G. I., O'Doherty, J. P., Dolan, R. J., & Schultz, W. (2008). Neuronal distortions of reward probability without choice. *Journal of Neuroscience, 28,* 11703–11711.

Todorov, E., & Jordan, M. I. (2002). Optimal feedback control as a theory of motor coordination. *Nature Neuroscience, 5,* 1226–1235.

Trommershauser, J., Maloney, L. T., & Landy, M. S. (2003a). Statistical decision theory and the selection of rapid, goal-directed movements. *Journal of the Optical Society of America, A, Optics and image science, 20,* 1419–1433.

Trommershauser, J., Maloney, L. T., & Landy, M. S. (2003b). Statistical decision theory and trade-offs in the control of motor response. *Spatial Vision, 16,* 255–275.

Trommershauser, J., Maloney, L. T., & Landy, M. S. (2008). Decision making, movement planning and statistical decision theory. *Trends in Cognitive Sciences, 12,* 291–297.

Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of Business, 59,* S251–S278.

Volz, K. G., Schubotz, R. I., & von Cramon, D. Y. (2004). Why am I unsure? Internal and external attributions of uncertainty dissociated by fMRI. *Neuroimage, 21,* 848–857.

Wittmann, B. C., Daw, N. D., Seymour, B., & Dolan, R. J. (2008). Striatal activity underlies novelty-based choice in humans. *Neuron, 58,* 967–973.

Downloaded from http://direct.mit.edu/jocn/article-pdf/23/6/1346/1775571/jocn.2010.21436.pdf by guest on 18 June 2021