**Fig S1. Related to Figure 3; Trial-by-trial correlations.** The left panel shows production intervals as a function of buildup rates, and the right panel shows buildup rates as a function of Set activity for all trials for which the buildup rate could be estimated. To combine values across neurons, variables were first normalized (zscore) for each neuron independently. The marginals of the the variables on the ordinate and abscissa are shown on the left and below each panel. The two ellipses (red) represent a 2D gaussian fit to the data points at 1 and 2 standard deviations from the mean.

## Supplemental Experimental Procedures

Monkeys were surgically implanted with a titanium head-holding device attached to the skull. During the experimental sessions, monkeys were comfortably seated in a dark quiet room with their heads fixed, and viewed stimuli binocularly from a distance of 57 cm. All stimuli were presented on a fronto-parallel 15-inch Hewlett Packard HP71 CRT monitor at a resolution of 1024x768 at a refresh rate of 85 Hz. We used the software Expo (http://corevision.cns.nyu.edu/) to deliver stimuli and control behavioral contingencies. Eye positions were sampled at 1kHz from signals recorded with an infrared camera (Eyelink 1000; SR Research Ltd.; Ontario, Canada). We recorded action potential extracellularly through a craniotomy over the right intraparietal sulcus with either glass-coated tungsten electrodes (Alpha Omega Co.; Alpharetta, GA, USA)

or quartz-platinum/tungsten microelectrodes (Thomas RECORDING GmbH; Giessen, Germany). Eye movement and extracellular signals were stored using a standard data acquisition system (Plexon Inc.; Dallas, Texas, USA).

## Behavioral task

In the main RSG task, each trial began with the appearance of the central fixation point (FP) that monkeys were required to fixate. While fixating and after a 500 ms delay, the saccade target was presented. After a variable foreperiod (0.2-2.0 sec, truncated exponential distribution), two 105-ms flashes (Ready and Set) were presented. Typically, Ready was presented diametrically opposite to the saccade target, and Set was presented 4 to 10 deg away from the FP, equidistant from Ready and the saccade target. Monkeys were trained to measure the sample interval ($t_s$) between Ready and Set, and reproduce that interval by a proactive saccade to the target. Production interval ($t_p$) was measured from midpoint during the flashed Set (i.e. ~ 50 ms after its onset) to when the monkey acquired the target.

When the difference between $t_s$ and $t_m$ was smaller than an experimentally specified window, the color of the saccade target changed to green and monkeys received an immediate liquid reward. To compensate for the increased variability associated with longer $t_s$ [S1,S2] the width of the rewarded window was scaled with the sample interval. The scaling constant was adjusted so that approximately 60% of trials were rewarded. The rewards were graded linearly such that the volume was largest for accurate reproduction and declined to 0 at the boundary of the acceptance window. Monkeys' behavior was stable and consistent in all recorded sessions. In some sessions, we increased reward rate near the end of the session by increasing the width of reward window by nearly 10% to encourage the monkey to complete more trials. These manipulations did not have a significant effect on the average root-mean-squared-error (RMSE) of production intervals.

The RG task we used as a control experiment was identical to the RSG task in sensory and motor events but required monkeys to make a saccade to the target 1588 ms after the Ready, irrespective of when Set was presented. The duration of Ready-Set was drawn from the same discrete uniform distribution that was used in the RSG task. The rewarded window in the RG task was also scaled with the production interval.

## Bayesian model for the RSG task

Our Bayesian model follows directly from our work on humans performing the RSG task [S3]. Briefly, a Bayesian observer makes a noisy measurement of the $t_s$. We modeled the measurement noise as a random variable with a zero-mean Gaussian distribution whose standard deviation scales with the $t_s$ (i.e., scalar variability). Accordingly, the relationship between the measured interval, $t_m$, and sample interval, $t_s$, can be written as:

(1)

$$p(t_m|t_s) = \frac{1}{\sqrt{2\pi(w_m t_s)^2}} e^{-\frac{(t_m - t_s)^2}{2(w_m t_s)^2}}$$

where $p(t_m|t_s)$ denotes the conditional probability of the measured interval ($t_m$) as a function of $t_s$, and $w_m$ is the Weber fraction associated with scalar variability in the measurement process.

The Bayesian observer combines the probabilistic information conveyed by $t_m$ with knowledge about the distribution of sample intervals, $p(t_s)$, which is a uniform distribution ranging between 529 and 1059 ms, to compute the posterior distribution over $t_s$:

(2)

$$p(t_s|t_m) = \frac{p(t_m|t_s)p(t_s)}{p(t_m)}$$

Following previous work on humans [S3], we assumed that the observer uses the posterior to infer an optimal estimate, $t_e$, that minimize the squared difference between the sample and estimate. In other words, we considered a quadratic cost function $(t_e - t_s)^2$, which is associated with a Bayes Least Squares (BLS) estimator. The BLS estimator uses the mean of the posterior as its optimal estimate:

(3)

$$t_e = f_{BLS}(t_m) = \int t_s p(t_s|t_m) dt_s$$

The Bayesian observer then attempts to reproduce the estimate, $t_e$. We assumed that the production process is also noisy, and modeled the production noise as a random variable with a zero-mean Gaussian distribution whose standard deviation scales with the estimate (i.e., scalar variability):

(4)

$$p(t_p|t_e) = \frac{1}{\sqrt{2\pi(w_p t_e)^2}} e^{-\frac{(t_p - t_e)^2}{2(w_p t_e)^2}}$$

where $p(t_p|t_e)$ denotes the conditional probability of the production interval, $t_p$, as a function of $t_e$, and $w_p$ is the Weber fraction associated with production process that may be different from $w_m$.

We fitted this model to the monkeys' behavior by maximizing the likelihood of the model parameters, $w_m$ and $w_p$, across all $t_s$ and $t_p$ values.

We quantified the monkeys' behavior and the corresponding behavior of the Bayesian model using two parameters, the average bias and the average variance. These two metrics, which we denote them by *BIAS* and *VAR*, respectively, are directed related to

the overall mean-squared error (MSE) of the production intervals as follows:

(5)

$$MSE = \sum_{j=1}^{M}\sum_{i=1}^{N}(t_p^{i,j} - t_s^j)^2 = \sum_{j=1}^{M} VAR(t_p^j) + BIAS^2(t_p^j) = VAR + BIAS^2$$

where index $j$ specifies the value of $t_s$ $(t_s = t_s^j)$, and $t_p^{i,j}$ refers to the production time in the $i^{th}$ repetition of $t_s = t_s^j$. As indicated by this equation, MSE can be written as the sum of two terms: the sum of squares of bias (i.e., *BIAS²*) and the sum of variances (i.e., *VAR*).

**Bayesian model for the RG task**

In the RG task, the animals' behavior was influenced by the time of the Set cue. This influence seems perplexing because the monkey is rewarded for reproducing a 1588 ms interval on all trials, regardless of the intervening Set cue. There seem to be two general possibilities for explaining this influence: Set presents itself as a distractor, or alternatively, the monkey actively uses the Ready-Set interval as a cue to improve performance. We can eliminate the first possibility by examining the variability of the production intervals in the RG task. If Set were to act as a distractor, it would increase the variability, but it appears to do the opposite. We base this conclusion on the following analysis.

We computed the predicted standard deviation of the $t_p$ in the RG task, which is equal to the mean $t_p$ multiplied by $w_p$. Our estimate of $w_p$ from the fits of the Bayesian model to the behavior in the RSG task ($w_p$ = 0.14) predicted the standard deviation of $t_p$ in the RG task to be ~224 ms. In comparison, the standard deviation we observed in the RG task was ~150 ms, which is approximately 33% below expectation. It is therefore, not plausible to assume that the Set was a simple distraction. Instead, monkeys seemed to

have used the Ready-Set interval to reduce variability in their behavior.

It may seem at first unintuitive that using the time of the Set cue would improve performance, but in fact it could. To understand why, it is useful to consider a hypothetical experiment in which the Set occurs at the same time on every trial. In that case, the superior strategy is to rely on Set because it would allow the monkey to produce the shorter time from Set to the target 1588 ms, which will be less variable, as per Weber's law. For example, if $t_s$ were fixed to half the 1588 ms (i.e., 794 ms), the subject could learn to wait until the Set, and then produce a fixed 794 ms, which would make production times twice as accurate as timing 1588 ms from the Ready cue.

What about the case in which the Ready-Set is drawn from a distribution? In that case, an optimal Bayesian observer would still not ignore the Set altogether; instead it would used it probabilistically and estimate the Ready-Set interval by integrating the observed interval with the prior distribution from which it is drawn. The benefit is of course attenuated by the additional uncertainty of the prior distribution but it is not eliminated altogether, so long as the prior distribution can be exploited. In other words, the same observer model we used to explain the behavior in the RSG task could also explain the effect of Set on the behavior of the two monkeys in the RG task. The only difference is that, in the RG task, the observer reproduces $1588^{ms} - t_e$.

**Electrophysiology**

A 9 Kg male (Yo) and a 6.8 Kg female (Wi) rhesus monkey (Macaca mulatta), both 7 years old, participated in this experiment. We recorded from 97 well-isolated LIP neurons in the ventral portion of area LIP (LIPv) in the right hemisphere of two monkeys (Yo and Wi), 58 in the RSG task (35 in Yo, and 23 in Wi), and 39 in the RG task (26 in Yo, and 13 in Wi). LIPv was identified based on (1) anatomical landmarks along the IPS [S4] and in registration with structural MRI scans distributed with the CARET software package [5] and (2) transition of white and gray matter during recording.

Most LIP neurons have oculocentric response fields (RF) [S6] and their firing rate

increases when monkeys plan a saccadic eye movement to the RF [S7]. Many LIP neurons also exhibit persistent activity in the memory period of the memory saccade task [8,9]. We reasoned that neurons with persistent activity would be a good candidate for studying temporal anticipation and planning, as they are capable of reflecting the animal's internal state in the absence of direct visual input [S10,11]. We thus recorded from cells that had spatially selective sustained activity in the memory period of a memory-saccade task [S8]. The cell's response during the memory period of the memory saccade task had to be at least 50% larger than firing rate for the memory period of saccades made to a target opposite to cells' response field (RF). We recorded from neurons with RF located at eccentricities ranging between 3.8 to 20 deg. For RFs located at eccentricities over 14 deg, we moved the fixation point to ensure that the target was not too close to the edge of the monitor. During the RSG and RG tasks, the saccade target was placed in the central part of the neuron's RF (ranging between 4 and 20 deg across sessions), and spiking activity was recorded. Ready and Set were displayed outside the RF. For many cells, we verified the stability of recordings by monitoring the spatial selectivity and persistent activity in the memory-saccade task intermittently throughout the session.

**Data Analysis**

We estimated response strength at the time of Set by averaging firing rates within a window of 50 ms before to 50 ms after Set. We estimated the slope of the ramping activity in the production epoch from firing rates within a window of 500 to 200 ms before saccade initiation. Results were qualitatively the same for moderate variations of the analysis windows (tested up to 50 ms). For each trial, we fitted a first-order rate function of an inhomogeneous Poisson process to each trial's spike times (maximum-likelihood procedure), and used the linear term of the fitted rate function as our estimate of the slope for that trial.

We used linear regression to quantify the correspondence between LIP activity and the

sample interval. To estimate the slope of the regression line, we used ordinary least squares for single cell analyses and weighted least squares for the population analysis. For the weighted least squares regression analysis, the error term associated with each cell was weighted by the reciprocal of the standard error of that cell's mean firing rate. Confidence intervals and p values for the fits were derived from the standard error of the slope of the regression line [S12].

**Models**

We developed a series of models to understand which of several candidate quantities would best capture the observed nonlinear response dynamics (firing rate vs. time) during the measurement epoch of the RSG task. In each case, we computed the quantity of interest as a function of time, $f(t)$, and developed a model in which the firing rate, $y(t)$, was a linear function of $f(t)$; i.e. $y(t) = Af(t)+B$. Importantly, the linearity of this transformation ensures that the response nonlinearlities can only be explained by the quantity of interest and not the parametric form of the model. Furthermore, it ensues that all models have the same number of parameters ($A$ and $B$) that simply control the scaling and offset of the fit to the data.

*1. Anticipation of Set.* To test this model, we computed the probability distribution of measured intervals, $p(t_m)$, from the prior probability, $p(t_s)$, and $p(t_m|t_s)$, as described in the Bayesian fit to the behavior (Equation 1), and used the hazard of $p(t_m)$ as a subjective hazard for Set, $h_{Set}(t)$.

(6)

$$p(t_m) = \int p(t_m|t_s)p(t_s)dt_s$$

(7)

$$h_{Set}(t) = \frac{p(t_m)}{1 - P(t_m)}$$

where *p(tm)* and *P(tm)* refer to the probability density and distribution function, respectively. We then modeled the firing rate, *y(t)*, as a linear function of $h_{Set}(t)$, and computed the best fit of this model to activity from 300 ms after Ready to the time of Set. Our conclusions were the same for a variant of this model in which firing rate are modeled as a linear function of the objective hazard (computed from $p(t_s)$ instead of $p(t_m)$).

(8)

$$y(t) = Ah_{Set}(t) + B$$

*2. Anticipation of Reward.* To test this model, we first computed the expected reward as a function of time in terms of the hazard of Set; i.e., the probability of receiving reward, given that Set has not yet occurred. This probability provides an empirical measure of the probability of reward as a function of time during the measurement epoch. At the beginning of the measurement epoch and until the time of the earliest possible Set, all sample intervals are equally probable, and therefore, the probability of reward is equal to its grand average across all trials. As time elapses beyond the earliest possible Set time, the shortest sample interval is no longer possible,  and the probability of reward gets updated to the average of all trials except those associated with the shortest sample intervals. This updating process continues after each possible sample interval so that the probability of reward dynamically reflects the animals' expectation based on the remaining probable sample intervals. If we denote the proportion of rewarded for trials as a function of the sample interval by $g_{Rew}(t_s)$, this hazard-like function for reward, $h_{Rew}(t)$, can be written mathematically as follows:

(9)

$$h_{Rew}(t) = \frac{\sum_{t_s > t} g_{Rew}(t_s) p(t_s)}{\sum_{t_s > t} p(t_s)}$$

This function has a piecewise linear form as it gets updated only at discrete times associated with each sample interval. We used a linear interpolation scheme to smooth the function, and then modeled firing rate, $y(t)$, as a linear function of this interpolated function. We computed the best fit of this model to activity from 300 ms after Ready to the time of Set. Our conclusions were the same for a variant of this model in which we replaced the proportion of rewarded trials by the expected reward magnitude computed as the reward magnitude multiplied by the probability (recall that, when the animal was rewarded, the magnitude of reward was dependent on accuracy).

(10)

$$y(t) = Ah_{Rew}(t) + B$$

*3. Anticipation of the expected time of reward.* In the production epoch of the RSG task, and throughout the RG task, activity increased linearly with a slope that allowed firing rate across trials to reach a terminal value shortly before the expected time of reward (and saccade). Motivated by this observation, we developed a model of LIP activity in the measurement epoch in which the instantaneous slope was adjusted to create a ramp that would reach the terminal firing rate at the expected time of reward. At the beginning of the measurement epoch and until the time of the earliest Set, all sample intervals are equally probable, and therefore, the expected time of reward is the twice as long as the mean sample interval. As time elapses, some of the sample intervals are no longer probable, and therefore, the expected time of reward gets progressively delayed to reflect the expected time of reward for the remaining possible sample intervals. Mathematically, this model can be written as:

(11)

$$\frac{dr}{dt} = \frac{R_{max} - r}{E[t_{Rew}|t] - \Delta t_{Sacc} - t}$$

in which $dr/dt$ is the instantaneous slope, $R_{max}$ is an arbitrary terminal (maximum) firing rate, $E[t_{Rew}|t]$ is the expected time of reward given time $t$, and $\Delta t_{Sacc}$ is the fixed delay

between when firing rate reached $R_{max}$ and when reward is delivered. We inferred the expected time of reward empirically from the animals' production intervals, as follows:

(12)

$$E[t_{Rew}|t] = \frac{\Sigma_{t_s>t} E[t_p|t_s]p(t_s)}{\Sigma_{t_s>t}p(t_s)}$$

where $E[t_p|t_s]$ is the expected (average) production interval for sample interval, $t_s$. To estimate $\Delta t_{Sacc}$, we noted that when average firing rates reached their common terminal firing, the response variance was at its minimum. Therefore, we computed the earliest time when response variance was within 2 standard deviation of this minimum value.

To compute response variance, we first constructed peri-event time histograms (PETH) for each neuron's spike times within a window of 400 ms before saccade initiation using a 100 ms boxcar filter. We then subtracted from PETH, the average firing rate for each time point across trials to compute residual (i.e., fluctuations) of firing rates around mean (PETH$_{res}$). To estimate the minimum variance, we combined PETH$_{res}$ across neurons, sorted trials into 20 bins based on the production intervals, measured average PETH$_{res}$ within each sorted bin, and computed the variance across the bins. To compute the standard deviation of the variance, we used 100 bootstrap estimates of the variance by resampling from the combined PETH$_{res}$ across cells and trials. Finally, we determined the earliest point before saccade initiation when the average variance was within 2 standard deviations of the minimum variance (i.e., the time point at which firing rates had reached their terminal value).

We then solved Equation 11 for $r(t)$, and modeled the observed firing rate, $y(t)$, as a linear function of $r(t)$, and computed the best fit of this model to activity from 300 ms after Ready to the time of Set.

(13)

$$y(t) = Ar(t) + B$$

Our conclusions were insensitive to reasonable variations of analysis parameters including 80 to 200 bootstraps, 50 to 150 ms boxcar filtering of the PETHs, and different measures of significance for change of variance (1 to 3 standard deviations away from the average variance).

*4. Bayesian estimate of the sample interval.* For this scheme, we modeled the firing rate as a linear function of the *BLS* estimate (Equation 3) and computed the best fit of this model to activity from 300 ms after Ready to the time of Set.

(14)

$$y(t) = Af_{BLS}(t) + B$$

*5. Preplanning the production dynamics.* According to this model, response dynamics in the measurement epoch provide a moment-to-moment (if Set were to occur now) prediction of the build-up rate in the ensuing production epoch. To formalize this model, we assumed that the instantaneous slope of the activity in the measurement epoch is linearly related to the slope of the build-up rate in the production epoch, which is determined by the ratio of a fixed firing rate excursion, $\Delta r$, to the duration of the build-up, which begins $\Delta t_{Set}$ after Set, and continues until $\Delta t_{Sacc}$ before saccade initiation. We can write this model mathematically as follows:

(15)

$$\frac{dr}{dt} = \frac{\Delta r}{t_p - \Delta t_{Sacc} - \Delta t_{Set}}$$

We estimated $\Delta t_{Sacc}$ as described above for model 3 (Anticipation of the expected time of reward). For $\Delta t_{Set}$, we used the same exact strategy: we first determine that responses reached their lowest variance approximately 150 ms after Set, and proceeded by looking for the latest point in time after Set that response variance was still within two standard deviation of the minimum value.

We then solved Equation 15 analytically for *r(t)*, and modeled the observed firing rate, *y(t)*, as a linear function of *r(t)*, and computed the best fit of this model to activity from 300 ms after Ready to the time of Set.

(16)

$$y(t) = Ar(t) + B$$

Our conclusions were insensitive to reasonable variations of analysis parameters including 80 to 200 bootstraps, 50 to 150 ms boxcar filtering of the PETHs, and different measures of significance for change of variance (1 to 3 standard deviations away from the average variance).

S1. Gallistel CR, Gibbon J. Time, rate, and conditioning. Psychol Rev. 2000;107: 289–344.

S2. Rakitin BC, Gibbon J, Penney TB, Malapani C, Hinton SC, Meck WH. Scalar expectancy theory and peak-interval timing in humans. J Exp Psychol Anim Behav Process. 1998;24: 15–33.

S3. Jazayeri M, Shadlen MN. Temporal context calibrates interval timing. Nat Neurosci. Nature Publishing Group; 2010;13: 1020–1026.

S4. Lewis JW, Van Essen DC. Mapping of architectonic subdivisions in the macaque monkey, with emphasis on parieto-occipital cortex. J Comp Neurol. Wiley Online Library; 2000;428: 79–111.

S5. Van Essen DC. Windows on the brain: the emerging role of atlases and databases in neuroscience. Curr Opin Neurobiol. 2002;12: 574–579.

S6. Colby CL, Duhamel JR, Goldberg ME. Oculocentric spatial representation in parietal cortex. Cereb Cortex. 1995;5: 470–481.

S7. Bracewell RM, Mazzoni P, Barash S, Andersen RA. Motor intention activity in the macaque's lateral intraparietal area. II. Changes of motor plan. J Neurophysiol. American Physiological Society; 1996;76: 1457–1464.

S8. Gnadt JW, Andersen RA. Memory related motor planning activity in posterior parietal cortex of macaque. Exp Brain Res. Springer-Verlag; 1988;70: 216–220.

S9. Andersen RA, Snyder LH, Bradley DC, Xing J. Multimodal representation of space in the posterior parietal cortex and its use in planning movements. Annu Rev Neurosci. 1997;20: 303–330.

S10.    Shadlen MN, Gold JI. The neurophysiology of decision-making as a window on cognition. The cognitive neurosciences. Citeseer; 2004;3: 1229–1441.

S11.    Shadlen MN, Kiani R. Decision making as a window on cognition. Neuron. 2013;80: 791–806.

S12.    Draper N, Smith H. 1966. Applied Regression Analysis. John Wiley & Sons, Inc.