SUPPORTING ONLINE MATERIAL

Materials and Methods

Two adult male rhesus monkeys (*Macaca mulatta*) were used in this study. Prior to physiological experiments each animal was prepared surgically with a head-holding device (*S1*), scleral search coil for monitoring eye position (*S2*) and a recording chamber over the intraparietal sulcus to allow access to area LIP. Each monkey was then trained on both the delayed saccade and matching tasks. During training and while engaged in experiments, daily water intake was controlled to maintain adequate levels of motivation. All surgical, behavioral, and animal care procedures complied with the U.S. Department of Health and Human Services (National Institutes of Health) Guide for the Care and Use of Laboratory Animals (1996).

Area LIP was identified based upon its anatomical location and on the characteristic physiological response properties of its cells and those of neighboring areas (*S3-S5*). In Monkey F, localization was aided by anatomical Magnetic Resonance Imaging studies. At the time of this report, both monkeys remain actively engaged in experiments, so precise histological identification of recording sites is not yet available. We employed standard methods to record the discharge of isolated single neurons using extracellular tungsten microelectrodes (FHC Inc). Real time experimental control was implemented in the *Rex* (*S6*) environment for the Qnx operating system, running on a PC compatible microcomputer. Visual stimuli were generated using the VSG graphics card (Cambridge Graphics) housed in a second PC compatible computer, and presented on a CRT display. After amplification, single unit spiking activity was identified and collected using either a dual voltage-time window discriminator (Bak Electronics) or the *Plexon* (Plexon Inc) data acquisition system operating in conjunction with *Rex*. Data pertaining to the timing of behavioral and task events, eye position, and single unit spike times were all digitized and recorded to disc for later offline analysis using custom software written in the MATLAB programming environment, running on Apple Macintosh computers.

Quantitative analysis of relationship between choice and reward history

Figure 1C in the main text qualitatively demonstrates the local temporal dependence of choice in our matching task on past rewards. To quantify this relationship we can turn to signal-processing methods such as cross-correlation or Weiner kernel analysis, which reveal the time course of any linear relationship between two time series (*S7*). Through these methods we can infer the form and temporal extent of the best linear operator that relates recent reward history to current choice. This problem is very similar to the problem faced by sensory neurophysiologists in relating neural responses to antecedent sensory stimuli, where the technique of spike-triggered averaging (STA) is commonly applied (*S8-S9*). As a direct analogy to this technique, we can employ a form of ‘choice-triggered averaging’ (CTA), to estimate the relationship between choice and preceding rewards. Both of these approaches are ultimately special cases of the more general technique of Wiener kernel analysis.

Conceptually, the choice triggered averaging procedure is quite simple. Consider the reward history that immediately precedes a particular choice. For each choice of that same color, there is an analogous 'choice-triggered' history. The average of these histories is, in a sense, the prototypical reward history that precedes choices of that color. If the time series of rewards had zero mean and was free of correlations (like the Gaussian white noise stimuli used in STA studies (*S10*)), this choice-triggered average history would be directly proportional to the best linear filter relating rewards to choice. In the case of our matching task, however, the blockwise changes in average reward rates introduce correlations in the time series of rewards whose influence must be removed to arrive at an unbiased estimate of this optimal linear filter.

Formally, we can use the Weiner-Hopf theorem (*S7*) to remove the influence of these autocorrelations from the CTA and estimate the optimal filter. Given the time series of rewards $r(t)$ and choices $c(t)$, Weiner-Hopf reconstructs the linear filter $h(l)$ that relates them by minimizing the squared error in predicting one time series from the other

$$E = \left\| c(t) - \sum_{l=1}^{\infty} h(l) r(t-l) \right\|^2 .$$

The filter that minimizes this error satisfies the Wiener-Hopf equation

$$\Theta_{cr}(i) = \sum_{l=1}^{m} h(l) \Theta_{rr}(i-l) ,$$

where Θ_{cr} is the cross-correlation between $c(t)$ and $r(t)$, $h(l)$ is the causal filter of length m, and Θ_{rr} is the autocorrelation of $r(t)$. We can compute the best filter relating choices to rewards, $h(l)$, by simply rewriting the previous equation in matrix form, and inverting the autocorrelation matrix of rewards.

$$\vec{h} = \Theta_{rr}^{-1} \vec{\Theta}_{cr}$$

In this formulation, the cross-correlation Θ_{cr} is the uncorrected CTA described above, while the inverted autocorrelation matrix Θ_{rr} is a correction term that accounts for the temporal structure inherent in the time series of rewards. This correction removes the influence of these correlations from the CTA, and reconstructs the best linear filter relating rewards to choice.

Figure S1 illustrates the corrected CTAs computed using the above approach, for each of the two monkeys in our study. In this context, one can interpret the CTA as measuring the influence of preceding rewards on the monkey's current choice. The precipitous monotonic decay in the magnitude of this filter as a function of time is evidence for an integration process that is, in fact, highly local in time. Importantly, the time scale over which this analysis shows effects of reward history corresponds very well with the time scale of integration suggested by our model fitting in Figure 2C of the main text. The shape of the CTA, however, somewhat departs from the single exponential filter employed in our model, suggesting the potential refinement in future models of the transformation between reward history and choice.

Obviously, the linear influence of reward history captured by the CTA is not a complete description of choice behavior in our task. Choice behavior in animals has strong nonlinear and stochastic aspects, which the CTA cannot directly capture (*S8*). As such, choice-triggered-averaging is best viewed as a useful descriptive tool, whose results provide an independent means of quantifying the short temporal window over which our animals integrate reward information.

References

S1. E. V. Evarts, *Methods Med. Res.* **11**, 241 (1968).

S2. S. J. Judge, B. J. Richmond, F. C. Chu, *Vision Res.* **20**, 535 (1980).

S3. J. W. Gnadt, R. A. Andersen, *Exp. Brain Res.* **70**, 216 (1988).

S4. S. Barash, R. M. Bracewell, L. Fogassi, J. W. Gnadt, R. A. Andersen, *J. Neurophysiol.* **66**, 1095 (1991).

S5. R. A. Andersen, in *Handbook of Physiology: The Nervous System*, V. B. Mountcastle, F. Plum, S. R. Geiger, Eds. (Amer. Physiol. Soc., Baltimore, MD, 1987), vol. 5, chap. 12.

S6. A. V. Hays, B. J. Richmond, L. M. Optican, *WESCON Conf. Proc.* **2**, 1 (1982).

S7. T. K. Moon, W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing* (Prentice Hall, Upper Saddle River, NJ, 2000), pp. 157.

S8. E. P. Simoncelli, L. Paninski, J. Pillow, O Schwartz, in *The New Cognitive Neurosciences*, M. Gazzaniga, Ed. (MIT Press, Cambridge, MA, 2004).

S9. P. Dayan, L. F. Abbott, *Theoretical Neuroscience* (MIT Press, Cambridge, MA, 2001).

S10. E. J. Chichilnisky, *Network: Computation in Neural Systems*, **12**, 199 (2001).

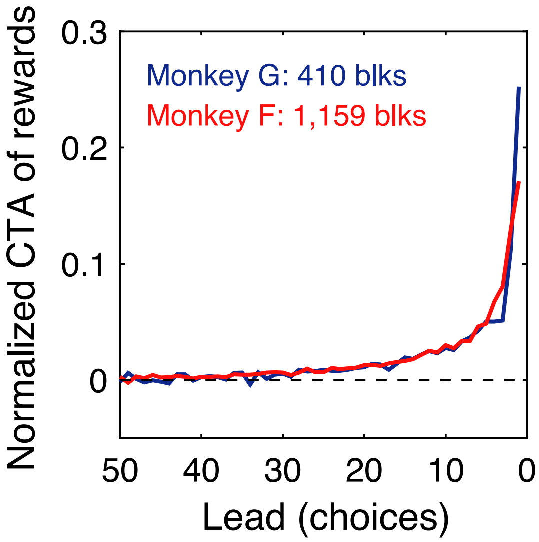

Figure S1. Quantifying the local relationship between choice and rewards. Corrected CTA quantifying the dependence of current choice on preceding rewards. Ordinate: the choice triggered average of rewards (normalized to have an integrated area of one). Abscissa: the temporal offset (in choices) at which this average applies. The CTA at each offset is directly proportional to the corresponding weight of the best linear filter relating the time series of choices and rewards. The dashed line at zero corresponds to the relationship expected by chance. Blue curve: CTA from 410 blocks of data from monkey G. Red curve: CTA from 1,159 blocks of data from monkey F.