J Neurophysiol 98: 1263–1277, 2007.
First published June 27, 2007; doi:10.1152/jn.00016.2007.

Quantitative Comparison Between Neural Response in Macaque Inferotemporal Cortex and Behavioral Discrimination of Photographic Images

Sarah R. Allred and Bharathi Jagadeesh
Program in Neurobiology and Behavior, Department of Physiology and Biophysics, University of Washington, Seattle, Washington

Submitted 5 January 2007; accepted in final form 26 June 2007

Allred SR, Jagadeesh B. Quantitative comparison between neural response in macaque inferotemporal cortex and behavioral discrimination of photographic images. *J Neurophysiol* 98: 1263–1277, 2007. First published June 27, 2007; doi:10.1152/jn.00016.2007. Inferotemporal (IT) cortex plays a critical role in the primate ability to perceive and discriminate between images, but the relationship between responses of single neurons and behavioral capacities is poorly understood. We studied this relationship by recording from IT neurons while monkeys performed a delayed-match-to-sample task with two images. On each day, two sample images were chosen to maximize the selectivity of the neuron and task difficulty was manipulated by varying sample duration and by masking the sample. On each trial, monkeys reported which of the two sample images was presented. Neural performance was described using an ideal-observer analysis. Across the population, neural and behavioral sensitivity to changes in sample duration were indistinguishable. Neural sensitivity was dependent on epoch used to analyze neural response; maximal neural sensitivity was achieved in the 128-ms epoch that began 85 ms after sample onset. At most sample durations, the epoch that yielded optimal neural performance was longer than the sample duration, suggesting that neural selectivity persisted after the presentation of the mask during performance of the task. A control experiment showed that neural and behavioral performance improved in the absence of the mask. These observations suggest that the responses of individual IT neurons contain sufficient information to allow behavioral discrimination of images in a demanding task.

INTRODUCTION

The ability to recognize and discriminate between complex objects is a signature feature of the primate visual system, and the responses of neurons in the inferotemporal (IT) cortex are thought to underlie this ability. However, the relationship between responses of individual IT neurons to complex images and behavioral choices made about those images is poorly understood. Results from previous studies indicate that IT neurons can modulate their responses along particular feature dimensions that are most relevant to primate behavior (Logothetis et al. 1995; Sigala and Logothetis 2002; Vogels et al. 2001). In parametric stimulus sets, IT neurons also exhibit graded selectivity that is determined by the relative location of stimuli in that parameter space (Freedman et al. 2003; Leopold et al. 2006; Op de Beeck et al. 2001). In addition, trial-by-trial variations in IT response can predict behavioral responses to ambiguous stimuli (Sheinberg and Logothetis 1997; Uka et al. 2005). Taken together, these results suggest that IT plays a critical role in behavioral discrimination of objects. To understand more completely the relationship between IT responses and behavior, we have examined how the spiking activity of neurons can be quantitatively related to behavior in a demanding delayed-match-to-sample (DMS) task. In this study, monkeys performed a threshold psychophysics task with many different images. We analyzed the responses of individual IT neurons to these images while monkeys performed the task to investigate whether the sensitivity of individual IT neurons could account for behavioral performance.

Monkeys performed an object-discrimination task by reporting which of two images was presented during a sample period in a DMS task. The DMS task has been used to probe a variety of visual abilities (Cahusac et al. 1989; Chelazzi et al. 1998; Sereno and Maunsell 1998). To increase task difficulty, we shortened sample duration and added a mask; these manipulations decreased available sensory information and elicited a range of behavioral performance that spanned psychophysical thresholds. Previous studies have shown that masking briefly presented images reduces neural selectivity (Bacon-Mace et al. 2005; Keysers et al. 2001; Kovacs et al. 1995; Rolls et al. 1999) and impairs discrimination performance in monkeys (Kovacs et al. 1995) and humans (Grill-Spector and Kanwisher 2005; Keysers et al. 2001; Rieger et al. 2005; VanRullen and Thorpe 2001).

Previous studies of responses in IT to masked stimuli have suggested that when images are presented for short durations and masked, the length of epoch that defines maximum neural sensitivity is similar to the length of stimulus presentation (Kovacs et al. 1995; Rolls et al. 1999). In these previous studies, data were collected while monkeys performed passive fixation tasks. Task demands are known to modulate the responses of IT neurons (Moran and Desimone 1985; Sheinberg and Logothetis 2001). In this study, monkeys engaged in a demanding task while we simultaneously collected neural data, and we investigated the epoch in which neurons achieved maximum selectivity and how that selectivity related to behavioral performance. Compared with previous studies, we found higher neural sensitivity for longer epochs, with the length of epoch for optimal neural performance significantly outlasting the length of sample presentation in most conditions. These results suggest that IT neurons may, in some conditions, temporally filter or limit the impact of masking stimuli on visual responses.

METHODS

Subjects

Two male adult monkeys (*Macacca mulatta*) served as subjects in these experiments. Recording hardware was implanted using previ-

Address for reprint requests and other correspondence: B. Jagadeesh, Physiology and Biophysics, University of Washington, Box 357330, Seattle, WA 98195 (E-mail: bjag@u.washington.edu).

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

ously described techniques (Allred et al. 2005). Briefly, surgeries on each animal were performed to implant a head restraint, a cylinder to allow neural recording, and a scleral search coil to monitor eye position (Fuchs and Robinson 1966). All animal handling, care, and surgical procedures were performed in accordance with guidelines established by the National Institutes of Health and approved by the IACUC committee at the University of Washington. Locations of neural recordings are discussed in a later section (see *Recording procedures*).

Behavioral tasks

DELAYED-MATCH-TO-SAMPLE (DMS). On each day, the monkey performed the DMS task with two stimuli. Both monkeys had been trained in variants of the DMS task for at least 1 yr before all data reported here. An example image pair and associated trials are illustrated in Fig. 1. Each trial was as follows: a red fixation square (0.3° width) appeared at the center of the monitor and was the cue for the trial to begin. After the monkey acquired fixation, there was a variable delay (456–721 ms) before the onset of the sample. The sample was presented for 0, 10, 20, 40, 80, 160, or 320 ms, and was followed immediately by a 300-ms presentation of the mask. Trials from different sample durations were randomly interleaved. On 0-ms trials, only the mask was presented, with no preceding sample stimulus. In some experiments, due to a corrected programming error, the mask was presented for only 270 ms on mask-only trials. After a delay period (116–518 ms), the choice array (which consisted of both images) was presented in the left half of the monitor, contralateral to the recording chamber (which was placed over the right hemisphere in both monkeys). One of the choice stimuli was centered 5° to the left and 5° above the fixation spot; the other was presented 5° to the left and 5° below the fixation spot. Location of choice stimuli was randomized between the two positions, so the monkey could not determine the location of correct saccade before choice array onset. If the monkey's gaze left the fixation window (2° width, centered on fixation square) at any time before the onset of the choice array, the screen turned black, the trial was discarded, and the monkey experienced a time-out for 3 s before the next trial began.

After the choice array appeared, the monkeys were required to wait for a variable delay period before making their choice of stimulus. We introduced this delay period to force the monkeys to avoid impulsive choices at the onset of the choice array. After a variable wait period (500–1,200 ms), the fixation spot changed color. If the monkey waited until after the fixation spot changed color, he was rewarded for a correct decision. If he made an eye movement before the fixation spot changed color, he was never rewarded. The change in color of the fixation point signals the availability of reward. As subsequently described, for a variety of reasons, including training history, the monkeys never learned the significance of the cue. Instead, they correctly deduced that by waiting longer periods of time, they were more likely to be rewarded for correct target choices. The time course of this increasing reward rate (for correct responses) was 0% at 500 ms (after choice array onset), increased linearly, and then reached 100% at 1,200 ms (after choice array onset). On trials where the monkey initiated a saccade before the cue to go, both target images turned off and the monkey experienced a time-out for 3 s before the next trial began. When a saccade occurred after the cue to go, the stimulus to which the monkey made a saccade remained on the screen for 2 s, whereas the other stimulus was turned off. If the monkey maintained fixation on the correct target for $\geq$50 ms, a reward of water or juice was administered, and the target stimulus remained on the screen for another 2 s. The screen then turned gray before the fixation square appeared as a cue for the next trial to begin. On error trials, the screen turned black for 3 s before the next trial began.

DMS TASK WITHOUT MASK. The DMS task without the mask was identical to the DMS task with mask, except that after the sample, no mask was presented and the trial immediately progressed to the delay period between sample presentation and choice array onset.

PASSIVE FIXATION WITHOUT MASK. This task was used exclusively to isolate selective IT neurons; no data from this task are reported. A fixation square (0.3° width) appeared. Sample onset occurred 617 ms after the monkey acquired fixation. Stimuli were presented for two successive 300-ms epochs, with an intervening 300-ms epoch where only the fixation square was presented. Monkeys were required to maintain fixation (2° width, centered on fixation square) throughout the trial and were rewarded with drops of water or juice at the end of the trial. Failure to maintain fixation resulted in a time-out period, followed by a repeat trial. Stimulus sets for the passive fixation task consisted of eight images that had been divided into four image pairs for possible use during the DMS task.

Stimuli

Stimuli were photographic images of people, animals, natural and manmade scenes, and objects. All stimuli were 90 $\times$ 90 pixels and were drawn from a variety of sources, including the world wide web, image databases, and personal photo libraries. Image pairs were drawn at random from our larger database of 4,832 images and were organized before recording sessions into 32 lists of four pairs. From these predefined lists of image pairs, selective neurons were found (see *Recording procedures*) for a total of 57 unique image pairs. Stimuli were presented on a computer monitor with 800 $\times$ 600 resolution (refresh rate 100 Hz) and images subtended 4°. Masks were produced by layering and scrambling either the effective and ineffective images (52 of 57 image pairs) or four images, two of which were the effective and ineffective images (five of 57 image pairs). One identical mask stimulus was used in each experimental session. Sample and choice stimuli were the same size.

Behavioral data analysis

TIME OF SACCADE INITIATION. We analyzed the time from choice array onset to initiation of saccade and discovered that, rather than using the cue to go to guide behavior on a trial-by-trial basis, both monkeys initiated saccades at a stereotyped time after the onset of the choice array, centered after the mean time of cue to go (mean cue to go 822 ms; mean time of saccade initiation, monkey L = 1,027 ms; mean time of saccade initiation, monkey G = 1,192 ms). For a variety of potential reasons, the monkeys never learned the significance of the cue. A post hoc analysis indicated that both monkeys' performance was poor (even on the easiest trials) when they responded soon after the onset of the choice array and performance in these early saccade trials did not reflect the ability of the monkey to complete the two-alternative forced-choice (2AFC)–DMS task.

To eliminate trials with these nonchoice-related errors from our data sets, we excluded trials that occurred before a cutoff time, determined separately for each monkey, by analyzing performance as a function of time of saccade initiation for the easiest trials, where monkeys are expected to perform nearly perfectly (320-ms sample duration). When time of saccade initiation was $>$800 ms (monkey L) or 1,200 ms (monkey G), performance on the easiest trials (320-ms sample duration) was nearly perfect (monkey L = 96%, monkey G = 92%). We therefore include trials from all sample durations when saccade initiation occurred $>$800 ms (monkey L) or 1,200 ms (monkey G) after onset of the choice array. Behavior and neural responses were calculated using this set of trials for all the data shown in the figures.

Behavioral performance was worse if all trials that occurred after the cue (all potentially rewardable trials) are included in calculating performance, rather than all trials that occurred after the cutoff latencies for each animal (10 ms = 0.55; 20 ms = 0.65; 40 ms = 0.79; 80 ms 0.84; 160 ms = 0.87; 320 ms = 0.85), as opposed to the values

shown in Figs. 9 and 10. Performance on all trials that occurred before the cutoff latency, including those that occurred before the cue, was near chance for all stimulus onset asynchronies (SOAs), indicating that the monkey guessed on these trials (10 ms = 0.50; 20 ms = 0.52; 40 ms = 0.51; 80 ms 0.52; 160 ms = 0.51; 320 ms = 0.48). Approximately 32% of trials were discarded because they occurred before the cutoff latencies for each animal. Even when these trials resulted in saccades toward the target stimulus, they did not result in a reward because most occurred before the cue. We excluded those trials to get an estimate of the monkeys' best performance with these images. The aim of these experiments was to examine the degree to which individual neurons might be sufficiently selective to account for behavioral performance and therefore it was important not to underestimate performance.

PSYCHOMETRIC FUNCTIONS. To obtain a psychometric function from behavioral data, we fit the proportion correct as a function of sample duration using the Weibull function. Fitting the data in this fashion reduces the noise associated with measurements at individual sample durations and thus allows more reliable comparisons between data sets

$$y = \gamma + (\lambda - \gamma)\{1 - \exp[-(x/\alpha)^{\beta}]\} \quad (1)$$

where y represents the proportion correct; γ and λ represent, respectively, the floor and ceiling parameters; α represents the location parameter; and β represents the slope. We used the psignifit toolbox, version 2.5.6 (see http://bootstrap-software.org/psignifit), a software package that implements the maximum-likelihood method reported by Wichmann and Hill (2001), to determine the best-fit parameters, and constrained γ to be 0.5 and λ to the range [0.6–1]. The minimum performance value is thus defined as 0.5 at 0 ms and the ceiling performance is represented by λ. The steepness of the curve is reflected by the slope parameter (β; larger values indicate steeper slopes) and the relative horizontal location of the curve is reflected by the location parameter (α). Threshold for each cell is the sample duration where performance reaches 0.75. The sensitivity of the monkey to changes in sample duration is proportional to the slope parameter, which measures how quickly the curve reaches the ceiling parameter, and is inversely proportional to threshold. For two behavioral data sets, the Weibull fit failed to converge because of poor behavioral performance. These data sets have been dropped where fit values are reported, but are included in reports of behavioral performance.

Recording procedures

On each day, an x–y stage for positioning and an electrode holder containing a sterile guide tube and tungsten microelectrode (Alpha-Omega, Nazareth, Israel) were attached to the top of the recording cylinder. The electrode was moved using a microdrive (David Kopf Instruments, Tujunga, CA) and after passing through a preamp, filter, and amplifier, signals from the electrode were sorted on-line using the Alpha-Omega spike sorter. Responses of single IT neurons were collected while monkeys viewed images and performed the DMS task (see *Behavioral tasks*). Coded spikes were stored on a PC at a rate of 1,000 Hz using Cortex, a program for neural data collection and analysis developed at the National Institutes of Health (Bethesda, MD). On-line histograms were created to qualitatively judge selectivity, but for the purposes of this paper all data analysis was performed post hoc on stored data. Eye movements were monitored and recorded (1,000 Hz) using a scleral search coil system obtained from Crist Instruments (Hagerstown, MD) and DNI (Newark, DE).

Neurons were selected using anatomical and physiological criteria. Structural magnetic resonance imaging (MRI) was used to guide placement of the recording chambers, which were centered in stereotaxic coordinates 16–17 mm lateral and 18–22 mm anterior, over the right hemisphere. Neural recordings were targeted to the center of the chamber, near the anterior middle temporal sulcus. The selection criterion for recording locations was the presence of cells that responded selectively to the types of stimuli used in this study, and recording locations were altered until such selectivity was found. Precise anatomical locations of recording sites are unavailable because both animals are still being used in other experiments. However, neurons were likely located in anterior IT on the ventral surface (and not either bank of the STS) because our recording procedure involved locating the anterior limit of IT and the ventral surface of the brain and targeting recording sites to that location. We have less information about the precise lateral/medial location, but based on recording with the structural MRI in hand, the sites were likely just medial to the anterior medial temporal sulcus.

On each day, the electrode was extended to the expected anatomical location of IT before the monkey began any task. After a waiting period of 10–30 min, depending on qualitative assessment of recording stability, the monkey began the passive fixation task without a mask while we attempted to isolate one or two neurons. On isolation of a neuron, we qualitatively judged whether the neuron was selective for one of the image pairs. If the neuron did not appear to be selective for any of the four image pairs, the electrode was advanced and we attempted to isolate another neuron. On isolating an apparently selective neuron, the monkey began performing the DMS task with the stimulus set for which the isolated neuron was selective.

In post hoc analysis we defined neurons as visually responsive if the neuron responded significantly above baseline (defined as the average spike count in the 300 ms before sample onset) in any of the 13 stimulus conditions during a 150-ms epoch (75–225 ms after sample onset). The epoch was offset 75 ms after sample onset to account for neural latency and significance was assessed at the $P < 0.05$ level (unpaired t-test; at least one trial per condition). We recorded from 166 visually responsive neurons. Neurons were classified as visually selective (96/166) if they responded significantly differently to the two stimuli in either the 160- or 320-ms conditions (75- to 225-ms epoch; $P < 0.05$, unpaired t-test). The sample that elicited the higher response in the 150-ms epoch (160- and 320-ms sample durations) was termed the "effective sample," whereas the other was termed the "ineffective sample." The neural data collected while monkeys performed the control task (DMS task with no mask) were also required to fulfill the criteria outlined earlier (visually responsive and selective). In the same epoch, we also calculated a Selectivity Index (*SI*) for each neuron at each sample duration, which was defined as

$$SI = \frac{E_{resp} - I_{resp}}{E_{resp} + I_{resp}} \quad (2)$$

where E_{resp} is the mean firing rate to the effective image in the indicated epoch (75–225 ms after sample onset) and I_{resp} is the mean firing rate to the ineffective image.

Neurophysiological data analysis

Neural performance was calculated using an ideal-observer analysis. This analysis rests on two theoretical assumptions. First, we assumed that for each recorded neuron, there existed an "anti-neuron" elsewhere in IT cortex with identical but opposite selectivity; that is, the anti-neuron responded to the ineffective sample in the same way that the recorded neuron responded to the effective sample and vice versa. Second, we assumed that on each trial, the ideal observer was able to compare the responses of the neuron and anti-neuron to determine whether the effective or ineffective image was presented. In practice, the ideal-observer analysis was performed by dividing the data from each experiment into groups based on sample duration, and then separating spike counts from trials into two distributions dependent on whether the effective or ineffective sample was presented during the sample period. Our measure of neural performance was the

area under the region of overlap curve (ROC) obtained from the two distributions (Green and Swets 1966). This performance value represents the proportion of trials that the ideal observer would correctly identify the effective sample by comparing a random pick from the neuron and "anti-neuron" spike count distributions. We also repeated this ideal-observer analysis at every sample duration except for the 0-ms (mask-only) condition.

We calculated the neural performance using many different neural epochs. In the ideal-observer analysis, spike counts on each trial were summed in epochs beginning between 50 and 150 ms after sample onset (to account for neural response latency) and lasting between 1 and 400 ms. All neurophysiological data analysis was confined to epochs time locked to sample onset. For each epoch, we obtained neurometric fits from neural performance values using Weibull function (*Eq. 1*) that was also used to fit behavioral data.

In certain epochs, some neurons classified as selective using the selection criteria (see *Recording procedures*) failed to reach threshold performance (0.75) in the Weibull fit. In any analysis where a neuron failed to reach threshold performance, or where threshold was >1,000 ms (threefold the longest sample duration), it was dropped from that analysis and this exclusion is reflected in the reduced number of neurons (n) for that that analysis. When performance values (rather than fit parameters) are reported, all neural data sets (96) are included.

Comparing psychophysical and neuronal sensitivity

We compared psychophysical and neuronal sensitivity for individual experiments using the bootstrap analysis described in Uka and DeAngelis (2003). For a given experiment, we resampled neural responses and behavioral choices by making random draws (with replacement) from spike counts and behavioral choices at each sample duration. The number of random draws for behavioral choices and spike counts was equal to the number of trials actually performed at each sample duration. We then created a psychometric and neurometric curve from the randomly resampled data. For each experiment, we created 1,000 psychometric–neurometric pairs. For a given experiment, neurometric and psychometric sensitivities were considered to be statistically different if the 95% confidence interval of the difference between calculated threshold for each psychometric–neurometric pair did not overlap with 0.

RESULTS

We recorded from 166 visually responsive cells, 96 of which were also selective for a pair of stimuli (see *Recording procedures* in METHODS). These IT neurons were recorded from two rhesus monkeys while they performed a DMS task (Fig. 1). We manipulated task difficulty by varying sample duration, and we made quantitative comparisons between behavioral performance and the activity of single neurons as sample duration changed. We first show that across the population, most individual neurons' sensitivity to changing sample duration was not statistically different from that of the behaving monkey. Second, we examine how the epoch used to analyze neural response affects neural performance and its relationship to behavioral performance. Finally, through a control task (a DMS task with no mask) we provide additional evidence that the magnitude and temporal duration of stimulus selectivity are related to behavioral performance.

Behavioral and neural performance

Behavioral accuracy in the DMS task was dependent on sample duration. The behavioral data for one stimulus set are shown in Fig. 2. Two different sample stimuli were presented during the performance of this task. On each day, stimuli were selected so that the neuron responded relatively strongly to one sample (effective sample) and relatively weakly to the other sample (ineffective sample; see *Recording procedures* in METHODS). We calculated the proportion of trials on which the monkey correctly identified the sample stimulus (by making a saccade to it in the choice array) as a function of sample duration. For the example stimulus pair shown in Fig. 2, behavioral performance improved with sample duration, ranging from 0.55 at the 10-ms sample duration to 0.97 at the 320-ms sample duration. The data were fit with a Weibull function (see *Eq. 1*). The resulting psychometric function is plotted in Fig. 2 and yielded a slope parameter (β) of 3.34 and a location parameter (α) of 31.48, with a ceiling parameter (λ)

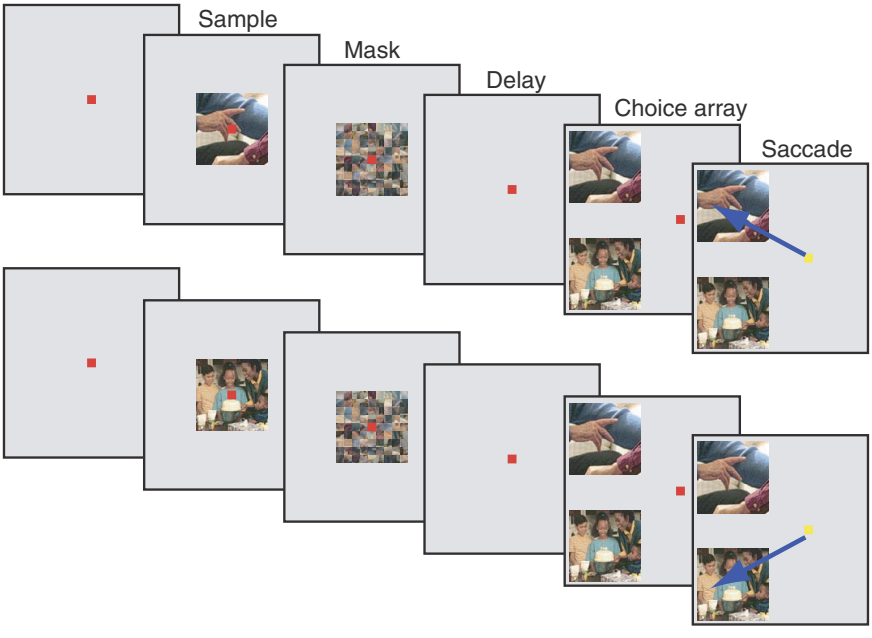

FIG. 1. Behavioral task. On each day the monkey performed the task with 2 possible sample images. For one example stimulus set, a representative trial for each sample image is illustrated in each succession of frames. After fixation was acquired, the sample was presented, followed immediately by a mask, followed by a delay period. Choice array, consisting of both images, then appeared in the periphery. After the fixation square changed color, the monkey received a liquid reward for making a saccade to the image in the choice array that was presented during the sample period. Images in the choice array were the same size as the sample image; for convenience, they are displayed here as smaller.

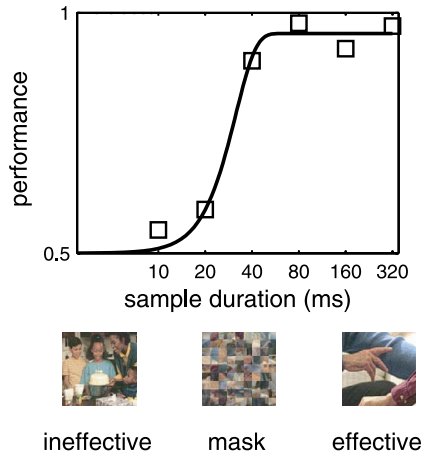

FIG. 2. Behavioral data and accompanying stimuli for one experiment. Proportion correct as a function of sample duration. Each square represents the actual proportion of trials on which the monkey chose the correct sample at that sample duration; the black curve represents the Weibull fit of the behavioral data (see *Eq. 1* in METHODS).

of 0.96. For this stimulus set, the monkey's threshold (sample duration required to reach 0.75 performance) was 29 ms.

As with behavioral performance, neural response varied as a function of sample duration. Longer sample durations were associated with increased selectivity for the sample stimulus, resulting from changes in both the magnitude and duration of neural response. Figure 3*A* illustrates that increasing the sample duration resulted in greater neural selectivity for the example stimuli in Fig. 2. Short presentations of either the ineffective or effective sample elicited weak responses from the neuron. Increasing the duration of the effective sample increased both the peak neural response and the duration of that response.

The sensitivity of the neuron for the stimuli shown in Fig. 2 was not different from the sensitivity of the monkey, as shown in the quantitative comparison between neural and behavioral sensitivity outlined in Fig. 3, *B* and *C*. We asked how well an ideal observer could determine whether the effective or ineffective sample was presented by analyzing the spike counts recorded on individual trials. Briefly, we calculated the spike count, at each sample duration, in the 128-ms epoch beginning 85 ms after sample onset (to account for neural latency). Choice of this epoch was motivated by additional analysis, described in the next section. For the example neuron, the average spike count elicited at each sample duration is plotted in Fig. 3*B*. For this neuron, high spike counts were rarely elicited on trials where the ineffective sample was presented; conversely, low spike counts were rarely elicited on trials where the effective sample was presented. Mask-only trials (0-ms sample duration) are not shown in Fig. 3*B*; because no decision was objectively correct on these trials, they have been omitted from the neural performance calculation in Fig. 3, *B* and *C*. For each sample duration, spike counts elicited on trials when the effective sample was presented were compared with spike counts elicited on trials when the ineffective sample was presented, using the ROC statistic (see *Neurophysiological data analysis* in methods). Like the behavioral data, neural performance values were also fit by the Weibull function (Fig. 3*C*: $\alpha = 27.02$; $\beta = 2.72$; $\lambda = 0.98$; threshold = 24 ms). The sensitivity of the neuron, defined by threshold and slope, was not different from the sensitivity of the monkey ($P > 0.05$, bootstrap analysis; see METHODS; Uka and DeAngelis 2003).

A comparison between neurometric and psychometric functions for four additional example cells is shown in Fig. 4. For a majority of experiments, neural and behavioral sensitivity were not statistically different from each other (Fig. 4, *A* and *C*). Occasionally, monkeys were statistically more sensitive than neurons (Fig. 4*B*). Even more rarely, individual neurons were statistically more sensitive to changing sample duration than was the behaving monkey (Fig. 4*D*).

Across the population, we found that neurons and monkeys were similarly sensitive to changes in sample duration. This relationship is evident in Fig. 5, where we plot threshold values (Fig. 5*A*), slope parameters (Fig. 5*B*, parameter "β" from the Weibull fit), and ceiling performance (Fig. 5*C*, parameter "λ" from the Weibull fit) for each experiment, and report fit parameters across the population (Table 1). For most data sets, neural and behavioral sensitivity were not distinguishable (sensitivity defined by threshold = 65/96 experiments, points not significantly away from the diagonal in Fig. 5, *A* and *B*). Occasionally, monkeys were statistically more sensitive than neurons (16/96 experiments, values above the diagonal in Fig. 5*A*). Infrequently, neurons were statistically more sensitive than the monkey (seven of 96 experiments, values below the diagonal in Fig. 5*A*). Across the population, neural and behavioral thresholds, slopes, and ceiling parameters were not statistically distinguishable (Table 1; thresholds: $P = 0.08$; slopes: $P = 0.38$; ceiling parameters: $P = 0.06$, two-tailed, paired *t*-test). We found no stimulus-by-stimulus correlation between neural and behavioral sensitivity defined by threshold, slope, or ceiling performance (Fig. 5*A*, $r = 0.12$, $P = 0.34$; Fig. 5*B*, $r = 0.1062$, $P = 0.31$; Fig. 5*C*, $r = -0.17$, $P = 0.10$); that is, image pairs and sessions that elicited higher behavioral sensitivity were no more likely to elicit higher neural sensitivity. Median threshold for neurons was 29 ms (see Table 1), whereas median threshold for the monkey was 26 ms (see Table 1).

Effect of epoch on the relationship between neural and behavioral sensitivity

In the preceding analyses (Figs. 2–5), neural performance was calculated using a standard epoch (beginning 85 ms after sample onset and lasting for 128 ms); that standard epoch was chosen because it minimized neural threshold (maximized neural sensitivity) across the population and across all sample durations. However, the selectivity of a neuron for a stimulus pair in any epoch will depend on the latency of individual neurons (which is variable) and the sample duration of the stimulus (which was manipulated in the experiment). In the first part of this section, we explore how the start of epoch and the length of epoch affect neural sensitivity to changing sample duration and its comparison to behavioral sensitivity; in the second part of this section, we investigate how the epoch that maximizes neural performance varies with sample duration.

The estimate of neural threshold depends on both the start of the epoch (estimation of neural response latency) and the length of the epoch that determines which spikes are included in the calculation. The dependency of neural threshold on the start and length of epoch is illustrated in Fig. 6. To quantitatively evaluate the relationship between neural threshold and

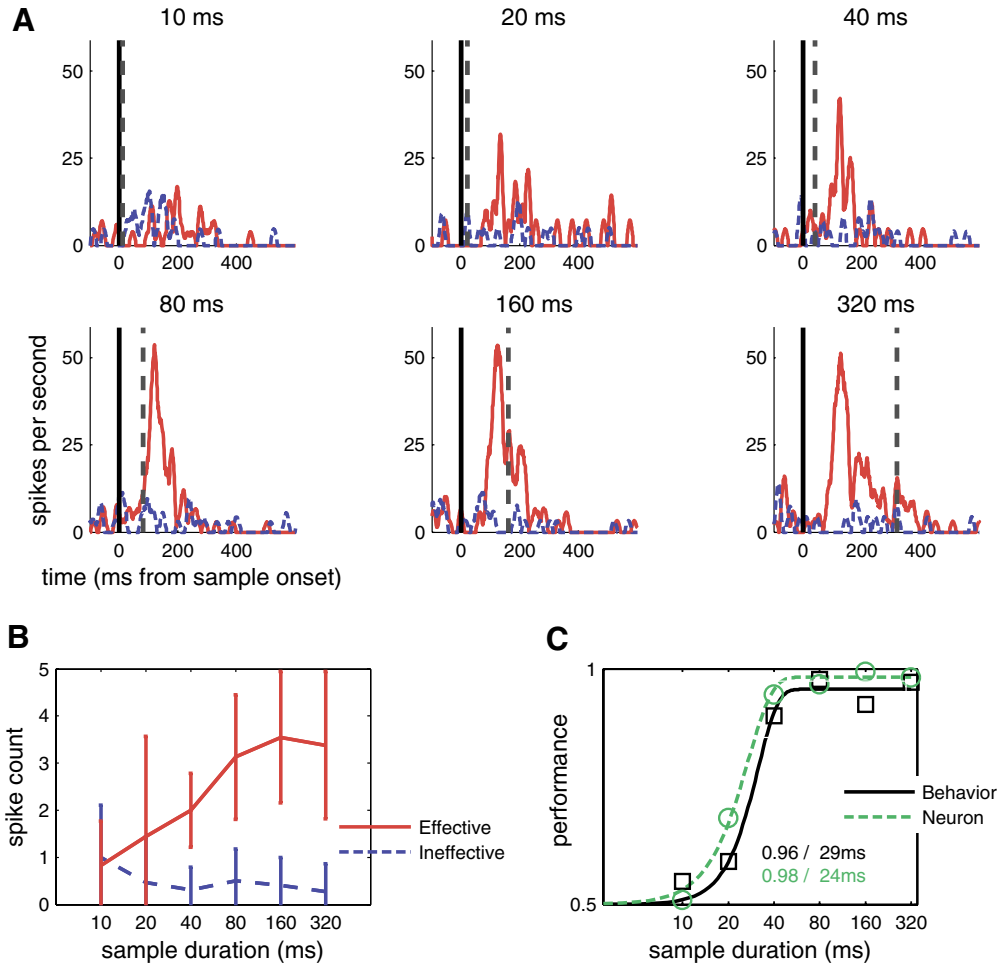

FIG. 3. Response of an inferotemporal (IT) neuron during the delayed-match-to-sample (DMS) task. *A*: each panel represents the average firing rate (spikes/s) as a function of time after sample onset for the sample duration indicated above the panel. *Blue dashed traces* represent neural response on presentations of the ineffective sample; *red traces* represent neural response on presentations of effective sample. Vertical black solid lines represent sample onset and vertical black dashed lines represent mask onset. To obtain the peristimulus time histograms (PSTHs) shown, the spike counts in each millisecond time bin were averaged across all trials for a given sample duration, converted to spike rates, and smoothed with a 20-ms Gaussian window. *B*: average spike count as a function of sample duration for effective (red, solid line) and ineffective samples (blue, dashed line). For each sample duration, the number of spikes in the 128-ms epoch beginning 85 ms after sample onset was averaged across trials. Error bars are SD. *C*: proportion correct as a function of sample duration for the example neuron (green dashed line, circles) and the monkey (black solid line, squares, data from Fig. 2). Neural performance was calculated using the ideal-observer analysis (see *Neurophysiological data analysis* in METHODS). Green curve represents the Weibull fit of the neural data. Ceiling parameters (λ) and threshold (*t*) are shown for neurometric (green text) and psychometric (black text) functions.

epoch (both start and length), we calculated neural threshold using epochs that started between 50 and 150 ms after sample onset and epochs that lasted between 1 and 400 ms; we refer to the start of the epoch as the estimate of neural response latency. For example, the 20-ms epoch for the 100-ms neural response latency lasted from 101 to 120 ms after sample onset. It takes time for information to propagate through the visual system to IT; thus an IT neuron will begin to respond to a stimulus many milliseconds after that stimulus was presented. To determine how length of epoch is related to length of sample stimulus, it is thus important to choose the correct start of the epoch. For each epoch, we then calculated the threshold for each neuron using the same ideal-observer analysis and Weibull fit described in Figs. 3–5 for the 85-ms latency and 128-ms epoch; we averaged across cells to obtain the values in Fig. 6. In Fig. 6*A*, we examine how average threshold varied with length of epoch for five representative examples of neural response latency. Conversely, in Fig. 6*B*, we plot average threshold as a function of epoch start, for five representative examples of epoch length. Neural threshold was dependent on both the start of the epoch and the length of the epoch; for the neural response latency used in Figs. 3–5 (85 ms), average threshold decreased as epoch increased, reached a minimum at 128 ms, and then increased again gradually. This minimum average threshold (which reflects maximum neural sensitivity) motivated the choice of the 128-ms epoch and 85-ms neural response latency used in Figs. 3–5.

Across the population, the lowest threshold occurred in the epoch beginning 85 ms after sample onset and lasting for 128 ms. For each estimate of neural response latency, threshold is initially high, reaches a minimum somewhere between 90 and 200 ms, and then increases again. The reason for this balance between start of epoch, length of epoch, and average threshold is illustrated in Fig. 3*A*. In this example cell, the selective response started at 96 ms. If we chose an earlier start of the epoch (decreased neural response latency), then the epoch would include time bins where the response to the effective sample and the response to the ineffective sample were iden-

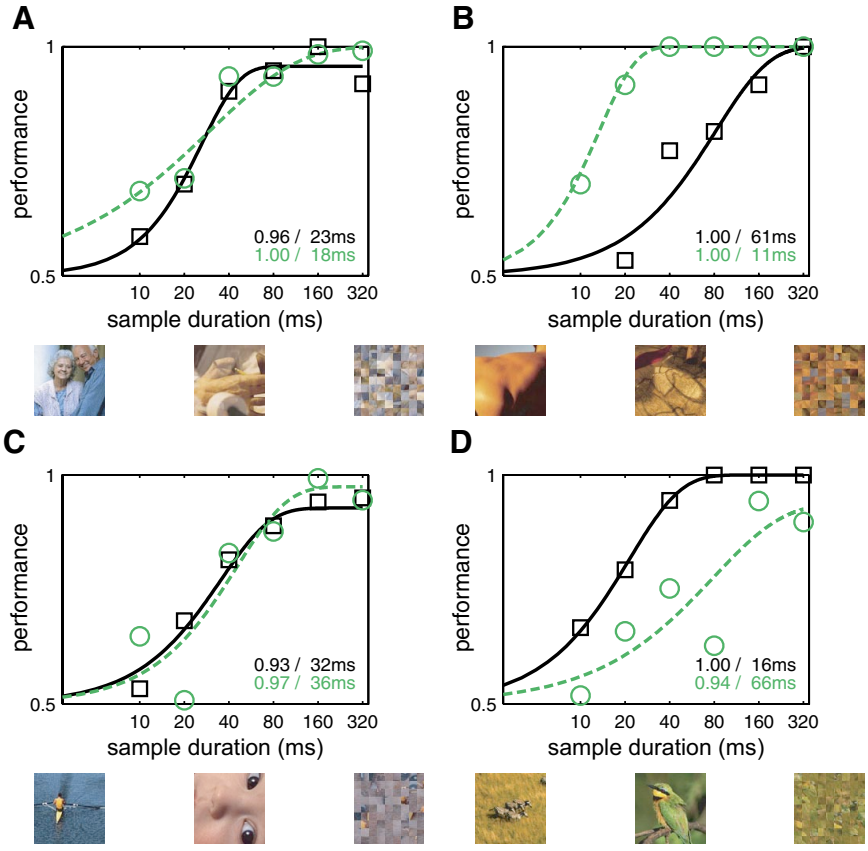

FIG. 4. Neurometric and psychometric functions for 4 experiments. *A–D*: for 4 different experiments, proportion correct as a function of sample duration for example neurons (green dashed line, circles) and monkey (black solid lines, squares). Lines represent the Weibull fits of data. Ceiling parameters (λ) and threshold (*t*) are shown for neurons (green text) and behavior (black text) in each panel. Each set of 3 images shows the ineffective (*left*), mask (*center*), and effective (*right*) stimuli used in that experiment.

tical; these spikes would add only noise to the calculation. If we chose an epoch that started long after stimulus onset (increased neural response latency), the threshold calculation would miss part of the selective response. In some cases (e.g., short epochs and short neural response latencies) the threshold calculation would miss all of the selective response. An example of this would be an epoch that started and ended before the effective and ineffective responses diverge (see Fig. 3*A*). In these epochs, many neurons were not selective for the stimulus, even at the longest sample duration. When neurons were insufficiently selective, they were discarded from the threshold calculation because they never reached threshold (see *Neurophysiological data analysis* in METHODS); thus the number of cells reaching threshold is also a measure of neural performance. In Fig. 6*A*, the majority of cells (50%, 48/96) in the optimal latency group (85 ms, blue line) had reached threshold by the 58-ms epoch and in Fig. 6*B* a majority of cells (69%, 66/96) in the optimal length of epoch group (128-ms epoch, blue line) had reached threshold for the first tested latency (50 ms). Many fewer cells were included at short epochs, indicating that neural performance across the population was very poor in these epochs.

In the plots of optimal thresholds shown so far, we have used a common epoch for all neurons [both start (85 ms) and length (128 ms)] and we have shown that for most data sets, neural and behavioral sensitivities were not statistically distinguish-

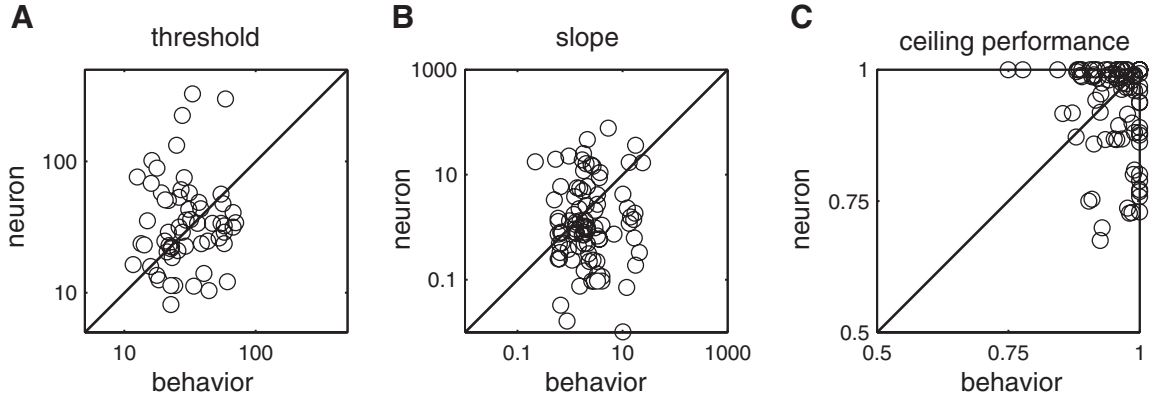

FIG. 5. Comparison of neural and behavioral data across the population. *A*: sample duration required to reach threshold performance (0.75) in behavioral (*x*-axis) and neural (*y*-axis) data sets (log scale). Shown are the 65 data sets where neural and behavioral data reached threshold in the epoch tested (start of epoch: 85 ms; length of epoch 128 ms). *B*: slope parameters (β) for Weibull fits to behavioral (*x*-axis) and neural (*y*-axis) data ($n = 96$). *C*: ceiling parameters (λ) for Weibull fits to behavioral (*x*-axis) and neural (*y*-axis) data sets ($n = 96$).

TABLE 1. *Average of parameters from the Weibull fits for behavioral and neural data*

	λ	α	β	*n*	Threshold	*n*
Neurons	0.93 / 0.98 / 0.0094	360 / 50 / 82	5.0 / 0.97 / 1.1	93	45 / 29 / 7	65
Monkey	0.95 / 0.96 / 0.0052	47 / 32 / 9	4.0 / 1.9 / 0.54.1	93	31 / 26 / 2	65

Ceiling parameters (λ, column 2), location parameters (α, column 3), and slope parameters (β, column 4) are shown for neural (*top row*) and behavioral (*bottom row*) data for the 93 data sets where the fitting routine successfully returned values for both neural and behavioral data sets (fitting was unsuccessful for two neural data sets and two behavioral data sets, one of which overlapped). Within each column, the first value represents the mean, the second represents the median, and the third represents the SE. The sample duration required to reach threshold performance of 0.75 (column 6) is shown for the 65 experiments where both neurometric and psychometric functions for the observed data had ceiling parameters >0.75 (29 neural data sets failed to reach threshold; two behavioral data sets failed to reach threshold). Thresholds are reported from the smoothed values used to determine choice of epoch (see Fig. 6).

able (Fig. 5). However, there was considerable variability in the start and length of the epoch that yielded the lowest thresholds from neuron to neuron (Fig. 6*C*, mean start of epoch = 99 ms, SD = 30 ms; mean length of epoch = 131 ms, SD = 94 ms), and it is possible that internal neural calculations may take into account that variability. When epochs are optimized on a cell-by-cell basis rather than an average basis, threshold values must be lower (neurons will be more sensitive). We observed that optimizing on a cell-by-cell basis decreased thresholds significantly, by a factor of 1.7 (Fig. 7*A*; median using average latency and epoch = 29 ms; median with individually optimized latency and epoch = 14 ms; $P < 0.00001$, paired *t*-test).

We evaluated the relationship between neural threshold and epoch to relate neural performance to behavioral performance, not to find the best possible neural performance. In fact, if the epoch used to calculate neural sensitivity was allowed to vary between neurons (the epoch used in Fig. 7*A*), then individual neurons outperformed the monkey at short sample durations. In Fig. 7, *B–F*, neural performance (in the epoch that minimized neural threshold for each cell in Fig. 7*A*) is plotted against the behavioral performance measured at each sample duration. At subthreshold sample durations, individual neurons performed significantly better than the behaving monkey (Fig. 7*B*; 10-ms sample duration: mean neural performance = 0.64 ± 0.02, mean behavioral performance = 0.56 ± 0.01, $P < 0.0001$, paired *t*-test; 20-ms sample duration: mean neural performance = 0.73 ± 0.02, mean behavioral performance = 0.67 ± 0.01, $P < 0.005$, paired *t*-test). At longer sample durations, optimization of epoch for each cell did result in significantly

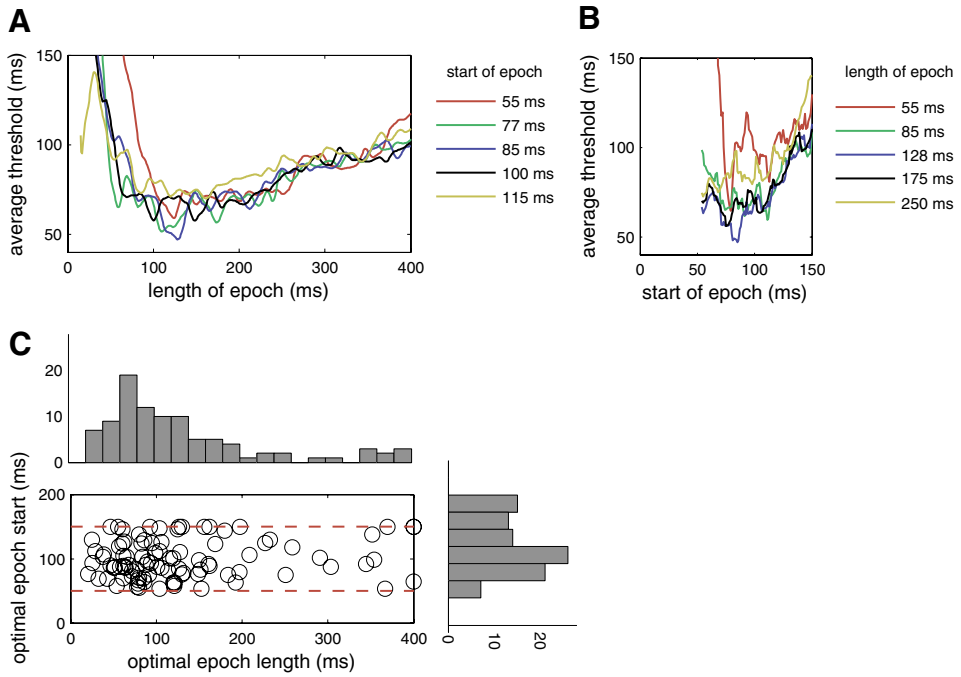

FIG. 6. Dependence of neural threshold on start and length of epoch. *A* and *B*: each line represents average neural threshold as a function of epoch length for 5 different epoch starts (*A*) or average neural threshold as a function of start of epoch for 5 different epoch lengths (*B*). Thresholds were calculated as follows: spike counts for effective and ineffective stimuli at each sample duration were collected in time bins for the specified start and length of epoch (e.g., 51–125 ms after sample onset for a 50-ms start of epoch and a 75-ms epoch length). Neural performance values at each sample duration were calculated with the resulting spike counts using the ideal-observer analysis and fit with a Weibull function (see *Neurophysiological data analysis* in METHODS). Threshold for 0.75 performance was then extracted from the resulting fit. For each cell, threshold as a function of epoch length was smoothed with a 10-ms Gaussian window, then threshold as a function of epoch start was smoothed with a 5-s Gaussian window. Average across cells is plotted. Cells that did not reach threshold in the tested epoch were discarded from the average. *C*: distribution of start and length of epochs that yielded the lowest threshold for each cell. As described in *A*, threshold for each cell was calculated in epochs that started between 50 and 150 ms after sample onset (indicated by red horizontal dashed lines) and lasted between 1 and 400 ms. Resulting 150 (number of epoch starts) × 400 (number of epoch lengths) matrix of threshold values for each cell was smoothed across length of epoch using a 10-s Gaussian window, then smoothed across start of epoch using a 5-s Gaussian window. Location of the lowest threshold in the 150 × 400 matrix of threshold values for each cell determined the length of epoch and the start of epoch depicted in *C* ($n = 96$; each cell reached threshold in at least one epoch).

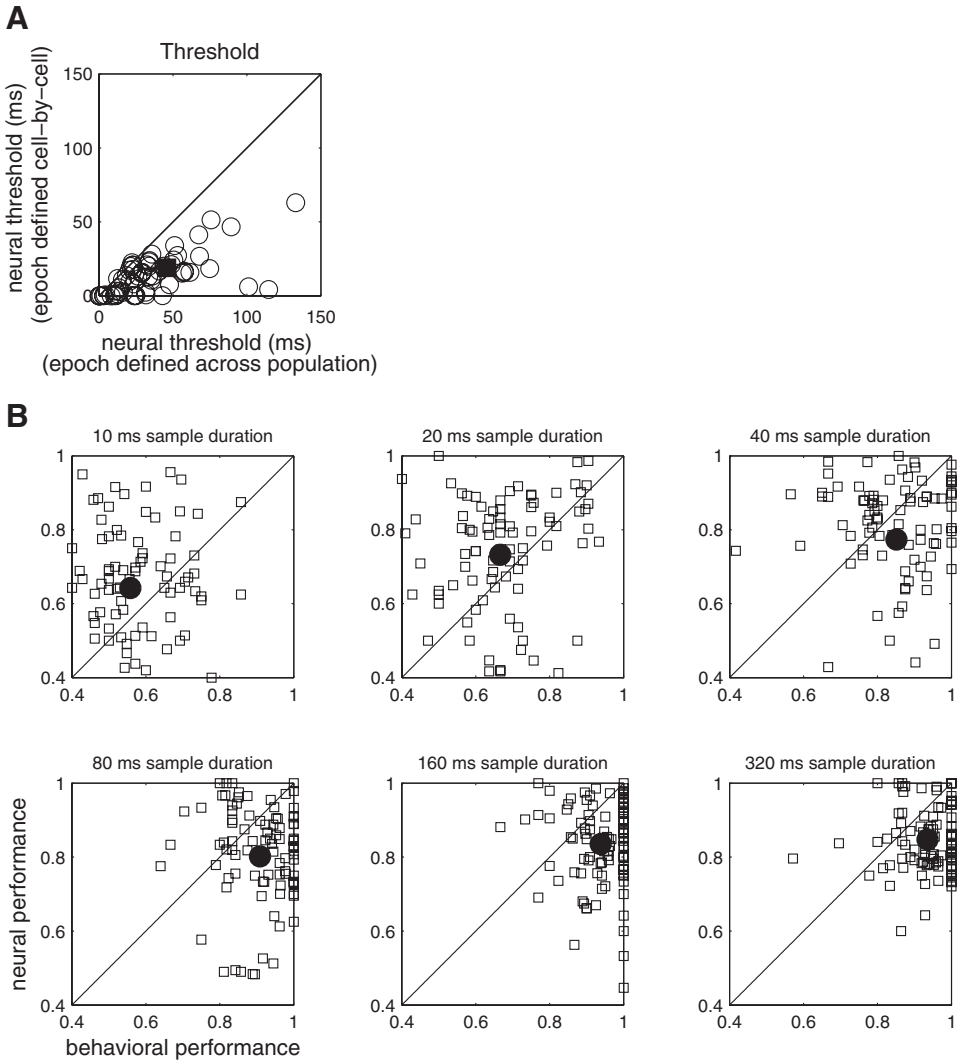

FIG. 7. Effect of optimizing start and length of epoch on a cell-by-cell basis. *A*: for each neuron, threshold is plotted for 2 different epochs. Choice of epochs is subsequently described. First, neural thresholds were determined using the ideal-observer analysis described in epochs beginning between 50 and 150 ms after sample onset and lasting for between 1 and 400 ms. Second, for each cell, threshold values were smoothed across length of epoch using a 10-ms Gaussian window and smoothed across start of epoch using a 5-ms Gaussian window. Finally, to obtain the cell-by-cell optimal epoch (epoch used to determined threshold on the *y*-axis), the epoch containing the lowest threshold value was determined separately for each cell (epochs shown in Fig. 6*C*). Threshold using this optimal epoch is plotted along the *y*-axis. To obtain the average optimal epoch (epoch used to determine threshold on the *x*-axis, epoch used in Figs. 3, 4, and 5), this matrix was averaged across cells and the epoch containing the lowest threshold value was used (see Fig. 6*A*). Threshold (*t*) using a common epoch for all cells is plotted along the *x*-axis ($n = 67$, cells that had *t*). *B*: comparison of neural and behavioral performance at each sample duration. Each panel plots neural performance (*y*-axis) and behavioral performance (*x*-axis) for each experiment (squares). Performance values are actual values, not extracted from fits to data. Neural performance values are taken in the epoch defined as optimal in the cell-by-cell optimization (Fig. 6*C*, epoch used to determine threshold on the *y*-axis in *A*). Solid circles represent means across all experiments. Error bars (SE) are plotted, but are smaller than circle size ($n = 96$).

higher neural performance than optimization of epoch across the population, but behavioral performance was still significantly better than neural performance (40-, 80-, 160-, 320-ms sample durations, $P < 0.005$, two-tailed, paired *t*-test).

Dependence of optimal epoch on sample duration

The epoch that minimized neural threshold across the population (128 ms) was longer than most of the sample durations used in this study (10, 20, 40, 80, 160, 320 ms); in the analysis shown in Figs. 2–7, we calculated performance using the same epoch for all sample durations. Somewhat surprisingly, even after accounting for neural response latency, using this strategy resulted in optimal epochs that included a substantial period of mask presentation, during which the visual stimulus was identical on all trials. For example, during the 10-ms sample duration trials, the stimulus changed from the sample to the mask during the epoch in which threshold was calculated (even after accounting for neural response latency); the sample was presented for 8% of the time during that epoch, whereas the mask was presented 92% of the time. Using a different epoch might thus affect neural performance at each sample duration differently because of the possible trade-off between optimal epoch and sample durations: longer epochs produce higher neural selectivity at longer sample durations, but result in an integration of response to a noninformative stimulus (mask) at shorter sample durations. The average threshold plotted in Fig. 6 relied on calculating neural performance at all sample durations and cannot disambiguate how epoch affects neural performance at each sample duration separately.

That the optimal epoch is on average longer than sample duration is illustrated in Fig. 8*A*, where we show population histograms for all 96 visually selective neurons at each sample duration, time-locked to sample onset. Across the population, the histograms show that at each sample duration, neural response to the sample began about 85 ms after sample onset (black arrow). In addition, although onset of the masking stimulus diminished neural selectivity, the responses to the effective and ineffective stimuli in the first four panels (10- to 80-ms sample durations, red and blue lines) remain separated past the sample duration. Green arrows indicate the end of sample duration, offset by the estimate of neural response latency (85 ms). The fact that selectivity persists beyond the sample duration is further emphasized in Fig. 8*B*, where the same data are plotted, but with each cell's response time-locked to its estimated neural response latency (start of epoch used in Fig. 7). Here the black solid lines represent the time of

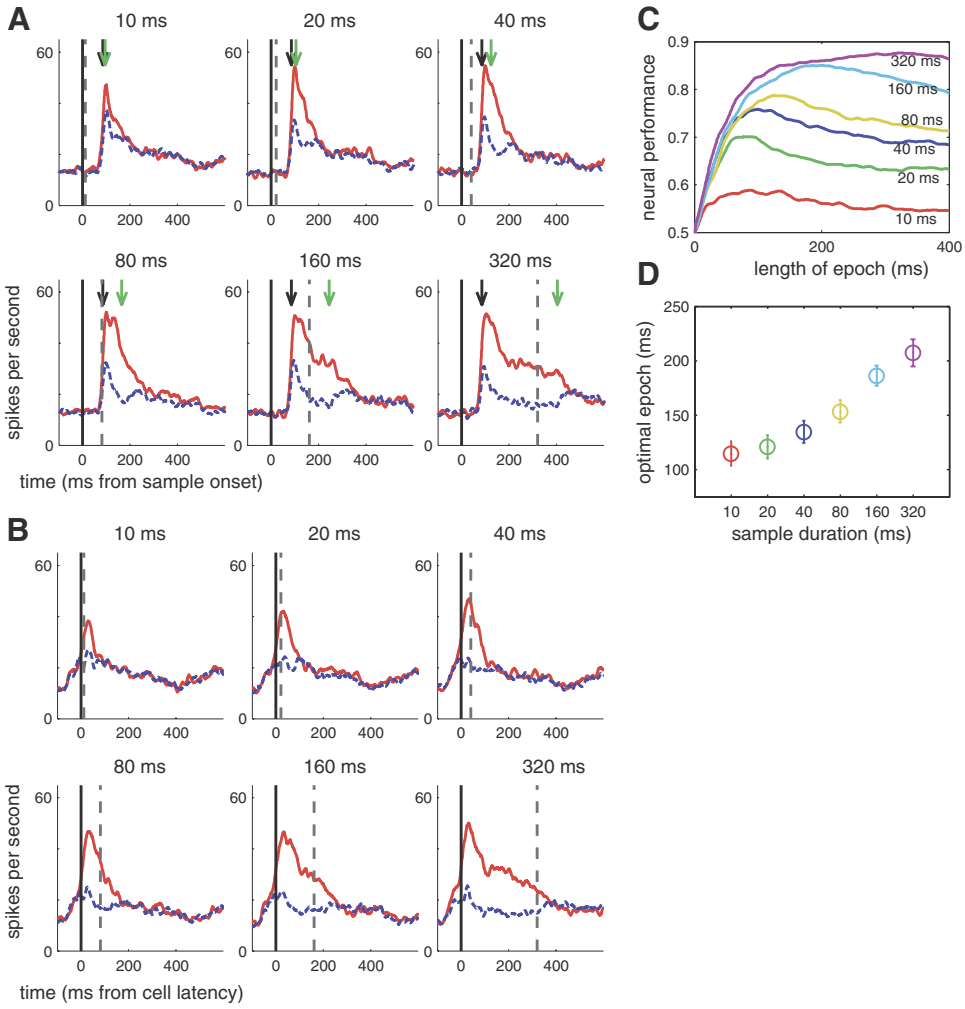

FIG. 8. Relationship between optimal epoch and sample duration. *A* and *B*: neural selectivity as a function of sample duration. Each panel shows the average PSTH of neural responses of all cells to all effective stimuli (red solid line) and ineffective stimuli (blue dashed line) for the indicated sample duration (10, 20, 40, 80, 160, 320 ms). Averages were created by smoothing the mean response (20-ms Gaussian window) across trials for an individual cell and then averaging across cells. *A*: responses are time locked to stimulus onset (long vertical black line) and mask onset is indicated (black dashed line). Arrows represent sample duration adjusted by estimated neural response latency (black arrows represent 85 ms after sample onset; green arrows represent 85 ms + sample duration). *B*: each cell's response is time locked to the start of epoch used in Fig. 7 (indicated by the long vertical black line) and the length of sample duration relative to the start of epoch is shown (black dashed line). *C* and *D*: dependence of optimal epoch on sample duration. Using a standard start of epoch (85 ms after sample onset) for each neuron, spikes were summed in epochs ranging from 1 to 400 ms; neural performance values at each epoch were calculated using the ideal-observer analysis. Resulting performance values for each cell were smoothed with a Gaussian window of 10 ms and the average across cells is plotted in *C*. For each cell, the epoch with the highest performance value was defined as the optimal epoch for that neuron at that sample duration. Optimal epochs were then averaged across neurons (*D*). Error bars are SE ($n = 96$).

neural response onset and the black dashed lines represent the time that the mask turned on, offset by the time of neural response onset.

To investigate how sample duration quantitatively affected the relationship between neural selectivity and epoch, we calculated neural performance as a function of epoch length (with a common start of epoch, 85 ms) for each sample duration separately (Fig. 8*C*). At each sample duration, there is an identical rise in neural performance as epoch increases; after this rise, the neural performance lines separate and peak at different epochs, showing that both optimal performance and the epoch that optimizes that performance increase with sample duration. When neural performance was calculated separately at each sample duration, the epoch that yielded the highest neural performance significantly outlasted the length of sample duration for all but the longest sample duration (Fig. 8*D*, 10- to 80-ms sample durations, $P < 0.00001$; 160-ms sample duration, $P < 0.005$; 320-ms sample duration, $P = 1$; one-tailed *t*-test).

As demonstrated in Figs. 6–8, the choice of epoch alters estimates of neural performance. To gain insight about what choice of epoch the brain might use to inform behavior, we optimized epoch in three different ways and compared neural performance at each sample duration to behavioral performance (Fig. 9, *A* and *B*). In the first example (green circles) a standard start and length of epoch was used for each cell (85-ms start, 128-ms length, epoch in Figs. 3–5); this strategy assumes that whatever algorithm decodes IT neural response is incapable of optimizing epoch on a cell-by-cell basis. In the second example (red crosses), optimal epochs were calculated cell by cell (epoch used in Fig. 7). However, only one optimal epoch per cell was calculated; that is, for a given cell, the same epoch was used at each sample duration. Calculations in other studies that estimate neural latency and use a different start of epoch for each neuron assume that the brain is capable of this type of optimization. In the third example (blue triangles), the optimal epoch was determined separately for each cell at each sample duration. Engaging in this optimization would require an explicit representation of sample duration; because trials of different sample durations are interleaved, it is not possible to predict sample duration until after mask onset. Thus it seems unlikely that the brain would be able to utilize this type of optimization. Finally, the behavioral performance at each sample duration is plotted for comparison (black squares). Figure 9*A* shows the corresponding epoch lengths that are used.

We can draw two conclusions from Fig. 9. First, even when the epoch is optimized separately for each cell at each sample duration, epoch length is longer than sample duration for the 10-, 20-, 40-, and 80-ms sample durations. In all cases, the length of epoch was offset by the estimate of neural response latency [either across the population (green circles) or for each cell individually (red crosses or blue triangles)], so this is not due to the delay before visual information reaches IT cortex.

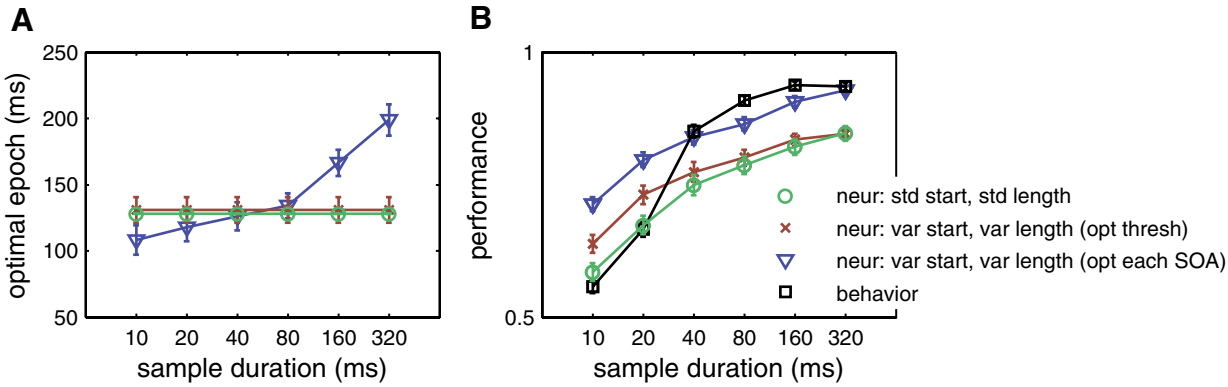

FIG. 9. Effect of optimization at each sample duration. Optimal epochs (*A*) and performance in the optimal epoch (*B*) for 3 different choices of epoch. *1*) Green circles: epoch was the same (start = 85 ms, length = 128 ms) for each cell (green circles) and sample duration; the chosen epoch minimized threshold across the population (Figs. 5 and 6). *2*) Red crosses: epoch varied from cell to cell; the chosen epoch minimized threshold for that cell (Fig. 7). For each cell, the epoch was the same at all sample durations. Epochs described with green circles and red crosses are not required to be identical. *3*) Blue triangles: epoch was chosen separately for each cell and each sample duration; the chosen epoch maximized performance of that cell at that sample duration. By definition, the epoch is constant across sample duration for green circles and red crosses (they are plotted at each sample duration for reference to possible changes in blue triangles). Performance values at each sample duration were smoothed in the same way that threshold values were smoothed (Figs. 6 and 7, smoothed with a 10-ms Gaussian window across length of epoch and 5-ms Gaussian window across start of epoch). *A*: average optimal epochs as a function of sample duration. *B*: average performance in the corresponding epoch from *A*. In addition, black squares show average behavioral performance at each sample duration. For all data points, error bars (SE) are plotted, but may be smaller than the data point ($n = 96$). There are no error bars for the green circles in *A* because the epoch used was that which minimized the average threshold across the population.

Second, neural performance is almost always better than behavioral performance at short sample durations (10 and 20 ms) and worse than neural performance at longer sample durations (40, 80, 160, and 320 ms). In the case of optimization cell by cell and at each sample duration, neural performance is identical to behavioral performance at the longer sample durations and much better than behavioral performance at the 10- and 20-ms sample durations. In fact, average neural performance at the 10-ms sample duration is 0.72 ± 0.01, which is very near threshold (0.75). Behavioral performance at the 10-ms sample duration is just above chance (0.56 ± 0.01).

Effect of cell population on the relationship between neural and behavioral performance

We have demonstrated that across the entire population of selective cells, the choice of epoch appreciably affects the relationship between neural and behavioral performance, with a tendency for individual neurons to outperform the behaving monkey at short sample durations and underperform (or equally perform) the monkey at long sample durations. Across the visual system it is not yet known precisely which populations of neurons influence perception; for example, are only the most selective neurons relevant or do less-selective neurons also influence perception? Our criterion for neural inclusion was simply *significant* selectivity at high stimulus strengths (160- and 320-ms sample durations); in contrast, in mediotemporal (MT) studies, stimulus parameters were tailored to achieve nearly *maximum* selectivity. We therefore wondered whether the overall selectivity of the neuron under study affected the relative mismatch across sample duration between neural and behavioral performance. To address this question, we divided the population into two groups based on overall selectivity at the longest two sample durations (160- and 320-ms SOA) and repeated the calculations in Fig. 9. When using a standard start and length of epoch for each cell, the population of most selective cells (SI >0.45, $n = 37$, Fig. 10*B*) had neural and behavioral performance levels that were indistinguishable at all sample durations (unpaired two-tailed *t*-test, $P > 0.05$).

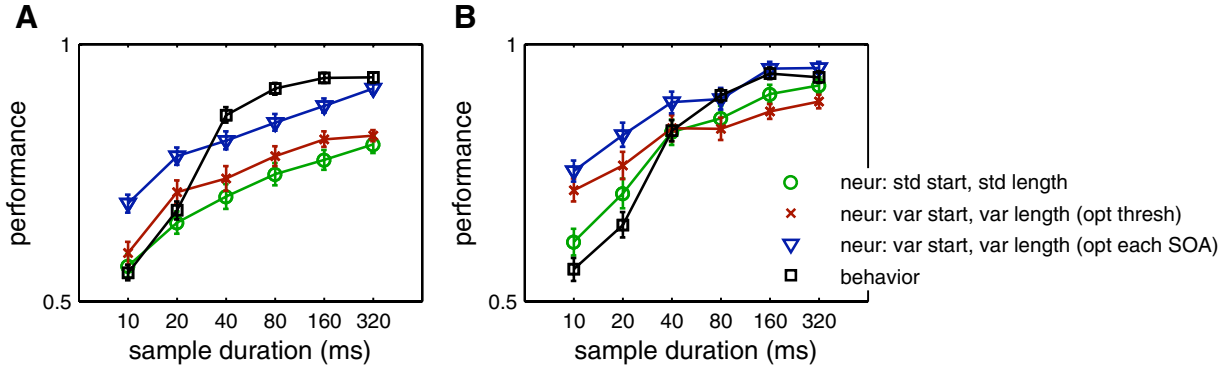

FIG. 10. Effect of selectivity requirements on neural performance. *A* and *B*: behavioral performance and neural performance as a function of sample duration in 3 different optimal epochs (described in Fig. 9) for highly selective neurons [*B*, Selectivity Index (SI) >0.45 for both the 160-ms sample duration and the 320-ms sample duration, selectivity determined in the epoch from 75 to 225 ms after sample onset; see METHODS] and all other neurons (*A*, neurons that did not meet selectivity requirements for *B*). $n = 37$ (*B*) and $n = 59$ (*A*).

Effect of mask on neural and behavioral performance

To further examine the relationship between neural and behavioral sensitivity, we performed a control task in which we recorded a subset of IT neurons while monkeys performed the DMS task with unmasked stimuli as well as masked stimuli. Masks were used to elicit a range of neural and behavioral performance; successful masks impair behavioral performance and disrupt the neural signal that differentiates between the sample stimuli. If individual neurons contain sufficient information to allow behavioral performance of the task, then in the absence of a masking stimulus, both neural and behavioral performances should improve.

For a subset of 19 neurons, after recording neural activity during performance of the DMS task with a mask (referred to as the "mask block"), we subsequently recorded neural response while the monkey performed the DMS task without the mask (referred to as the "no-mask block"). The no-mask block of trials was identical to the mask block, except that after sample onset, no mask was presented (see METHODS). In this subset of data, the neural selectivity should be greater, especially at short sample durations, because when stimuli are not masked, neural responses in the visual system persist after the presentation of a stimulus (Kovacs et al. 1995). To quantify changes in the magnitude and duration of neural selectivity, we used the ideal-observer analysis and Weibull fit (see *Neurophysiological data analysis* in METHODS) to calculate neural performance and neural threshold for each of the 19 cells where we collected data in mask and no-mask blocks.

At short sample durations, neural performance in the no-mask block of trials was much higher than neural performance in the mask block of trials. Figure 11, *A–C* shows the performance in the mask and no-mask blocks of trials for the 10-, 20-, and 80-ms sample durations. At the 10-ms sample duration, neural performance in the no-mask block of trials was significantly higher ($P < 0.05$) than neural performance in the mask block of trials in epochs ranging from 52 to 400 ms; at the 20-ms sample duration, the performance difference was significant ($P < 0.05$) for epochs ranging from 200 to 400 ms. For longer sample durations (40, 80, 160, 320 ms; data not shown except for 80-ms duration), neural performance in the no-mask block of trials was not significantly better than performance in the mask block of trials in any epoch ($P > 0.05$); the similarity in performance for the mask and no-mask conditions in the 80-ms sample duration (Fig. 11*C*) is representative of the other remaining sample durations (40, 160, 320 ms). In the complete set of mask data (of which the subpopulation of 19 cells is shown in Fig. 11, *A–C*), behavioral performance had reached a plateau at the long sample durations (80, 160, and 320 ms), and neural performance in the mask and no-mask blocks was similar for these long sample durations.

To investigate how these performance improvements affected neural thresholds and their comparison to behavioral thresholds, we optimized the epoch for each cell (procedure described in Fig. 7) separately in the mask block and the no-mask block. In Fig. 11*D*, the performance in the optimal epoch was significantly higher in the no-mask block for the

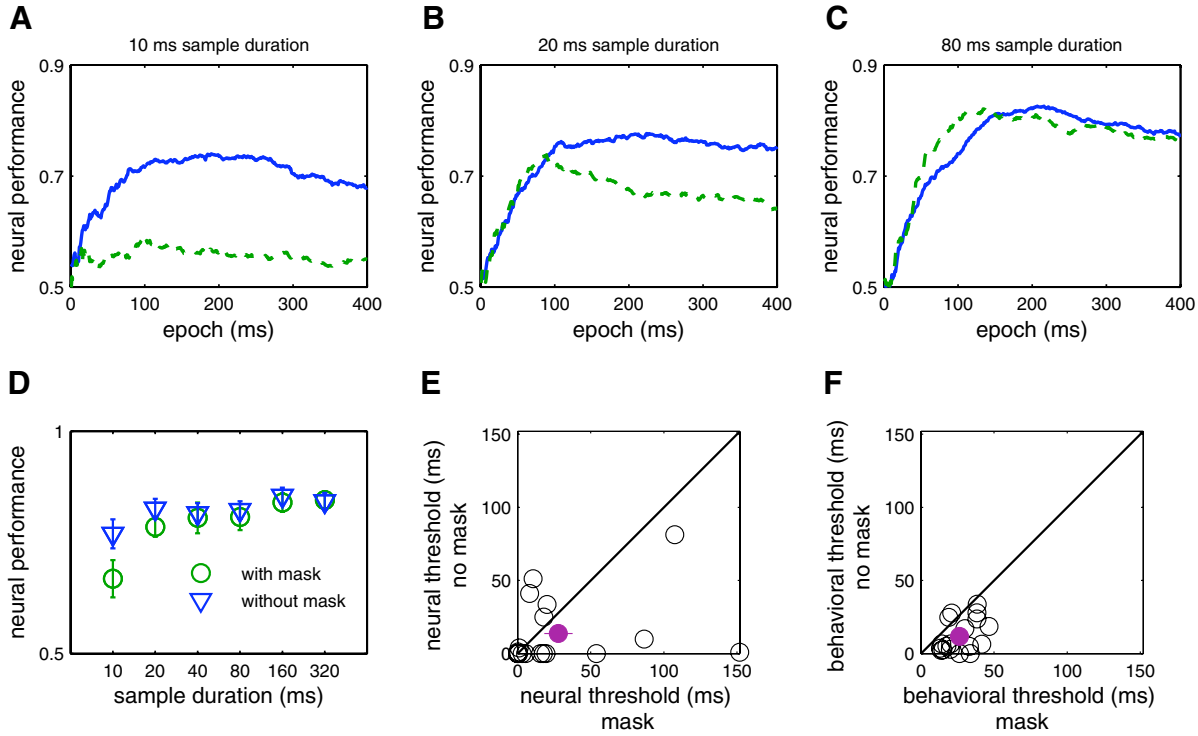

FIG. 11. Effect of mask on neural and behavioral responses. *A–C*: average neural performance ($n = 19$) as a function of epoch for the 10- (*A*), 20- (*B*), and 80-ms (*C*) sample durations in the mask (green dashed line) and no-mask (blue solid line) conditions. Neural performance values were calculated as in Fig. 6, and a standard start of epoch was used (85 ms after sample onset). *D*: average performance values in the optimal epoch for the mask (green circles) and no-mask (blue triangles) condition. Optimal epochs were defined as the epoch that minimized threshold for that cell (Fig. 7, red crosses in Fig. 9), and were defined for mask and no-mask conditions separately. *E* and *F*: individual thresholds for all neural (*E*) and behavioral (*F*) data collected in the mask (*x*-axis) and no-mask (*y*-axis) conditions. Thresholds were calculated using the cell-optimized epoch from *D*. Solid purple circles represent mean; error bars are SE in both *x*- and *y*-directions and are smaller than the data point in some cases.

shortest sample duration (mean, no-mask: 0.77 ± 0.02; mean, mask: 0.67 ± 0.04; $P < 0.005$). The improvement at the 10-ms sample duration resulted in neural performance values that were above threshold (mean = 0.76). Thus neural thresholds in the no-mask block were frequently very near 0. Neural thresholds in the no-mask block were lower than in the mask block, as shown in Fig. 11*E*.

In the no-mask block of trials, neurons reached threshold performance with sample durations of only 14 ± 5 ms. In the mask block of trials, the same neurons needed an average of 28 ± 10 ms to reach threshold. Behavioral performance was similarly influenced, with behavioral threshold changing from 12 ± 3 ms in the no-mask block of trials to 26 ± 2 ms in the mask block of trials (Fig. 11*F*). Taken together, these results suggest that our masks disrupted neural and behavioral response to some degree; in the absence of the mask, we observed comparable improvements in neural and behavioral thresholds. Furthermore, the optimal epoch lengths in the no-mask condition tended to be longer than in the mask condition (mean epoch length, no-mask: 192 ± 30 ms; mean epoch length, mask: 134 ± 25 ms), although the difference was not significant ($P = 0.12$). The start of the optimal epoch (neural response latency) was no different between mask and no-mask conditions (mean, no-mask = 100 ± 7 ms; mean, mask = 108 ± 7 ms; $P = 0.30$).

DISCUSSION

The primary focus of this study was to investigate the quantitative relationship between the responses of IT neurons to complex images and primate behavior with those images. To that end, we recorded the response of individual IT neurons while monkeys performed a threshold psychophysics task with photographic images. Consistent with previous masking studies in which neural responses were collected separately from behavior, we found that increasing sample duration improved both neural and behavioral performance (Fig. 3, Table 1; Grill-Spector and Kanwisher 2005; Keysers et al. 2001; Kovacs et al. 1995; Rieger et al. 2005; Rolls et al. 1999). Comparing simultaneously collected neural and behavioral data allows us to examine the ways in which neural responses in IT are quantitatively related to behavioral responses.

Across the population, we found individual neurons and the behaving monkey required a similar sample duration to reach threshold performance (Fig. 5, Table 1). This suggests that across stimulus sets, the responses of individual neurons to images are sufficient to underlie behavioral choices made about those images. Many previous studies have shown that IT neurons are selective for dimensions that are expected to be perceptually relevant (Baylis and Driver 2001; Freedman et al. 2003; Vogels et al. 2001), but they have not addressed the degree to which the responses of individual neurons could account for behavioral choices about images. This type of quantitative comparison of neural and behavioral sensitivities allows us to make proposals about neural populations required for behavioral function, as has been done in area MT, in the dorsal stream. In studies of MT, comparing neurometric and psychometric functions has provided valuable insights into relationships between neural and behavioral responses for a variety of visual mechanisms, such as motion perception (Britten et al. 1992, 1996), speed perception (Liu and Newsome 2005), and stereo discrimination (Uka and DeAngelis 2003). For example, Britten et al. (1996) reported that many individual neurons discriminate the direction of random-dot motion with sensitivity comparable to the behaving monkey. In contrast, Liu and Newsome (2005) reported that in the context of a speed discrimination task, psychometric thresholds were much lower than neurometric thresholds, suggesting that individual neurons are much *less* sensitive to changes in relative speed than is the behaving monkey. From these comparisons, one can infer that the neural population required to make relative judgments about speed may be larger than the neural population required to judge direction of motion (Liu and Newsome 2005).

We found that individual neurons were at least as sensitive to changes in sample duration as is the behaving monkey; by analogy to the MT studies, this would suggest that the neural population required to perform the DMS task is similar in size to the neural population required to judge direction of motion, given the observed correlation between individual neurons in IT of similar magnitude to the correlation between individual neurons in MT (although the reported correlation is slightly less in IT than in MT; for comparison see Erickson et al. 2001; Shadlen et al. 1996). However, a more detailed comparison of behavior to the dynamics of neural performance at each sample duration provides a more complex picture of the relationship between neural and behavioral sensitivities.

Surprisingly, given the complexity of the stimuli the relatively weak selectivity requirements imposed in this data, most individual neurons outperformed the monkey at short sample durations even for conservative choices of epoch (Fig. 9). There was also a tendency for the entire population of neurons to underperform the monkey at long sample durations (Fig. 9). Optimization of start and length of epoch cell by cell and sample duration by sample duration further exacerbated the discrepancy between neural and behavioral performance at short sample durations, while reducing the discrepancy between neural and behavioral performance at long sample durations (Fig. 9). There are several potential reasons for the differential mismatch between neural and behavioral performances. First, behavioral performance (in the 2AFC–DMS task) might not correctly reflect perception of the images; for example, the monkey might make nonstimulus-related errors. However, to fully account for the mismatch, nonstimulus-related errors would have to be much higher at short sample durations than at long sample durations. This seems unlikely, given that behavioral performance at short sample durations in this task was comparable to primate behavior (both monkey and human) found in other studies of backward visual masking (Grill-Spector and Kanwisher 2005; Kovacs et al. 1995). Second, it is possible that there is not a one-to-one map between the "features" represented by individual IT neurons and the "features" used by monkeys to perform the task; thus the monkey could rely partially on features not represented by the neurons under study and this could explain why neural performance was better than behavioral performance at short sample durations. This idea is supported by the observation that even though there was substantial variability in behavioral thresholds, there was no correlation between neural and behavioral thresholds (see *Behavioral and neural performance* in RESULTS). Finally, the population of neurons we recorded from might be larger than the population most relevant to percep-

tion. For example, our inclusion criterion was relatively lax compared with the MT studies (significant selectivity rather than maximal selectivity; see RESULTS). This explanation is supported by the fact that there was little discrepancy between neural and behavioral performance at each sample duration when we examined only the most selective cells (green circles and black squares, Fig. 10).

In addition to relating the magnitude of neural and behavioral sensitivities, this study also provides insight into how neural sensitivity evolves over time. Here we show that neural sensitivity is dependent on the epoch used to collect the spike counts that determine neural performance. Previous studies of IT response dynamics have suggested that early portions of neural response reflect general information, such as object category or constituent parts, whereas later activity reflects more specific information such as object identity (Brincat and Connor 2006; Sugase et al. 1999). However, the dynamics of stimulus representation and how it relates to behavior have yet to be explained when backward masking affects normal response dynamics. Previous studies of the neural effects of backward masking in IT, obtained while monkeys maintained passive fixation on images, have suggested that mask presentation either interrupts neural response to the sample image or integrates the response to the mask and the sample stimulus together (Enns and Di Lollo 2000, 2002; Keysers and Perrett 2002; Keysers et al. 2001; Kovacs et al. 1995; Rolls et al. 1999). In either case (interruption or integration), previous studies have reported that masking confines informative neural response in IT to epochs similar to (or ~60 ms longer than) the length of sample or target presentation (Keysers et al. 2001; Kovacs et al. 1995). In the case of short sample durations, only the early signal, putatively more general, would be present.

In the present study, we report that the epoch that maximizes neural sensitivity is considerably longer than the sample duration, even after adjusting cell by cell for the onset of neural response (Figs. 6 Figs. 7 Figs. 8 Figs. 9). For the 10-, 20-, and 40-ms sample durations, the optimal epoch is significantly >60 ms past the length of sample duration ($P < 0.001$). Long optimal epochs are not solely an artifact of insufficient masking because neural and behavioral thresholds both decreased when the mask was not shown (Fig. 11). The optimal epoch lengths, which ranged from about 100 to 200 ms depending on the analysis parameters, are long enough to include the later signal that is putatively more object specific (Brincat and Connor 2006; Matsumoto et al. 2005; Sugase et al. 1999), which could allow performance of the task (albeit imperfect) at short sample durations. One possible explanation for the prolonged selectivity of IT neurons for masked stimuli while monkeys are engaged in the task is recurrent processing within IT in combination with feedforward inputs that could be selectively gated by attention.

Recurrent or reentrant processing has been hypothesized to explain the dynamics of neural response is several visual areas, including V1, MT, and IT (Brincat and Connor 2006; Lamme et al. 2002; Pack et al. 2001). In the data reported here, recurrent processing could prolong selectivity to masked stimuli in the following fashion. Immediately after stimulus onset, an individual IT neuron would favor synaptic inputs from neurons in lower cortical areas, neurons that carry unmasked signals from V1, and suppress synaptic inputs from within IT. Studies of V1 and masking suggest that signals from V1 are unmasked under the conditions used in this IT study: the selectivity of V1 neurons for simple stimulus properties, such as orientation, is not affected by masking (Lamme et al. 2002). After mask onset, IT neurons would selectively favor synaptic inputs from neurons within IT or higher cortical areas and selectively dampen or weaken synaptic inputs from neurons in lower cortical areas, which are now dominated by the response to the mask presentation and are thus noninformative. Although the data in this study do not directly bear on the neural mechanism for selectively favoring one type of synaptic input over another, it could be implemented by low-level neurons that detect transients. To summarize, neural sensitivity to a sample stimulus could be partially maintained, even after onset of a masking stimulus, by a temporal gating mechanism that allows IT neurons to partially block the feedforward input from the mask and recurrently process the unmasked information that has already reached IT. Such a temporal gating mechanism could be mediated by attention.

Attention can modulate the behavioral effects of backward visual masking. For example, when psychophysical observers attend to a target stimulus, they can ignore distracting stimuli presented closely in time (Sheppard et al. 2002; Shih 2000) and attention can almost eliminate otherwise powerful behavioral-masking effects (Enns and Di Lollo 2000). In addition, attention can modulate the magnitude of neural response in IT (Moran and Desimone 1985; Olson 2001). Based on long optimal epochs and close correspondence between sensitivity of individual neurons and behavioral performance in the DMS task, we posit that attention modulates the temporal dynamics of IT neural response, as well as the magnitude of response, and this modulation may account for the behavioral effects of backward visual masking.

Decades of experimental work have shown that IT cortex is important for performing behavioral tasks with complex objects (Ungerleider and Mishkin 1982). The results from this study show that calculations made from subpopulations of IT neurons can mimic performance of the monkey in making perceptual judgments about complex images. Particularly at short sample durations, most individual neurons perform at least as well as the behaving monkey; furthermore, the dynamics of the relationship between neural and behavioral sensitivities provide insight into how neural signals are combined in IT. These results demonstrate that quantitative tools derived to describe dorsal stream processing of motion and other modalities also provide useful insights into ventral stream processes.

ACKNOWLEDGMENTS

We thank J. Skiver Thompson and K. M. Ahl for technical help and M. N. Shadlen, F. Rieke, and J. I. Gold for comments on an earlier version of this manuscript. In addition, we are pleased to make our data set available to other scientists; please contact the corresponding author with requests.

GRANTS

This research was supported by the Royalty Research Foundation (University of Washington), the Sloan Foundation, the McKnight Foundation, the Whitehall Foundation, and National Center for Research Resources. S. R. Allred was supported by National Eye Institute Vision Research Training Grant T32 EY-07031.

REFERENCES

Afraz SR, Kiani R, Esteky H. Microstimulation of inferotemporal cortex influences face categorization. *Nature* 442: 692–695, 2006.

Allred S, Liu Y, Jagadeesh B. Selectivity of inferior temporal neurons for realistic pictures predicted by algorithms for image database navigation. *J Neurophysiol* 94: 4068–4081, 2005.

Allred SR. The neural basis of visual object recognition. In: *Neurobiology and Behavior*. Seattle, WA: Univ. of Washington, 2006.

Bacon-Mace N, Mace MJ, Fabre-Thorpe M, Thorpe SJ. The time course of visual processing: backward masking and natural scene categorisation. *Vision Res* 45: 1459–1469, 2005.

Baylis GC, Driver J. Shape-coding in IT cells generalizes over contrast and mirror reversal, but not figure-ground reversal. *Nat Neurosci* 4: 937–942, 2001.

Brincat SL, Connor CE. Dynamic shape synthesis in posterior inferotemporal cortex. *Neuron* 49: 17–24, 2006.

Britten KH, Newsome WT, Shadlen MN, Celebrini S, Movshon JA. A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Vis Neurosci* 13: 87–100, 1996.

Britten KH, Shadlen MN, Newsome WT, Movshon JA. The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J Neurosci* 12: 4745–4765, 1992.

Cahusac PM, Miyashita Y, Rolls ET. Responses of hippocampal formation neurons in the monkey related to delayed spatial response and object-place memory tasks. *Behav Brain Res* 33: 229–240, 1989.

Chelazzi L, Duncan J, Miller EK, Desimone R. Responses of neurons in inferior temporal cortex during memory-guided visual search. *J Neurophysiol* 80: 2918–2940, 1998.

Enns JT, Di Lollo V. What's new in visual masking? *Trends Cogn Sci* 4: 345–352, 2000.

Enns JT, Di Lollo V. What competition? *Trends Cogn Sci* 6: 118, 2002.

Erickson CA, Jagadeesh B, Desimone R. Clustering of perirhinal neurons with similar properties following visual experience in adult monkeys. *Nat Neurosci* 3: 1143–1148, 2000.

Freedman DJ, Riesenhuber M, Poggio T, Miller EK. A comparison of primate prefrontal and inferior temporal cortices during visual categorization. *J Neurosci* 23: 5235–5246, 2003.

Fuchs AF, Robinson DA. A method for measuring horizontal and vertical eye movement chronically in the monkey. *J Appl Physiol* 21: 1068–1070, 1966.

Green DM, Swets JA. *Signal Detection Theory and Psychophysics*. New York: Wiley, 1966, p. 455.

Grill-Spector K, Kanwisher N. Visual recognition: as soon as you know it is there, you know what it is. *Psychol Sci* 16: 152–160, 2005.

Keysers C, Perrett DI. Visual masking and RSVP reveal neural competition. *Trends Cogn Sci* 6: 120–125, 2002.

Keysers C, Xiao DK, Foldiak P, Perrett DI. The speed of sight. *J Cogn Neurosci* 13: 90–101, 2001.

Kourtzi Z, DiCarlo JJ. Learning and neural plasticity in visual object recognition. *Curr Opin Neurobiol* 16: 152–158, 2006.

Kovacs G, Vogels R, Orban GA. Cortical correlate of pattern backward masking. *Proc Natl Acad Sci USA* 92: 5587–5591, 1995.

Leopold DA, Bondar IV, Giese MA. Norm-based face encoding by single neurons in the monkey inferotemporal cortex. *Nature* 442: 572–575, 2006.

Liu J, Newsome WT. Correlation between speed perception and neural activity in the middle temporal visual area. *J Neurosci* 25: 711–722, 2005.

Logothetis NK, Pauls J, Poggio T. Shape representation in the inferior temporal cortex of monkeys. *Curr Biol* 5: 552–563, 1995.

Matsumoto N, Okada M, Sugase-Miyamoto Y, Yamane S. Neuronal mechanisms encoding global-to-fine information in inferior-temporal cortex. *J Comput Neurosci* 18: 85–103, 2005.

Moran J, Desimone R. Selective attention gates visual processing in the extrastriate cortex. *Science* 229: 782–784, 1985.

Olson CR. Object-based vision and attention in primates. *Curr Opin Neurobiol* 11: 171–179, 2001.

Op de Beeck H, Wagemans J, Vogels R. Inferotemporal neurons represent low-dimensional configurations of parameterized shapes. *Nat Neurosci* 4: 1244–1252, 2001.

Rainer G, Lee H, Logothetis NK. The effect of learning on the function of monkey extrastriate visual cortex. *PLoS Biol* 2: E44, 2004.

Rieger JW, Braun C, Bulthoff HH, Gegenfurtner KR. The dynamics of visual pattern masking in natural scene processing: a magnetoencephalography study. *J Vis* 5: 275–286, 2005.

Rolls ET, Tovee MJ, Panzeri S. The neurophysiology of backward visual masking: information analysis. *J Cogn Neurosci* 11: 300–311, 1999.

Sereno AB, Maunsell JH. Shape selectivity in primate lateral intraparietal cortex. *Nature* 395: 500–503, 1998.

Shadlen MN, Britten KH, Newsome WT, Movshon JA. A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *J Neurosci* 16: 1486–1510, 1996.

Sheinberg DL, Logothetis NK. The role of temporal cortical areas in perceptual organization. *Proc Natl Acad Sci USA* 94: 3408–3413, 1997.

Sheinberg DL, Logothetis NK. Noticing familiar objects in real world scenes: the role of temporal cortical neurons in natural vision. *J Neurosci* 21: 1340–1350, 2001.

Sheppard DM, Duncan J, Shapiro KL, Hillstrom AP. Objects and events in the attentional blink. *Psychol Sci* 13: 410–415, 2002.

Shih SI. Recall of two visual targets embedded in RSVP streams of distractors depends on their temporal and spatial relationship. *Percept Psychophys* 62: 1348–1355, 2000.

Sigala N, Logothetis NK. Visual categorization shapes feature selectivity in the primate temporal cortex. *Nature* 415: 318–320, 2002.

Sugase Y, Yamane S, Ueno S, Kawano K. Global and fine information coded by single neurons in the temporal visual cortex. *Nature* 400: 869–873, 1999.

Uka T, DeAngelis GC. Contribution of middle temporal area to coarse depth discrimination: comparison of neuronal and psychophysical sensitivity. *J Neurosci* 23: 3515–3530, 2003.

Uka T, Tanabe S, Watanabe M, Fujita I. Neural correlates of fine depth discrimination in monkey inferior temporal cortex. *J Neurosci* 25: 10796–10802, 2005.

Ungerleider L, Mishkin M. Two cortical visual systems. In: *Analysis of Visual Behavior*, edited by Ingle DJ, Goodale MA, Mansfield RJ. Cambridge, MA: MIT Press, 1982, p. 549–586.

VanRullen R, Thorpe SJ. Is it a bird? Is it a plane? Ultra-rapid visual categorisation of natural and artifactual objects. *Perception* 30: 655–668, 2001.

Vogels R, Biederman I, Bar M, Lorincz A. Inferior temporal neurons show greater sensitivity to nonaccidental than to metric shape differences. *J Cogn Neurosci* 13: 444–453, 2001.

Wichmann FA, Hill NJ. The psychometric function: I. Fitting, sampling, and goodness of fit. *Percept Psychophys* 63: 1293–1313, 2001.