SUPPLEMENTARY RESULTS

Controls for differences between MT and LIP sessions and activity

Here we address several potential differences between the MT and LIP data sets, and demonstrate that they can not account for the observed attentional latency difference.

Firing rate

There were no significant differences in the mean sustained spike rate (pre-switch) between the LIP and MT populations (LIP vs. MT mean in spikes/s ± SD, t test; monkey B: 53 ± 29 vs. 58 ± 32, $p = 0.50$; monkey M: 37 ± 25 vs. 42 ± 26, $p = 0.35$).

Stimulus placement

Stimulus placement was dictated by the receptive field location of the individual neurons. The eccentricities of the MT and LIP receptive fields were well balanced (median eccentricity in degrees, Wilcoxon rank sum test: monkey B: MT: 10.1, LIP: 11.2, $p = 0.52$; monkey M: MT: 11.3, LIP: 12.0, $p = 0.53$). MT cells had receptive fields in the lower visual field more frequently than LIP cells (monkey B: MT: 29/31 (93%), LIP: 22/63 (35%); monkey M: MT: 32/36 (89%), LIP: 38/55 (69%)). This is consistent with the known bias of area MT's spatial representation towards lower visual fields (Maunsell and Van Essen, 1987).

We asked whether differences in stimulus placement could be responsible for the areal differences in attentional latency that we observed. Within each monkey's LIP population there was no difference in the attentional latencies for those cells with receptive fields above versus below the horizontal meridian (above vs. below median in ms, rank-sum test, monkey B: 222 vs. 251, $p = 0.49$; monkey M: 179 vs. 156, $p = 0.10$). Consistent with this, the subset of MT and LIP

cells with receptive fields below the horizontal meridian showed similar latency differences (LIP vs MT median latency in ms, rank-sum test, monkey B: 251 vs. 289, $p = 0.02$; monkey M: 157 vs. 230, $p < 0.001$).

Behavior

Single session behavior was broadly overlapping between MT and LIP sessions. However, there were differences in average behavior between MT and LIP sessions that, though small, were generally statistically significant (Supplementary Table 1). To quantify behavior in each session we used four different behavioral metrics, (1) fraction of speed pulses detected on valid trials, (2) reaction time on valid trials, (3) difference in fraction detected between valid and invalid trials and (4) difference in reaction time between valid and invalid trials. These four behavioral metrics for the two monkeys are shown in Supplementary Table 1.

It is possible that the difference in attentional switch latencies between MT and LIP populations was not related to areal differences but rather to differences in behavior between the MT and LIP sessions. To address this possibility we calculated correlation coefficients between single unit latencies and single session behavior for each neural population and behavioral metric. Of these 16 comparisons (4 metrics * 2 monkeys * 2 neural populations) only one exhibited a statistically significant correlation ($p < 0.05$, not corrected for multiple comparisons). Specifically, the fraction of speed pulses detected was weakly negatively correlated with monkey B's attentional latencies across the LIP population ($r = -0.41$, $p = 0.01$).

Is this a spurious correlation emerging from multiple comparisons? Notably, there was not even a weak correlation between single unit latencies and behavior for monkey B's MT population ($r = -0.09$, $p = 0.80$), or either population in monkey M (MT: $r = -0.17$, $p = 0.41$, LIP:

$r = -0.07$, $p = 0.68$). Regardless, what is important to our conclusions is whether variance in this behavioral metric accounted for any of the difference in attentional latency that we ascribe to a difference in brain area. In a multiple regression analysis we see that this is not the case. The variance accounted for by each variable is nearly independent of the other; the R^2 values from the single-variable regressions sum to equal the R^2 of the multiple regression (Supplementary Table 2). The addition of the Area variable improves the model significantly ($p = 0.002$, partial F test for addition of the Area variable). In summary, we did not observe a general trend towards correlation between a cell's attentional latency and the animal's behavior in that session. Variation in attentional switch latencies is well explained by the areal identity of the cells, and is not explained by any of the behavioral metrics tested.

Intersession differences

For monkey B, MT and LIP neurons were recorded in interleaved blocks of sessions (3 blocks of MT sessions and 2 blocks of LIP sessions). An additional test of the consistency of our findings is whether the interareal latency difference is evident in the pairwise block-by-block comparisons. Supplementary Figure 3 shows the mean attentional latency +/- standard error for the neurons in each of the 3 MT recording blocks and 2 LIP recording blocks in the order they were collected. The latency difference is apparent in the pairwise comparisons between different blocks. Although limited by the small number of cells in each block, several of the pairwise comparisons were statistically significant ($p < 0.05$, t test, not corrected for multiple comparisons) denoted by bridging lines in Supplementary Figure 3. The results for the individual comparisons are listed below with the number of neurons for each session shown in parentheses. None of the intra-areal comparisons were statistically significant.

	MT session 1 (n=7)	MT session 2 (n=7)	MT session 3 (17)
LIP session 1 (n=27)	p = 0.047	p = 0.316	p = 0.147
LIP session 2 (n=36)	p = 0.003	p = 0.070	p = 0.020

Analysis of visual response latencies in MT and LIP

We have shown that attentional signals emerge earlier in LIP than MT during rapid shifts of attention suggesting that they could arrive in MT at least in part via a pathway that includes LIP. In contrast, we would predict based on anatomical projections that visual signals would appear first in MT before LIP. Although a latency hierarchy for visual response latencies is well described, the latencies for dorsal stream visual areas (MT, MST, V3 and even FEF) are broadly overlapping (Schmolesky et al., 1998). To our knowledge no study has compared latencies between MT and LIP in a head-to-head fashion. However, LIP latencies can exhibit mean latencies as fast as 40-50 ms depending on the stimulus (Bisley et al., 2004) which overlaps with previous estimates in MT (Bair et al., 2002; Lisberger and Movshon, 1999; Raiguel et al., 1999).

We measured the visual response latency to the onset of the moving dots in our attention task. We refined the convolve-threshold method to make it more appropriate for detecting abrupt visual response latencies by reducing the degree of smoothing (Gaussian with 2 ms standard deviation) and requiring the spike rate to cross a rate threshold 3-standard-deviations above a baseline calculated in the 300 ms before dot onset, and to remain there for at least 10 ms. Latencies were considered reliable if they exhibited a standard deviation of less than 10 ms (by bootstrap, see Methods).

Consistent with previous work showing that attention does not affect response latency (Lee et al., 2007), there was no difference between latencies in the attend-in versus attend-out

condition (all intra-areal comparisons $p > 0.8$, paired signed rank test) and we combined attend-in and attend-out data to more accurately estimate single-neuron latencies. Supplementary Figure 4 shows the population response and distribution of single-neuron latencies for each area in each monkey. The visual latency was significantly earlier in MT in monkey B but not in monkey M (median MT vs. LIP latency, rank-sum test, monkey M: 114 vs. 111 ms, $p = 0.93$, monkey B: 115 vs. 144 ms, $p = 0.007$).

It is not clear why the results were different in the two animals. Our main task was not well suited to studying visual response latencies for several reasons. We used a low- luminance stimulus (appreciably dimmer than those used previously) that resulted in long latency responses with less abrupt onsets (Maunsell et al., 1999; Maunsell and Gibson, 1992). In addition each low luminance dot patch was surrounded by a brighter annulus that had a greater effect on LIP than MT baseline activity (Herrington and Assad, 2009). In theory, either or both of these could have contributed to our not detecting an appreciable latency difference in one of the animals.

This analysis also revealed an unexpected nuance in the stimulus onset response that had previously been obscured by smoothing. Some of the neurons exhibited a biphasic response consisting of an initial burst, a brief pause and then a second burst with a more sustained response. This was particularly evident in the population response of MT neurons in monkey B (Supplementary Figure 4). This response profile was not the result of averaging across two populations of neurons with distinct latencies as a similar response was evident in single-neuron responses (Supplementary Figure 5). This response profile was common in monkey B MT neurons (roughly ~30% by eye). In contrast it was less common and less convincing in monkey B LIP (5-10%) and rare in either area in monkey M (<10% MT, <5% LIP) although there were still a few clear examples (Supplementary Figure 5E). Such biphasic responses have previously

been described in LIP (Bisley et al., 2004) but not, to our knowledge, in MT. We considered the possibility that this could reflect a small eye movement time-locked on stimulus onset. Indeed, we have previously described neurophysiologic effects of microsaccades in this data set (Herrington et al., 2009). However, elimination of trials with microsaccades from -300 to +300 ms around dot onset did not substantially alter the response profile.

To assess whether the response profile and latency findings were stimulus-specific, we also determined visual response latencies using data from our direction-mapping task. This task involved passive fixation during display of a sequence of random dot fields equivalent to those used in the main task except that they were brighter (although still relatively dim, monkey M 0.9 cd/m^2: monkey B: 1.7 cd/m^2) and were not surrounded by annuli. The dot fields appeared and remained stationary for at least 300 ms before motion started, and it is the response onset to these stationary dot fields that we analyzed. Median latencies were shorter in MT than in LIP for both monkeys (median MT vs. LIP latency, rank-sum test, monkey M: 67 vs. 83.5 ms, $p = 0.007$; monkey B: 58 vs. 77 ms, $p = 0.003$). Although the earliest latencies were similar in the two areas (as evidenced in the population response), LIP neurons had a broader distribution of response latencies and hence a longer median latency.

Across the population, LIP activity arose more gradually than MT activity. In theory this could confound our latency detection as more slowly rising responses may be detected later even if they began simultaneously. However, the slower rise could have arisen if either 1) individual LIP neurons rose to peak more slowly or 2) LIP neurons had a wider distribution of response latencies but each neuron's response onset was as steep as in MT. To distinguish these possibilities we recalculated the population response after aligning each neuron on its response latency. The resulting population response shows a very similar slope of the initial response in

MT and LIP (Supplementary Figure 6E,F). We conclude that LIP neurons had a similar response profile to MT neurons, but a broader distribution of latencies and hence a longer median latency.

In conclusion, visual responses appear earlier in MT than LIP after the onset of a visual stimulus. This is consistent with the expected visual hierarchy and is in contrast to the emergence of attentional modulation which appears first in area LIP, consistent with a top-down flow of attentional signals.

SUPPLEMENTARY REFERENCES

Bair, W., Cavanaugh, J.R., Smith, M.A., and Movshon, J.A. (2002). The timing of response onset and offset in macaque visual neurons. J Neurosci *22*, 3189-3205.

Bisley, J., Krishna, B.S., and Goldberg, M. (2004). A rapid and precise on-response in posterior parietal cortex. J Neurosci *24*, 1833-1838.

Herrington, T.M., and Assad, J.A. (2009). Neural activity in the middle temporal area and lateral intraparietal area during endogenously cued shifts of attention. J. Neurosci. In press.

Herrington, T.M., Masse, N.Y., Hachmeh, K.J., Smith, J.E., Assad, J.A., and Cook, E.P. (2009). The effect of microsaccades on the correlation between neural activity and behavior in middle temporal, ventral intraparietal, and lateral intraparietal areas. J Neurosci *29*, 5793-5805.

Lee, J., Williford, T., and Maunsell, J.H. (2007). Spatial attention and the latency of neuronal responses in macaque area V4. J Neurosci *27*, 9632-9637.

Lisberger, S.G., and Movshon, J.A. (1999). Visual motion analysis for pursuit eye movements in area MT of macaque monkeys. J Neurosci *19*, 2224-2246.

Maunsell, J.H., Ghose, G.M., Assad, J.A., McAdams, C.J., Boudreau, C.E., and Noerager, B.D. (1999). Visual response latencies of magnocellular and parvocellular LGN neurons in macaque monkeys. Vis Neurosci *16*, 1-14.

Maunsell, J.H., and Gibson, J.R. (1992). Visual response latencies in striate cortex of the macaque monkey. J Neurophysiol *68*, 1332-1344.

Maunsell, J.H., and Van Essen, D.C. (1987). Topographic organization of the middle temporal visual area in the macaque monkey: representational biases and the relationship to callosal connections and myeloarchitectonic boundaries. J Comp Neurol *266*, 535-555.

Raiguel, S.E., Xiao, D.K., Marcar, V.L., and Orban, G.A. (1999). Response latency of macaque area MT/V5 neurons and its relationship to stimulus parameters. J Neurophysiol *82*, 1944-1956.

Schmolesky, M.T., Wang, Y., Hanes, D.P., Thompson, K.G., Leutgeb, S., Schall, J.D., and Leventhal, A.G. (1998). Signal timing across the macaque visual system. J Neurophysiol *79*, 3272-3278.

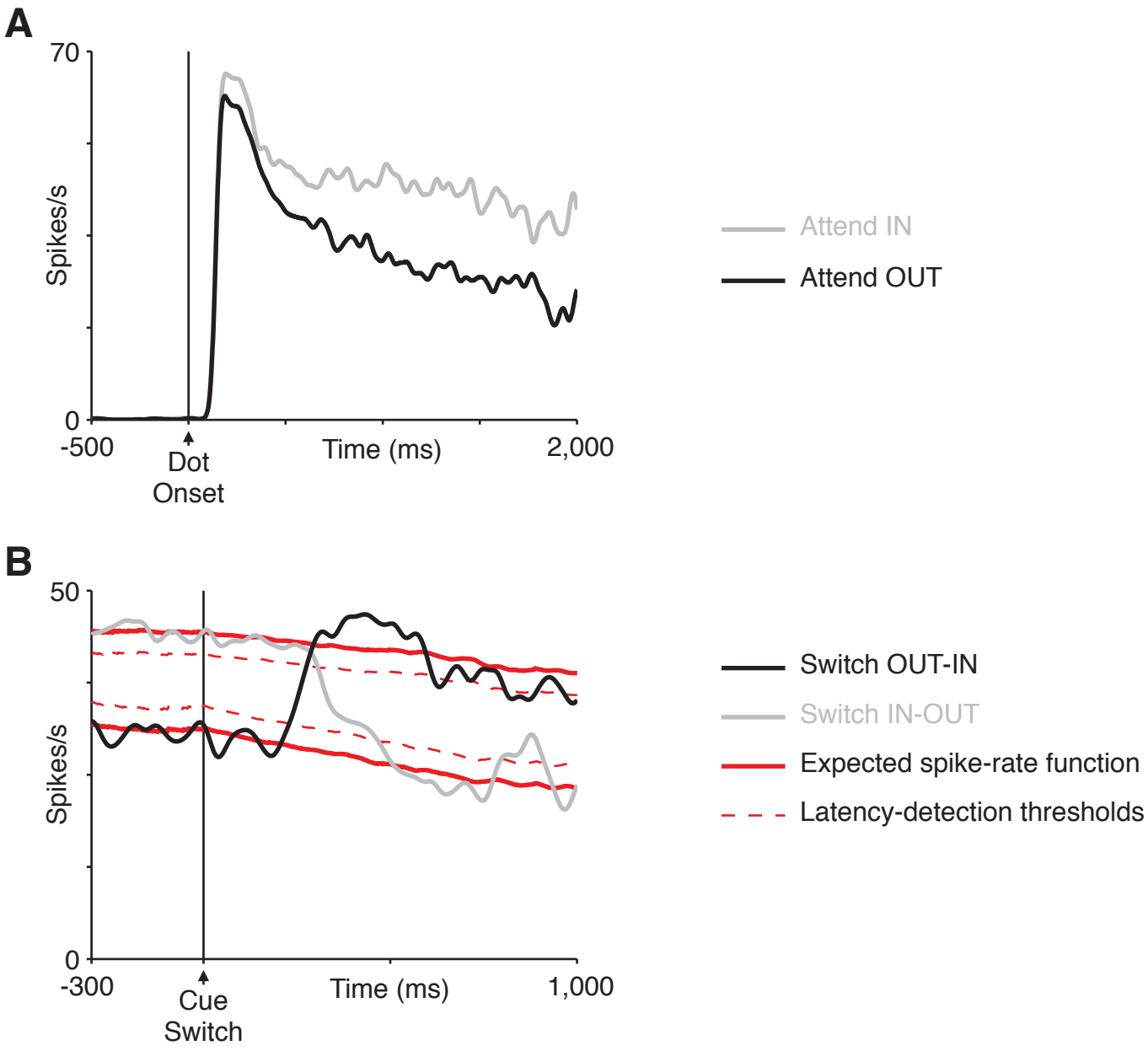
Monkey	Behavioral metric	MT mean ± SD	LIP mean ± SD	t-test p-value
B	RT: valid	477 ± 21	467 ± 22	0.043
B	FD: valid	0.67 ± 0.09	0.70 ± 0.10	0.158
B	RT: valid-invalid	-5.4 ± 7.9	-10.9 ± 7.7	0.002
B	FD: valid-invalid	0.19 ± 0.07	0.24 ± 0.08	0.002
M	RT: valid	420 ± 21	412 ± 17	0.002
M	FD: valid	0.77 ± 0.06	0.80 ± 0.06	0.042
M	RT: valid-invalid	-20.8 ± 13.5	-23.6 ± 18.0	0.422
M	FD: valid-invalid	0.49 ± 0.09	0.45 ± 0.13	0.103

Supplementary Table 1: Behavior in MT and LIP sessions. Quantification of single session behavior by four different metrics: reaction time (RT) on valid trials, fraction of speed pulses detected (FD) on valid trials, difference in RT between valid and invalid trials and differences in FD between valid and invalid trials. Data is presented as mean ± s.d.; p-value from t-test.

Model	Variables	Regression coefficient	95% C.I.	R^2
1	Area (MT/LIP)	-54	-85 … -22	0.19
2	FD: cued	-205	-351 … -58	0.13
3	Area (MT/LIP)	-49	-79 … -19	0.29
	FD: cued	-177	-313 … -43	

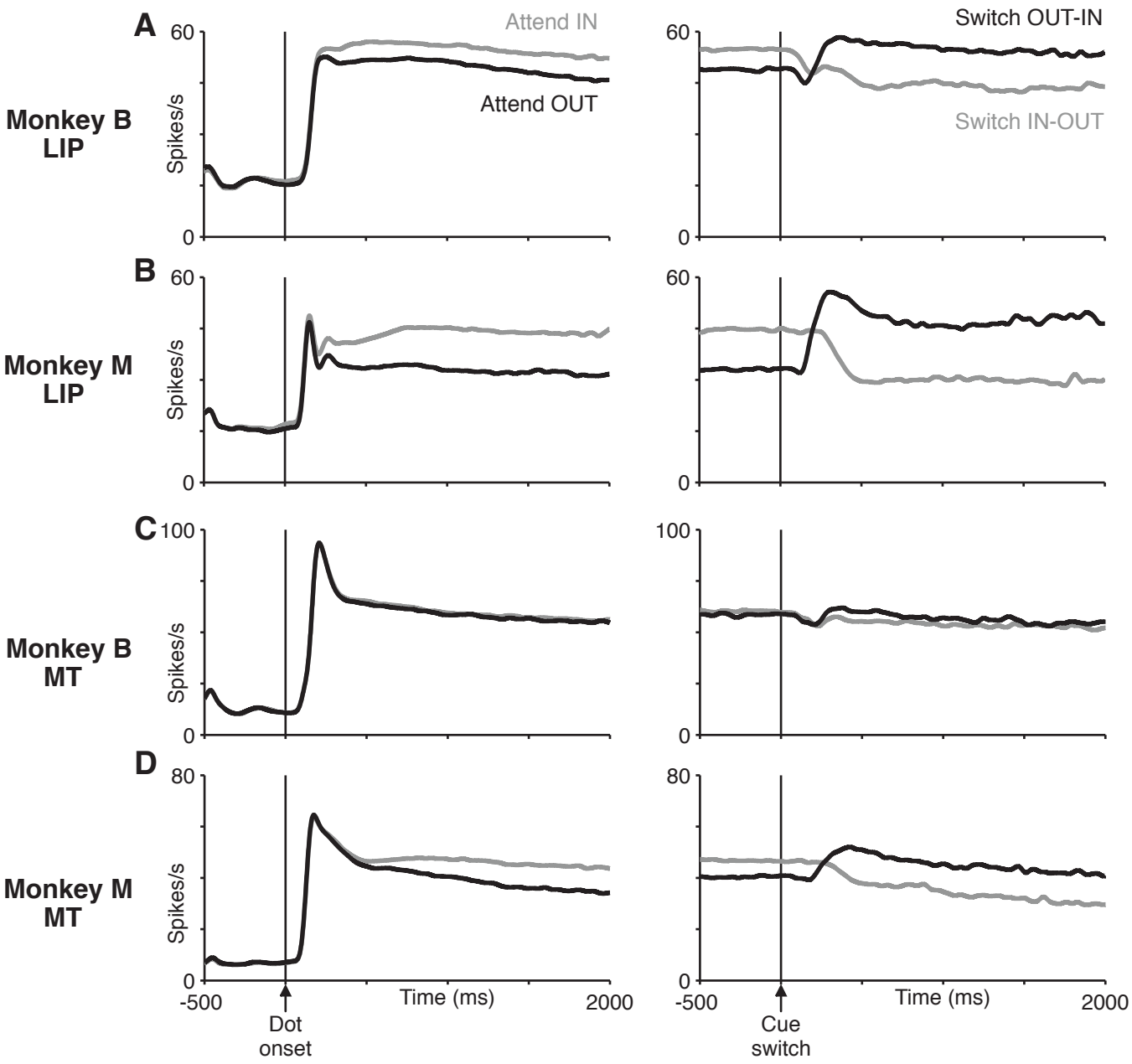
Supplementary Table 2: Linear regression model for monkey B. Single-variable regression and multivariable-regression model coefficients, 95 % confidence intervals for the coefficients and R^2 values for the model. Areal variable was (MT = 0, LIP = 1) and FD variable was fraction of speed pulses detected on validly cued trials each regressed against the single unit attentional latencies either separately (models 1 and 2) or together (model 3).

Supplementary Figure 1



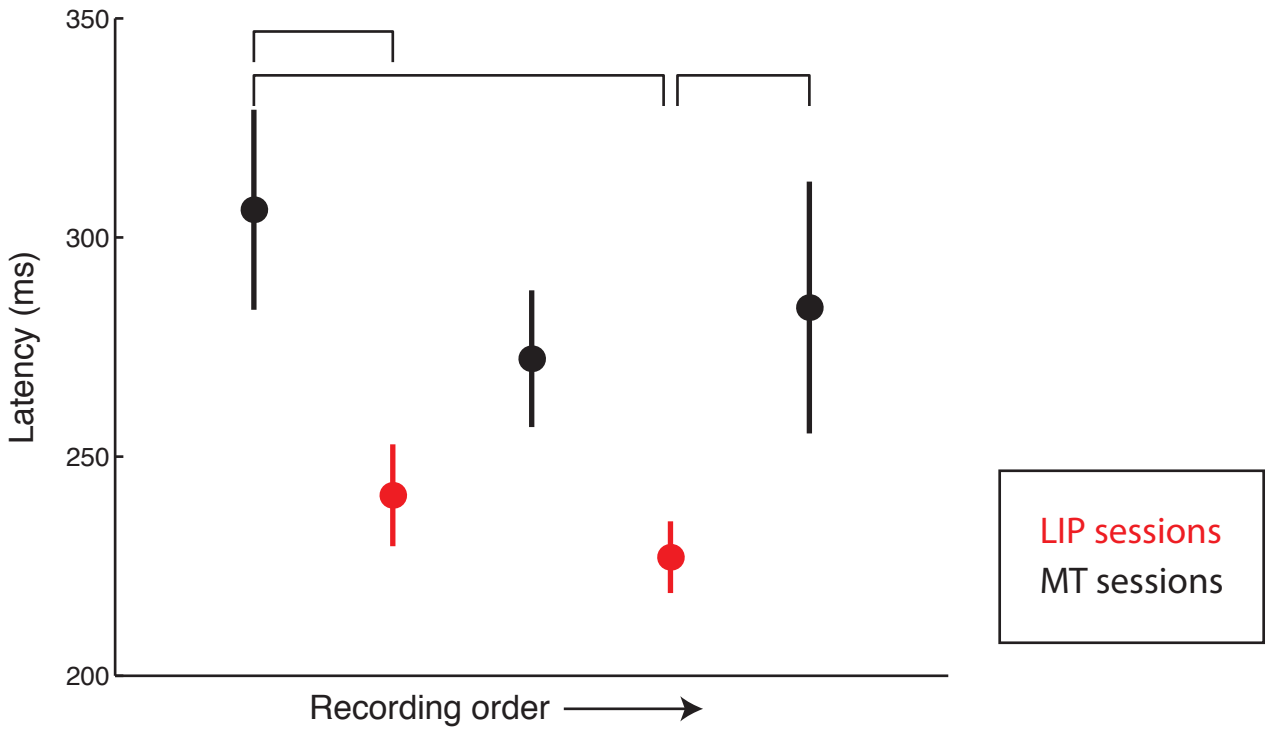
Supplementary Figure 1: Deviation-threshold method for detecting single neuron attentional latencies. See Methods for a full description of the analysis. **A,** Activity for an example neuron (area LIP, monkey M) aligned on dot onset. **B**, Peri-switch activity for the same neuron as in (**A**). The expected spike-rate function (red) is superimposed on the actual spike-rate function. The latency thresholds (dashed red lines) are equal to the expected spike-rate function ± 3 times the standard deviation of the actual spike rate function from -300 to 0 ms before the cue switch.

Supplementary Figure 2



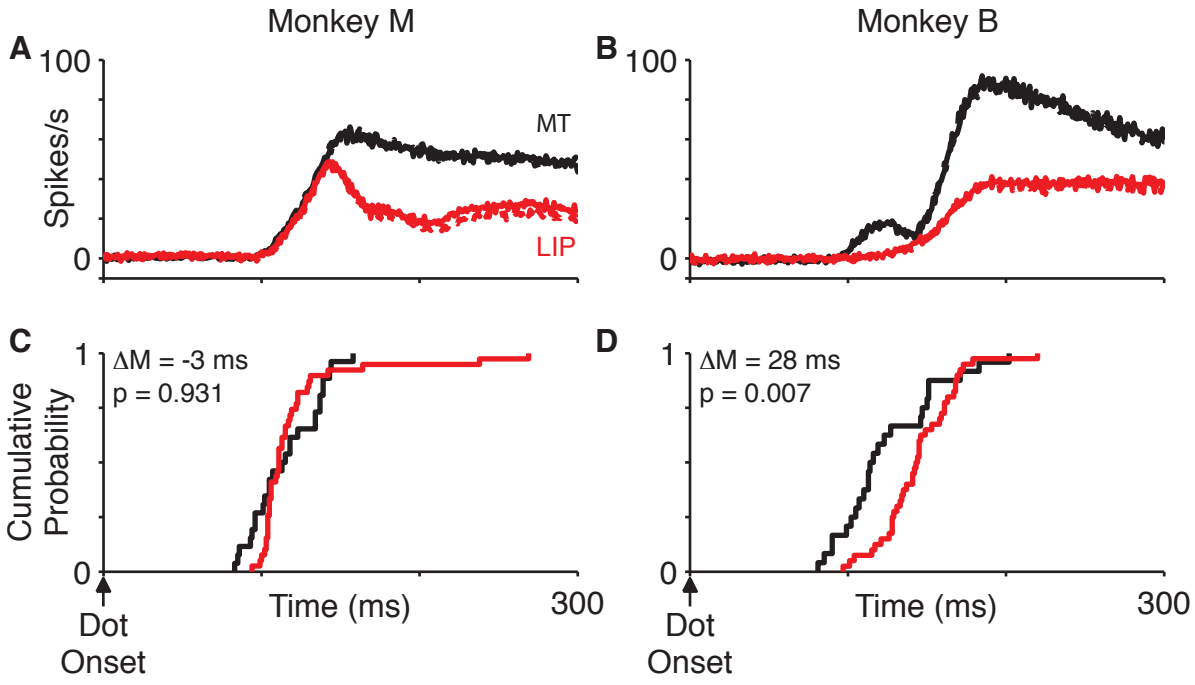
Supplementary Figure 2: Population response aligned on dot onset (left panels) or cue switch (right panels) for monkey B LIP (**A**), monkey M LIP (**B**), monkey B MT (**C**) and monkey M MT (**D**). Attentional modulation begins to emege prior to dot onset in LIP but not in MT. See the letter to reviewers for details.

Supplementary Figure 3



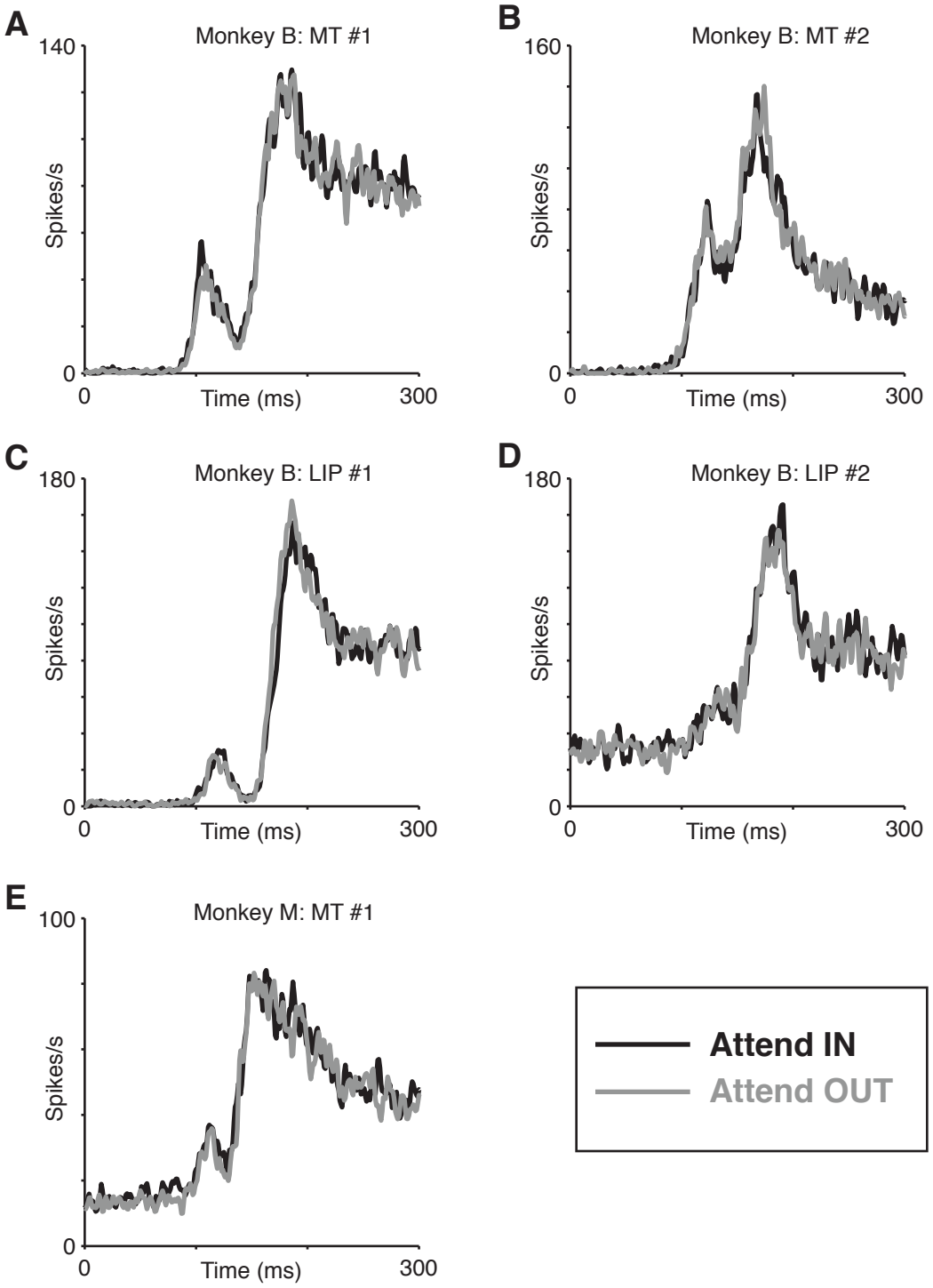
Supplementary Figure 3: Attentional latency across behavioral sessions for money B. MT and LIP recordings were conducted in interleaved blocks of sessions. The mean attentional latencies +/- standard errors are shown for each of the 3 MT (black) and 2 LIP (red) recording sessions. The latency difference is apparent in the pairwise comparisons between different sessions. The pairwise combinations reaching statistical significance are shown by bridging lines ($p < 0.05$, t test, not corrected for multiple comparisons). None of the intra-areal comparisons were statistically significant. See Supplementary Results for details.

Supplementary Figure 4



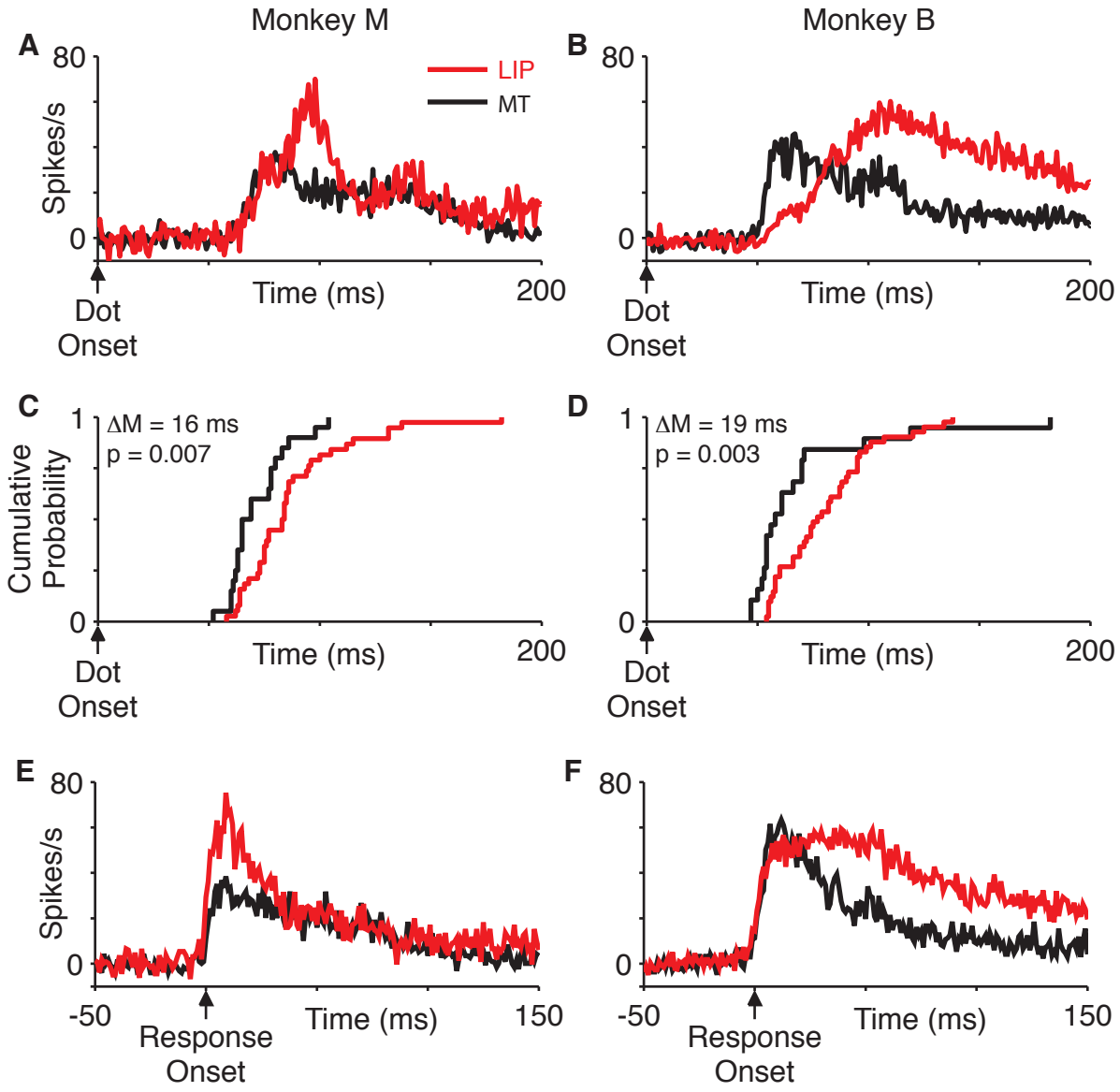
Supplementary Figure 4: Visual response latencies in the attention-switch task. A,B, Unsmoothed (1-ms-binned) population responses aligned on dot onset during the attention shift task. Responses are shown separately for MT (black) and LIP (red) neurons, as well as for the attend-in (solid) and attend-out (dotted line) conditions. The mean baseline activity from the 300 ms before dot onset was subtracted from each spike rate function. **C,D,** Cummulative distribution functions of the single-neuron visual response latencies in each area. See Supplementary Results for details. Inset values are the median latency difference (LIP - MT) and p value from a rank-sum test.

Supplementary Figure 5



Supplementary Figure 5: Examples of biphasic visual responses from single neurons in monkey B MT (**A, B**), monkey B LIP (**C, D**) and monkey M MT (**E**). All responses are shown aligned on dot onset and were smoothed with a 2 ms standard deviation Gaussian. Responses from attend-in (black) and attend-out (gray) trials are shown separately. See Supplementary Results for details.

Supplementary Figure 6



Supplementary Figure 6: Visual response latencies in the direction-mapping task. A,B, Unsmoothed (1-ms-binned) population responses aligned on dot onset during the attention-shift task. Responses are shown separately for MT (black) and LIP (red) neurons, as well as for the attend-in (solid) and attend-out (dotted line) conditions. The mean baseline activity from the 100 ms before dot onset was subtracted from each spike rate function. **C,D,** Cumulative distribution functions of the single-neuron visual response latencies in each area. See Supplementary Results for details. Inset values are the median latency difference (LIP - MT) and p value from a rank-sum test. **E,F,** Population average response from the same neurons as in panels A and B but with each neuron aligned on its visual response latency. The initial responses in MT and LIP populations have a similar slope.