Supplementary Modeling for *Serial dependence in visual perception* by J. Fischer and D. Whitney

Negative aftereffect model

To evaluate whether the serial dependence we report could arise simply as a result of classic negative aftereffects, we ran a simulation that estimated the strength of serial dependence we would expect to find in the presence of negative aftereffects alone. Using the counterbalanced trial sequences presented to subjects in Experiment 1, we estimated the perceived orientation on each trial based on the presented orientation and negative aftereffects from the orientations seen on previous trials. The modeled negative aftereffects peaked in strength for previous orientations that were 10 rotational degrees away from the present stimulus and fell off in strength for smaller or larger separations[51]. We allowed the overall strength of the aftereffects to vary as a free parameter in the simulation, ranging from zero (no aftereffect) to 100%, meaning that a previous stimulus rotated 10° from the present stimulus could repel the perceived orientation of the present stimulus by 10° (100% of the orientation difference). The falloff of the influence of negative aftereffects over time was a second free parameter in the model: aftereffects fell off in a Gaussian fashion with the standard deviation of the Gaussian ranging from zero (instantaneous falloff, hence no negative aftereffects) to one minute (trials seen several minutes ago still had some influence on the perception of the present trial). For a given pair of aftereffect strength and temporal falloff parameters, we computed the perception of the present stimulus as the orientation of the present stimulus (correct response) plus the Gaussian-weighted sum of negative aftereffects from all previous trials. After estimating the perceived orientation for each trial, we constructed a serial dependence plot (see **Supplementary Fig. 1**) and fit a DoG (first derivative of a Gaussian) curve to the data to measure the strength of serial dependence. We recorded the predicted serial dependence at each point in the parameter space to determine whether any combination of parameters could lead to results that resemble the serial dependence we found in subjects' responses. The results are shown in **Supplementary Figure 8**, and demonstrate that serial dependence does not emerge as a product of a negative aftereffect of any strength or duration.

Labeled-line models

How might serial dependence in orientation perception arise? We considered two possibilities: cells tuned to the orientations of recently-viewed stimuli might have increased sensitivity for a period of time following the stimulus presentation (a possible result of lingering feature-selective attention), or orientation tuning of single units might shift away from previously viewed stimuli. For each case we constructed a labeled-line model to examine whether such a phenomenon would result in serial dependence in the decoded orientation from a population of modeled cells. Each model consisted of 180 orientation-tuned Gaussian channels with a standard deviation of 28.2 degrees[52], tiled over orientation space in increments of one degree. Each channel always signaled the orientation for which it was optimally tuned prior to any stimulus presentation (hence a "labeled line" model). That is, in the "shift" model, the range of orientations to which a channel was responsive could shift, but the orientation signaled by a response from that channel remained constant. Perceived orientation was computed as the centroid of the population response: a Gaussian curve was fit to the responses of all channels (allowing for a skewed distribution by estimating sigma for the rightward and leftward tails separately), and the centroid of the distribution was computed as $-\sqrt{\frac{2}{\pi}}\,(\sigma_1 - \sigma_2)$.

In the "gain" model (**Supplementary Fig. 9a**), channels tuned to the orientation of a recently seen stimulus became more sensitive, i.e. they responded more strongly to subsequent stimuli that fell within their tuning range. The amount of gain applied across channels was governed by a Gaussian distribution centered on the orientation of the most recent stimulus, and the amplitude and standard deviation of the Gaussian distribution were free parameters in the model. A third free parameter was a relaxation parameter, which determined how quickly the gain in channels relaxed back to zero after stimulation.

In the "shift" model (**Supplementary Fig. 9b**), the orientation tuning of the channels was shifted away from a recently seen orientation. The distribution of shifts over channels was determined by the first derivative of a Gaussian (DoG) centered at the stimulated location, such that the shift at the stimulated location was zero and neighboring channels were repelled away with an amplitude that peaked at the maxima of the DoG function and fell to zero beyond that. The width and amplitude of the DoG function were free parameters in the model; subjects were allowed to have different maximal shifts at different distances from the stimulated orientation. As in the "gain" model, a third relaxation parameter determined how quickly the shifted channel tuning relaxed back to its default state after stimulation.

Predicted responses for a given set of parameters were generated by using the trial sequences presented to subjects in the fully randomized version of Experiment 1. For each trial, the model generated a prediction for the perceived orientation of the stimulus based on the present and prior stimuli in the trial sequence. We split the data from the randomized version of Experiment 1 into two halves for each subject and conducted model fitting on the first half of the data and model testing on the second half. Model fitting was conducted with least-squares fitting, finding the parameters that minimized the summed squared difference between the model's predicted responses and subjects' actual responses. Within the second half of the data, model performance was evaluated based on the correlation between the model's errors and subjects' errors. Testing performance based on errors rather than raw responses required the model to make the same pattern of errors that subjects made (rather than simply reporting the true stimulus orientations) in order to perform above chance.

Supplementary Figure 9c shows model performance for the gain model (blue data) and the shift model (red data). Chance performance was determined by permuting the trial correspondence between model errors and subjects' errors 5,000 times and recording the correlation on each iteration; p values were taken as the proportion of the permuted null distribution that was larger than the true correlation between model errors and subjects' errors. While the gain model tended to perform somewhat better than the shift model, both models performed significantly above chance for all subjects. Both models also have a plausible basis in known neural phenomena. If feature-selective attention remains tuned to the channels that responded to recent stimuli[30], it could increase sensitivity within those channels as posited in the gain model. There is also evidence that orientation tuning of single cells can shift in the manner specified in the shift model[4], although such shifts resulted from longer adapting stimuli and may or may not arise for the short stimulus presentation in our paradigm. Further tests will be needed to distinguish among these and other models (e.g., a Bayesian estimation framework[31]) that could accommodate our findings, but our labeled-line models demonstrate that serial dependence in orientation perception can arise from simple tweaks to neural tuning based on recent visual input.