Supplemental material

Perception-memory interactions reveal a computational strategy for perceptual constancy

Maria Olkkonen, Toni P. Saarela, and Sarah R. Allred

Comparison between blank-interval and distractor conditions

We employed two versions of the delay conditions (memory and joint). In the blank interval version, the delay between the reference and test was blank, only showing the fixation cross (see Figure 2 of the main text). In the distractor version, two distractor stimuli were displayed for 0.5 s in the middle of each delay period. The distractors had the same spatial dimensions and locations as the reference and test stimuli. Distractor luminances were selected from a Gaussian distribution approximately 1.5 just-noticeable-differences (JND) from a given reference luminance in either direction. There was no task related to the distractors. We included the distractor condition because we hypothesized that the distractors might increase estimation noise and thus exacerbate any potential biases.

The top panels in Figure S1a show the average biases for the memory and joint conditions for the blank interval (left) and distractor (right) versions. The independence predictions are shown with pink thick lines. The biases appear more pronounced in the distractor conditions, but plotting the biases for the two conditions against each other in Figure S1b shows that there are no systematic differences. Moreover, there was no difference in additivity of memory and context effects between the two conditions (1-way repeated measures ANOVA, F(1,50)=0.08, p=0.8). As the additivity, or independence, of memory and context effects is the main focus of this paper, we decided to pool the results over the two delay versions.

The bottom panels of Figure S1a and the left panel of Figure S1b show the thresholds for the two delay conditions. There were no systematic differences across the delay versions.

Figure 3f and g of the main text shows aggregate data pooled across the two delay versions. The other data figures (Figures 4,5,6) show the blank-interval and distractor data as separate data points (not differentiated in the plots for clarity).
Figure S1: Comparison between blank interval and distractor conditions. (a) The top panels show the average memory (blue) and joint (red) biases as a function of reference luminance for the delay conditions with blank interval (left) and distractor interval (right). Thick pink lines show the independence predictions for context and memory. The bottom panels show the average thresholds for the same conditions. Error bars in all panels show ±1SEM. (b) Bias (left) and thresholds (right) are plotted for the distractor conditions against the blank-interval conditions. Best-fitting linear regression lines are shown.
Modeled discrimination thresholds

As we fitted complete psychometric functions to the simulated data, we were able to extract model thresholds in addition to bias. Figure S2 shows a comparison between the observed and modeled discrimination thresholds. Figure S2a plots the modeled and observed thresholds for each of the four conditions. Figure S2b and c plot the errors in the modeled thresholds as a function of the observed threshold.

There were no large differences between the models in the pattern of thresholds. The thresholds were higher in the conditions involving memory, a trend consistent with human data. There were, however, systematic deviations between modeled and observed thresholds. First, in the human data, the thresholds tended to slightly decrease with reference intensity in the baseline condition. The models do not reproduce this effect. This discrepancy leads to a negative slope in the pattern of errors panels b and c: as the observed threshold increases, the error becomes more negative. The two “groups” of points correspond to the conditions with and without memory (conditions involving memory are to the right; the thresholds were higher). Second, the modeled thresholds in the memory conditions were on average too low. In Figure S2b and c, most of the errors in the rightmost group are negative.

There was, however, notable variation between individual observers in thresholds. This was especially true for the memory conditions, where the thresholds were on average higher.
Figure S2: Comparison of discrimination thresholds from the human data and the models. The figure is analogous to Figure 9 in the main text, but it compares thresholds instead of PSE-values. (a) Thresholds for the four conditions in the factorial design are plotted as a function of log reference intensity. Human data are in black symbols and lines, models in thick colored lines. Blue, reflectance model; green and red, contrast models. (b),(c) Analysis of errors in the modeled thresholds. The error (difference between the modeled and observed thresholds) is plotted against the observed thresholds. The histograms to the right show the distribution of errors. Blue, reflectance model; green and red, contrast models. The reflectance model errors are plotted in both (b) and (c) for ease of comparison with both contrast models.
Model parameters

Reflectance model

These are the best-fitting parameters for the reflectance model. The input to the model is log light intensity (relative, unitless scale). The noise parameters are defined on this scale. The priors were modeled as Gaussians, on log-reflectance and log-illuminance scales.

The prior means were not free variables; they were kept fixed. The measurement noise \( \sigma_m \) and time-dependent noise \( \sigma_t \) were free variables. The three prior width parameters were fit with two variables; see main text for description.

Table 1: Reflectance model parameters.

<table>
<thead>
<tr>
<th>Prior</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Illumination</td>
<td>-1.32</td>
<td>0.10</td>
</tr>
<tr>
<td>Reflectance, center</td>
<td>-1.32</td>
<td>0.44</td>
</tr>
<tr>
<td>Reflectance, surround</td>
<td>-0.57</td>
<td>0.20</td>
</tr>
<tr>
<td>( \sigma_m )</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>( \sigma_t )</td>
<td>0.45</td>
<td></td>
</tr>
</tbody>
</table>

Contrast model 1

Table 2 lists the best-fitting parameters in the first contrast model, which had a single prior. The input to the model is log contrast (log luminance ratio at the border of the stimulus patch).

Prior mean was fixed. Prior width, measurement noise \( \sigma_m \) and time-dependent noise \( \sigma_t \) were free variables.

Table 2: Contrast model parameters.

<table>
<thead>
<tr>
<th>Prior</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contrast</td>
<td>0.75</td>
<td>0.38</td>
</tr>
<tr>
<td>( \sigma_m )</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>( \sigma_t )</td>
<td>0.32</td>
<td></td>
</tr>
</tbody>
</table>

Contrast model 2

Table 3 lists the Best-fitting parameters in the second contrast model, with a variable prior. The input to the model is log contrast (log luminance ratio at the border of the stimulus patch).
Prior means were fixed. Measurement noise $\sigma_m$ and time-dependent noise $\sigma_t$ were free variables. Prior widths were fit with a single variable—the width depended on the range of contrasts used in the block. See main text for description.

**Table 3:** Contrast model 2 parameters.

<table>
<thead>
<tr>
<th>Prior</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contrast 1</td>
<td>0.52</td>
<td>0.37</td>
</tr>
<tr>
<td>Contrast 2</td>
<td>0.75</td>
<td>0.53</td>
</tr>
<tr>
<td>Contrast 3</td>
<td>0.98</td>
<td>0.37</td>
</tr>
</tbody>
</table>

$\sigma_m$: 0.08  
$\sigma_t$: 0.30