Supplemental Information:

Changing expectations about speed alters perceived motion direction

Grigorios Sotiropoulos, Aaron R. Seitz, and Peggy Seriès

Supplemental Data

Detailed Results (Figure S1)

Figure S1A indicates that perceptions of motion direction for both groups was accurate at high contrast, and initially biased towards perpendicular judgments at low contrast and short durations. For the low-speed group, this illusion was unaltered after exposure to slow speeds. However, in the high-speed group, exposure to high speeds modulated this initial bias gradually until the illusion reversed and the motion direction was most often perceived as being more oblique. The rate of learning in the high-speed group was very close to linear. To quantify these observations, linear models (both separately for pre- and post-training blocks and with these blocks combined) were fit to the 133 and 266 ms data (separately and combined) for each group. The fitted models were then evaluated for significance via linear hypothesis tests (linhyptest, MATLAB).

All linear fits for the high-speed group were significant at the 5% level, with the exception of the pre-training fit for the 133 ms condition, which only approaches significance. The most significant fits are observed for the 266 ms condition. Also, models that combine pre- and post-training blocks are more significant. In particular, the p values for the 133 ms condition were 0.0503, 0.0339, 0.0049 for the first, the second, and both test blocks within the session, respectively. The respective values for the 266 ms condition are 0.012, 0.0284, 0.0011 and for the combined 133 and 266 ms conditions, the values are 0.0242, 0.0437, 0.0037. In the low-speed group, $p > 0.398$ for all conditions.

Fig S1B indicates that the high-speed group continues to be biased towards perpendicular judgments during the training block, when it is exposed to high speeds, with a tendency for the bias to decrease with time. However, this tendency is not significant: a one-way ANOVA does show a mild but significant difference in p_o across sessions for the training block in the test group (p = 0.0111) but there is no consistent upward tendency as the one observed in the test blocks (p = 0.28, 0.2 and 0.09 for the 133, 266 and 532 ms conditions, respectively, linear hypothesis test). These results provide evidence against the observed changes found in the testing blocks being due to a response bias that develops through the training sessions.

Fig S1C and D show that the changes in the perception of direction of motion affects both oblique and horizontal motion conditions: with exposure to high speeds, subjects show less and less bias towards perceiving oblique motion as being perpendicular (Fig S1D), but also develop a new bias towards perceiving perpendicular motion as being oblique (Fig S1C).

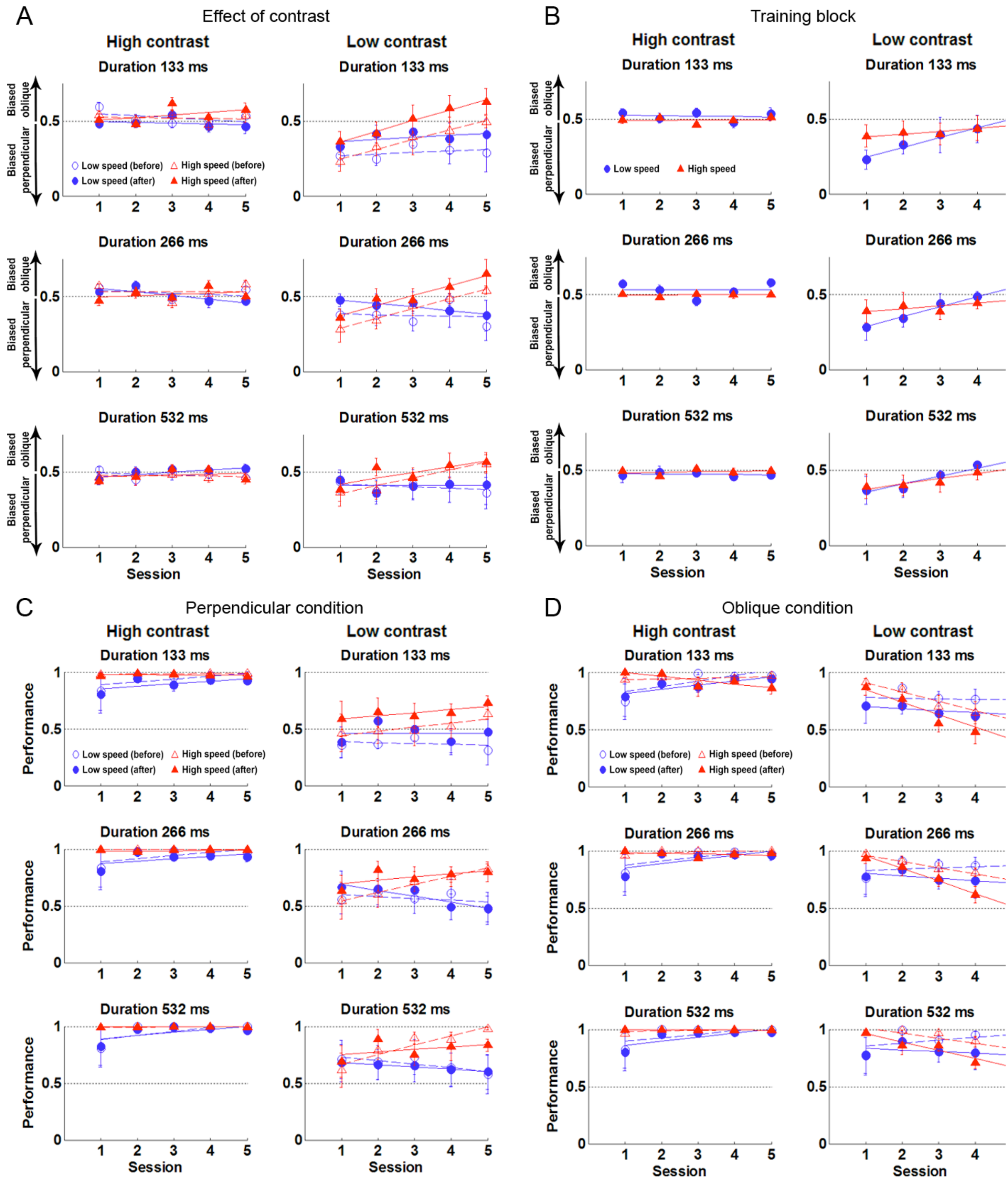

Figure S1: (A) Proportion of oblique responses plotted against session number for the low-speed and high-speed groups and for all durations and contrast levels. Dashed lines correspond to data from the first test block in each session (before training) whereas full lines correspond to the third block (after training). A three-way ANOVA on proportion of oblique responses with factors session, duration and block showed that for the low-speed group, neither session number nor duration had a significant effect *(p* > 0.85 and $p > 0.29$, respectively), whereas the effect of block (before vs after training) was borderline significant at the 5% level *(p=0.046)*. For the high-speed group, duration also had no effect *(p=0.52);* however, the effect of session number was highly significant ($p<0.001$), as was the effect of block *(p=0.0047)*. There were no higher-level interactions. (B) Proportion of oblique responses in the training (middle) block for each group. (C) and (D) Performance in trials with normal and oblique (respectively) veridical motion. Error bars are ±1 between-subjects SEM.

Bayesian Model

To investigate quantitatively whether our results are consistent with the idea of a changing speed prior, we adapted the Bayesian model of motion perception proposed by Weiss et al (2002). This model suggests that motion perception can be described as an optimal estimation of object velocities under the assumption of local measurement noise and an *a priori* preference for slower velocities (in Weiss et al (2002), the prior is centered at zero). The idea of our extension of this model is that the speed prior is initially centered close to zero but varies further away from zero from session to session due to exposure. As described below, the model and parameters of this learning were determined to best fit the group-averaged data.

When receiving as inputs the moving stimulus image, the model produces a velocity estimate, which is given by a variant of the solution of Weiss et al (2002), generalized to include a velocity prior with a non-zero mean:

$$\hat{v} = -\begin{pmatrix} \sum I_x^2 + \frac{\sigma^2}{\sigma_p^2} & \sum I_x I_y \\ \sum I_x I_y & \sum I_x^2 + \frac{\sigma^2}{\sigma_p^2} \end{pmatrix}^{-1} \begin{pmatrix} \sum I_x I_t - \frac{\sigma^2}{\sigma_p^2}\mu_x \\ \sum I_y I_t - \frac{\sigma^2}{\sigma_p^2}\mu_y \end{pmatrix}$$

where I_x, I_y, I_t are the spatial (two dimensions) and temporal (partial) derivatives of the image intensity function, σ^2/σ_p^2 is the ratio of the likelihood and prior variances, the only free parameter in the original model of Weiss et al (2002), and μ_x, μ_y are the means (two dimensions) of the velocity prior. The sums were computed over the pixels in the stimulus images. The spatial derivatives were computed using MATLAB's `gradient()` function, which performs a simple subtraction of the values of neighboring pixels along each axis. The temporal derivative is given by:

$$I_t = -(I_x v_x + I_y v_y)$$

where v_x, v_y are the local velocity measurements in the two axes (and equal to the veridical velocity of the stimulus and to each other, since the stimulus translates rigidly).

Following Weiss et al (2002), we simulated our 2-alternative forced choice experiment by assuming that the decision ('up' or 'down') is corrupted by Gaussian noise. The model response in the presence of this 'decision noise' is given by:

$$r = \text{sign}(\phi_{est} + \eta)$$

where ϕ_{est} is the angle between the estimated velocity vector and the horizontal and η a zero-mean Gaussian random variable of standard deviation σ_D. The direction estimate $\phi_{est}>0°$ corresponds to right-upward motion and $\phi_{est} < 0°$ to right-downward motion.

Updating the Prior.

The mean of the velocity prior $\boldsymbol{\mu}_p = (\mu_x, \mu_y)$ was allowed to vary within a session and between session. We modeled this learning using the following assumptions:

(i) Based on observed performance differences between successive sessions, we assumed that only a proportion (1-γ) of the shift of the prior within a session would be retained in the next session. We call γ the 'unlearning rate' parameter.

(ii) To describe the trajectory of the prior, two variants of the model were investigated:

• In the first (non-parametric) model, the prior mean is determined independently for each session, by fitting to the data. In total, this model has 10 parameters: 5 for the prior mean locations after training (one for each session), 1 for the 'unlearning rate' γ, 1 for the "decision noise" standard deviation σ_D and 3 for the ratio σ^2/σ_p^2 for each duration.

• In the second (parametric) model, the post-training prior mean was assumed to vary linearly with time (or session). This model was chosen after observing a roughly linear relation between prior means and session number using the non-parametric model above. The prior mean was modeled as: $\mu_x(s) = as + b$, where s denotes the session number. This reduces the number of the prior-related parameters from 5 to 2.

(iii) Only the horizontal component μ_x, of $\boldsymbol{\mu}_p$ was allowed to vary; μ_y was fixed at zero. Given that in our model v_y is a Gaussian, a nonzero value of μ_y would correspond to an artificially imposed vertical directional bias that would mask the "up/down" bias that we wish to investigate, which is a result of a prior favoring high speeds in general (without assuming a preference for either vertical direction). A more accurate alternative would be to use a bimodal distribution for v_y. Such a prior could potentially track exactly the stimulus distribution (which is defined by 2 points during training for the high-speed group: $\{\mu_{x=} 8\cos(\theta_1); \mu_{y=}8 \sin(\theta_1)\}$; $\{\mu_{x=} 8\cos(\theta_2); \mu_{y=}8 \sin(\theta_2)\}$, where θ_1 = -20 deg and θ_2 = 20 deg represents the oblique and perpendicular conditions). However, the closed-form solution of Eq. 1 would no longer apply, as it is based on assumptions of Gaussianity.

(iv) All aforementioned parameters were fit with MATLAB's `fminsearch` function using the low-contrast data group averaged data[1], using Maximum Likelihood estimation.

The results shown in Figure 1D of the main text correspond to the non-parametric version of the model.

Figure S2 shows the means of the speed prior (i.e. the magnitudes of the velocity means) fitted to the group-averaged data both parametrically and non-parametrically. It is observed that in the high-speed group the prior shifts towards speeds approaching the testing speed with training -- reaching a value of 6.2 deg/s at the end of the last session -- whereas the prior for the low-speed group remains almost fixed. The fitted σ^2/σ_p^2 values decrease monotonically with duration. This reflects the fact that with longer durations, as the visual evidence is becoming more reliable, the likelihood becomes sharper.

The other best-fitting parameters were σ_D =5 deg (both groups) for the motor noise and γ=0.17 and 0.99 (high-speed and low-speed group, respectively) for the unlearning rate. This large difference in γ across the groups reflects the fact that, for the low speed group, the within-session learning seems to be forgotten from one day to another. One explanation for this might be that the amount of learning should exceed a threshold in the short-term for it to be consolidated in the long-term.

The root mean squared error (RMSE) of the non-parametric model fit for the high-speed and low-speed group was 0.026 and 0.0403, respectively. The respective values for the parametric variant are 0.0304 and 0.0486. As expected, the non-parametric model fits the data better at the cost of greater complexity (number of parameters). To compare the two models, we used the 'corrected Akaike information criterion' [2], defined as

$$\mathrm{AICc} = 2kn/(n-k-1) - 2L$$

where L is the likelihood of the data given the model, k the number of model parameters and n the number of data points. The model favored by AICc (the one with the lowest score[2]) was the parametric one in the case of the high-speed group (8709.3 parametric vs 8717.5 non-parametric) but

[1] Fitting including the high-contrast data too gave slightly worse results, presumably due to fitting more noise: probabilities at high contrasts are theoretically all close to 0.5 and thus are not informative.

[2] The reported values are up to an additive constant that is common to all models (only depends on the data) and thus ignored.

the non-parametric one for the low-speed group (8603.8 parametric vs 8601.6 non-parametric).

In theory, not only the mean of the prior could change with exposure, but also its shape. Variants of the model where the prior variance (σ_p^2) was also allowed to vary across sessions were also examined. In the first variant, σ_p^2 is a linear function of session number (requiring 3 additional parameters); in the second variant, σ_p^2 is allowed to vary freely in each session (requiring 6 additional parameters). These models predict that the prior mean increases in the experimental group, qualitatively similar to the previous model, while the variance of the prior remains mostly stable (data not shown). The RMSEs of the best-fitting linear-variance model for the high-speed and low-speed group were 0.026 and 0.0388, respectively. The respective RMSEs for the free-variance model were 0.0249 and 0.0384. Both variants fit the data slightly better than the fixed-variance model, however the AICc of the former are greater. In particular, the AICc for the fixed-variance model were 8717.5 and 8599.2 (high-speed and low-speed group, respectively); the respective AICc for the linear-variance model were 8730.9 and 8605.2; and the AICc for the free-variance model were 8758.8 and 8637.2. This suggests that the model which varying prior mean and fixed variance is a better description of the data. However, it was found that these models were difficult to fit as the optimization process was prone to local minima, due to the increased number of parameters compared to data points. Future work will aim at a more detailed characterization of the prior variance with exposure.

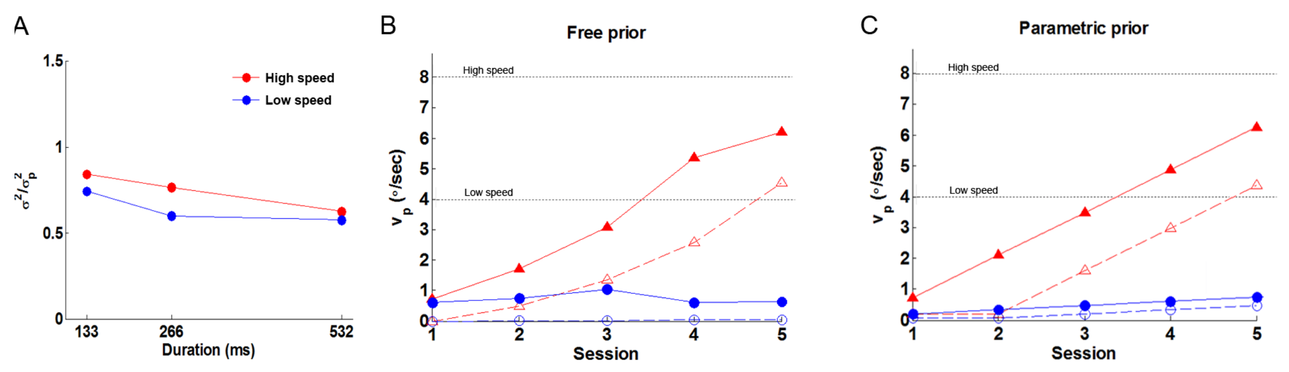

Figure S2: (A) Ratio of the variances of the likelihood (visual evidence) and the prior (expectations) as a function of trial duration. With longer durations, visual evidence becomes more reliable and thus the likelihood is sharper and σ^2/σ_p^2 decreases. (B, C) Evolution of the speed prior (magnitude of the velocity prior) through training: mean of the speed prior as a function of session number for the non-parametric and parametric (linear) prior model, respectively. For comparison, the dotted line shows the actual stimulus speed during testing.

Supplemental Experimental Procedures

The stimulus consists of a field of parallel lines, translating rigidly and coherently (see Figure 1A in the main text, and movie of stimulus - here shown for 0.5 seconds and at a contrast that is much higher than in the actual experiment). Line elements had a length of 4° of visual angle and an orientation of 110° i.e. approximately 11 o'clock. Motion direction was either *perpendicular* to the line, which means 20° from horizontal (upward motion), or *oblique,* that is, -20° from horizontal (downward motion). The equal distance of the two directions from the horizontal ensured that any bias towards the horizontal [1] would have the same effect in both conditions ('up' and 'down'). In each trial, the direction was picked randomly without replacement. Stimulus presentation durations were uniform randomly chosen from the set {133, 266, 532 ms}. Stimulus speed was either 4 deg/sec (test session and training session of low-speed group, see below) or 8 deg/sec (training session of high-speed group). With the exception of line length, all above parameters were the same as in the first experiment in Lorenceau et al (1993).

The field of parallel lines was visible through a circular mask of 24° (of visual angle) in diameter, as in previous work (Lorenceau et al, 1993). The circular mask is much larger than individual lines, so most of the lines (excluding those near the circumference of the aperture) are visible in their entirety (ie including their endpoints). The background was black (luminance 0.29 cd/m^2) whereas the lines were gray. Two gray levels were used for the lines: they were either shown with a luminance of 0.342 cd/m^2 corresponding to a Michelson contrast of 8% ("low" contrast) or with a luminance of 0.948 cd/m^2 corresponding to a Michelson contrast of 53% ("high" contrast). Lines were 2.4' (minutes of arc) thick.

Twelve naive subjects participated in this experiment, evenly divided in two groups, test and low-speed. Each subject conducted 5 high-speed sessions, each on a separate day with each session at the same time of day for a given subject. Each session lasted 35-45 minutes and was divided into 3 blocks - a small test block, followed by a large training block, followed by a small test block identical to the first. Stimulus contrasts, durations and directions were randomly interleaved in each block. Motion speed was kept constant and equal to 4 deg/sec in the test blocks and in the training block of the low-speed group, while it was fixed at 8 deg/sec in the training block of the high-speed group.

The number of trials in each test block was chosen so that each condition was presented 18 times. Since there were 2 (equiprobable) contrast levels, 3 (equiprobable) durations and 2 directions, the total number of trials was 2 x 3 x 2 x 18 = 216. The number of trials in the training block was chosen to be larger than both test blocks. In particular, each condition was presented 60 times, so that each training block consisted of 2 x 3 x 2 x 60 = 720 trials. In total, the number of trials in a session was 1152.

The structure of an individual trial was as follows:

- A 200 ms fixation period, during which the static stimulus (line field) was displayed, together with a red central fixation dot (diameter = 24 arcmin).
- A stimulus presentation period. The fixation dot disappeared at the beginning of this period and motion of the line field was initiated. All lines moved in the same direction and at the same speed, so the apparent motion is a global translation of the entire field. The duration of this period was either 133, 266 or 532 ms.
- A response period. The stimulus disappeared and the fixation dot reappeared. Subjects reported the perceived direction of motion by pressing the 'Up' /'Down' arrow key on the keyboard. When a response was made or when 3 seconds had elapsed, the trial ended.

Data analysis

Following Lorenceau et al. (1993), the proportion of oblique responses (hereafter referred to as p_o) was measured for each duration and contrast level. Performance (proportion of correct responses in the task) before and after training was also calculated separately for the upward- and downward-motion trials. ANOVA was used to determine whether there was a change in p_o across sessions and training blocks and whether there was a difference between test and low-speed groups. Significant results from ANOVA were followed up by Tukey's HSD tests in order to examine these differences.

Supplemental References

[1] G. Loffler and H.S. Orbach (2001). Anisotropy in judging the absolute direction of motion. *Vision Research*. 41(27), 3677-3692.

[2] Hurvich, C. M., and Tsai, C.-L. (1989). Regression and time series model selection in small samples. *Biometrika*, 76: 297–307.