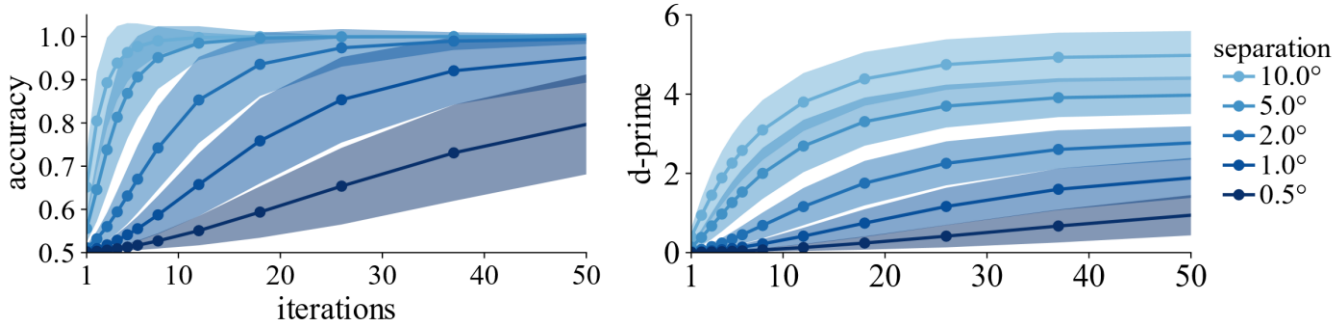

Figure 2-1. Learning performance trajectory for first 50 iterations, same data as Figure 2 but plotted as mean±std of accuracy (left) and d-prime (right). Performance increased approximately linearly at the beginning and more slowly for harder tasks.

Statistical test on training and test precision on transfer performance

transfer conditions		ori + 45	ori + 90	SF × 0.5	SF × 2.0
train and test at the same angle separations	effect of log angle separation	**β=0.15 t(2878)=27.58 p=8.9×10^{-149} R^2=0.21**	**β=0.184 t(2878)=34.57 p=2.2×10^{-219} R^2=0.29**	**β=0.21 t(2878)=59.29 p≈0.0 R^2=0.55**	**β=0.17 t(2878)=42.78 p=5.7×10^{-310} R^2=0.39**
train and test at all combinations of angle separations	effect of log training angle separation	**β=-0.086 t(14397)=-36.09 p≈3.9×10^{-273} R^2=0.052**	**β=-0.050 t(14397)=-20.94 p≈6.8×10^{-96} R^2=0.018**	**β=-0.12 t(14397)=-78.93 p≈0.0 R^2=0.11**	**β=-0.12 t(14397)=-71.41 p≈0.0 R^2=0.11**
	effect of log testing angle separation	**β=0.23 t(14397)=96.32 p≈0.0 R^2=0.37**	**β=0.23 t(14397)=77.50 p≈0.0 R^2=0.40**	**β=0.29 t(14397)=191.92 p≈0.0 R^2=0.64**	**β=0.27 t(14397)=160.26 p≈0.0 R=0.57**

Table 2-1. Linear regression test details of training and test (log) angle separations on transfer performance. When transfer was tested on the same angle separation as training, the transfer performance increased as angle separation increased. When the network was trained and tested on all combinations of angle separations, significant effects of both predictors were seen in all transfer conditions, but the effect size (R^2) of training angle separation in orthogonal orientation transfer (ori+90) was smaller than in other conditions. Some p-vales are less than machine precision and are indicated as p≈0.0.

Statistical test details for effect of layer number on PSI

separation	0.5°	1.0°	2.0°	5.0°	10.0°
effect of layer number (1-5)	**β=-4.73 t(2878)=-8.44 p=4.95×10^{-17} R^2=0.024**	**β=-0.52 t(2878)=-4.27 p=2.0×10^{-5} R^2=0.0063**	β=-0.094 t(2878)=-1.42 p=0.16 R^2=0.00070	β=-0.014 t(2878)=-0.32 p=0.75 R^2=3.6×10^{-5}	β=-0.013 t(2878)=-0.46 p=0.65 R^2=7.3×10^{-5}

Table 3-1. Linear regression test details for effect of layer number (excluding layer 6) on the iteration at which a layer reached its maximum speed of change (PSI). Significant effects of layer number were found under the two most precise conditions (bold), implying that higher layers started to asymptote before early layers.

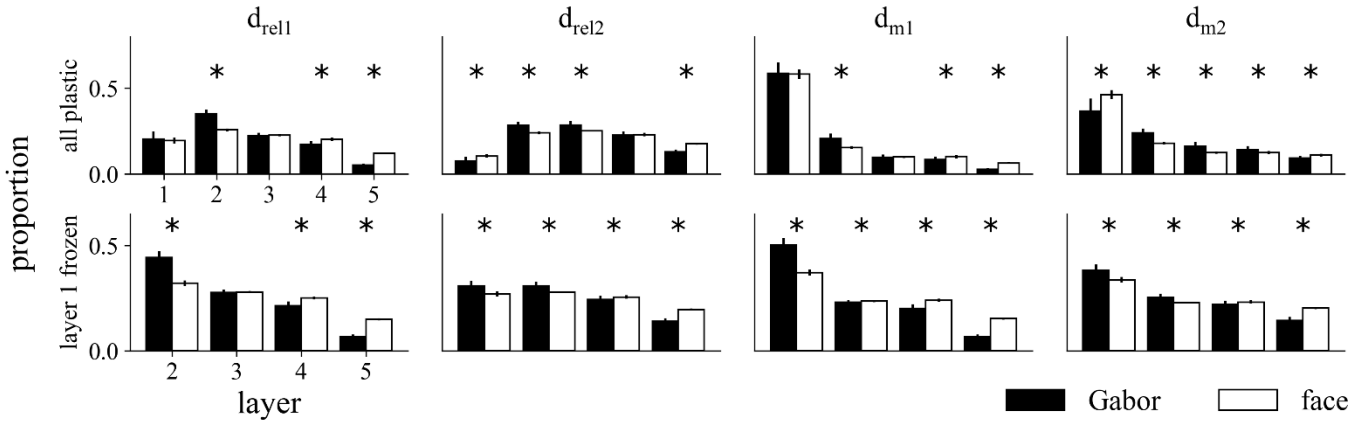

Figure 5-1. Weights change proportions at each layer when all layers were plastic (top row) and when all but the first layers were plastic (second row) using four different layer change measures (columns). The first normalization operation in the network was after the first layer, so freezing layer 1 avoided learning directly on the pixel values. The four measures are defined in Equations 4, 4a, 4b and 4c, respectively. These changes were then normalized to sum to 1 to obtain the proportions. The interaction of layer × task on layer change proportion was significant in both conditions using any of the four measures ($p<0.0001$, see Table 5-1 for details). Learning the face task produced more change in the higher layers when the first layer was frozen. (*) indicates significant difference in layer change proportion (threshold $p=0.01$, Mann-Whitney U, Bonferroni-corrected for 4 or 5 plastic layers).

Statistical test details for the interaction of layer × task on layer change proportion

Freeze layer 1?	d_{rel1}	d_{rel2}	d_{m1}	d_{m2}
No	**$F(4,590)=263.65$** **$p=8.7\times10^{-130}$** **$R^2=0.12$**	**$F(4,590)=198.09$** **$p=1.4\times10^{-107}$** **$R^2=0.058$**	**$F(4,590)=48.89$** **$p=1.6\times10^{-35}$** **$R^2=0.0055$**	**$F(4,590)=135.89$** **$p=3.1\times10^{-82}$** **$R^2=0.052$**
Yes	**$F(3,472)=908.80$** **$p=1.2\times10^{-195}$** **$R^2=0.13$**	**$F(3,472)=240.37$** **$p=1.2\times10^{-94}$** **$R^2=0.11$**	**$F(3,472)=990.63$** **$p=3.3\times10^{-203}$** **$R^2=0.1$**	**$F(3,472)=252.61$** **$p=9.6\times10^{-98}$** **$R^2=0.074$**

Table 5-1. ANOVA test details for the interaction of layer × task on layer change proportion. The individual main effects were included in the analyses. All interactions were significant.

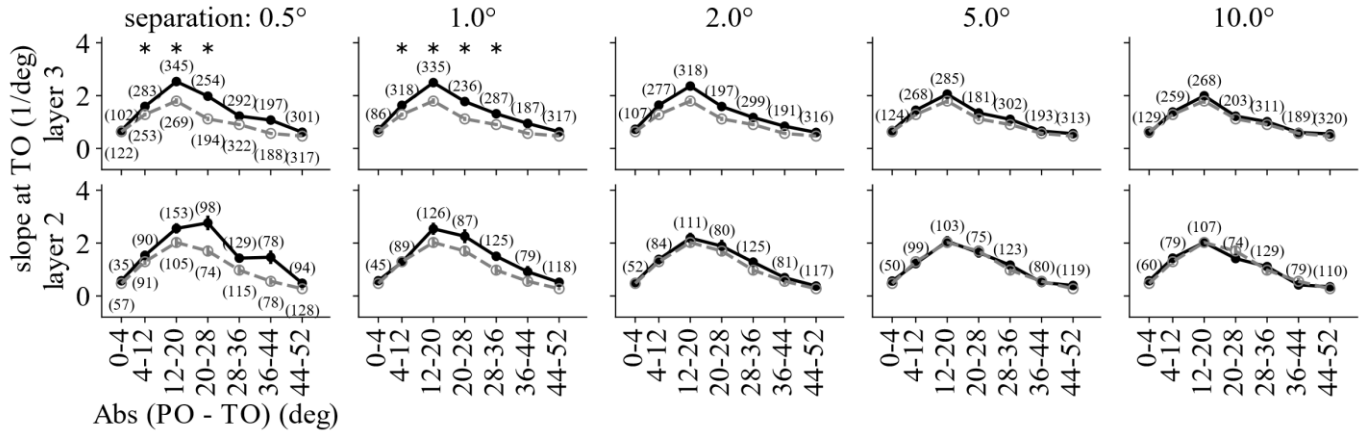

Figure 6-1. Slope of (max-normalized) tuning curve at the trained orientation (TO) grouped by preferred orientation (PO) relative for the trained (black solid) and naïve (gray dashed) model unit populations in the two lower layers. Two rows correspond to layers 3 and 2; columns are different training angle separations. There was a significant interaction of orientation × training on the slope at TO for units in layer 3 when trained at angle separations of at 0.5 and 1.0° ($p<0.01$) but not for larger angle separations ($p>0.07$, see Table 6-1 for details). Layer 2 failed to reach significance for all angle separations ($p>0.1$). (*) indicate significant increase in slope at TO (threshold $p=0.01$, Mann-Whitney U, Bonferroni-corrected for 7 bins).

Statistical test details for the interaction of orientation × training on slope at TO

separation		**0.5°**	**1.0°**	**2.0°**	**5.0°**	**10.0°**
layer	**3**	**$F(6,3425)=5.17$ $p=2.7\times10^{5}$**	**$F(6,3417)=2.99$ $p=0.0065$**	$F(6,3356)=1.90$ $p=0.077$	$F(6,3317)=0.47$ $p=0.83$	$F(6,3330)=0.39$ $p=0.89$
	2	$F(6,1311)=1.71$ $p=0.11$	$F(6,1303)=0.82$ $p=0.55$	$F(6,1284)=0.15$ $p=0.99$	$F(6,1283)=0.24$ $p=0.96$	$F(6,1272)=0.51$ $p=0.8$

Table 6-1. ANOVA test details for the interaction of orientation × training on slope at TO. The individual effects were included in the analyses. There was a significant interaction of orientation × training on the slope at TO for units in layer 3 when trained at angle separations of at 0.5 and 1.0° but not for larger angle separations.

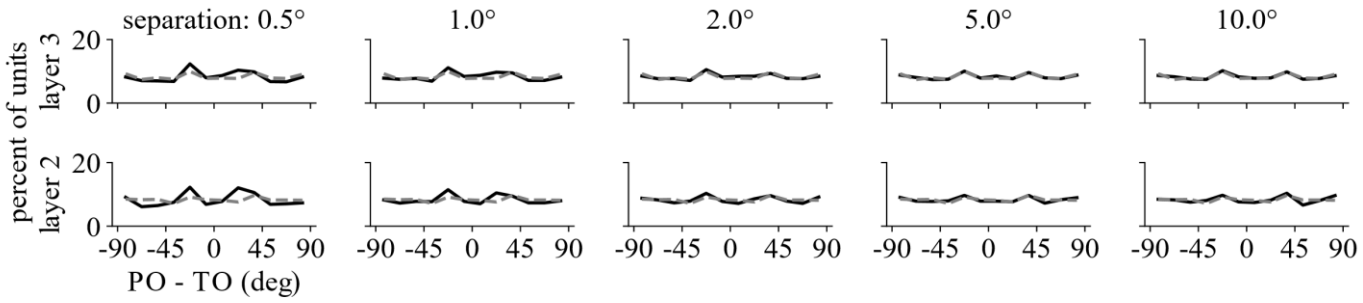

Figure 6-2. Distributions of preferred orientation (PO) relative to trained orientation (TO) for the trained (black solid) and naïve (gray dashed) model unit population in the two lower layers. Two rows correspond to layers 3 and 2; columns are different training angle separations. No significant changes in distribution were found ($p>0.05$, see Table 6-2 for details).

Statistical test details for change in distribution of preferred orientation

separation		**0.5°**	**1.0°**	**2.0°**	**5.0°**	**10.0°**
layer	**3** (n_{naive}=2848)	d=0.0353 p=0.059 n=2787	d=0.0278 p=0.22 n=2839	d=0.0144 p=0.93 n=2845	d=0.00805 p=1.0 n=2845	d=0.00888 p=1.0 n=2843
	2 (n_{naive}=1092)	d=0.0383 p=0.4 n=1078	d=0.0231 p=0.93 n=1089	d=0.0126 p=1.0 n=1103	d=0.0113 p=1.0 n=1105	d=0.0210 p=0.97 n=1100

Table 6-2. K-S test details for the interaction of orientation × training on slope at TO. No significant changes in distribution were found ($p>0.05$ for all layers and angle separations).

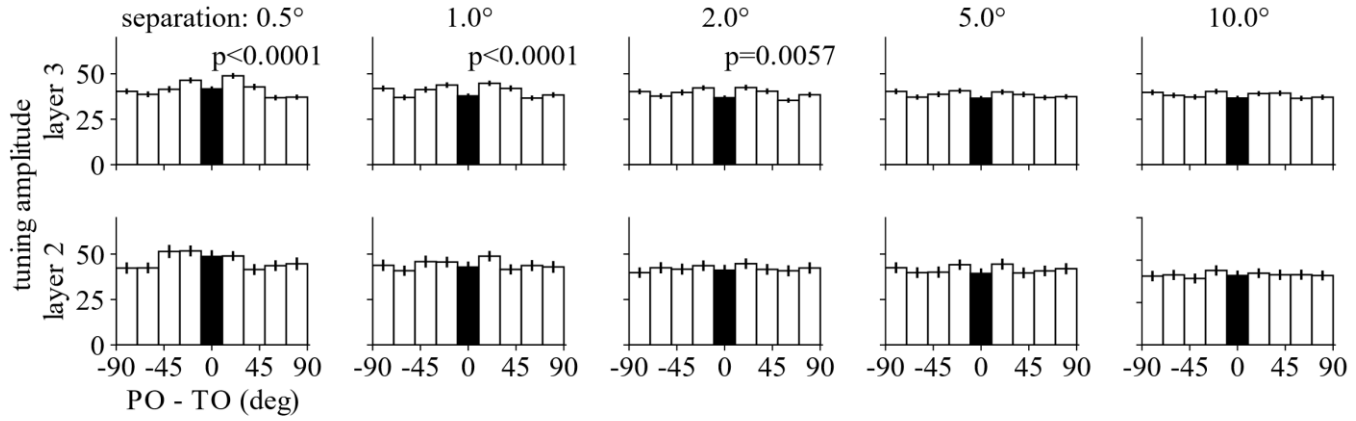

Figure 7-1. Tuning amplitude grouped by preferred orientation (PO) relative to trained orientation (TO) for the trained model unit population in the two lower layers. Two rows correspond to layers 3 and 2; columns are different training angle separations. Black bar indicates the orientation bin that contains the trained orientation. Preferred orientation had significant effects on tuning amplitude only in layer 3 when trained up to 2.0° ($p<0.01$, see Table 7-1 for details).

Statistical test details for effect of orientation on tuning amplitude

separation		**0.5°**	**1.0°**	**2.0°**	**5.0°**	**10.0°**
layer	**3**	**$F(9,2777)=6.94$ $p=4.4\times10^{-9}$**	**$F(9,2829)=4.00$ $p=0.00010$**	**$F(9,2835)=2.71$ $p=0.0057$**	$F(9,2835)=1.19$ $p=0.30$	$F(9,2833)=1.00$ $p=0.44$
	2	$F(9,1068)=1.70$ $p=0.095$	$F(9,1079)=0.72$ $p=0.67$	$F(9,1093)=0.26$ $p=0.98$	$F(9,1095)=0.44$ $p=0.90$	$F(9,1090)=0.19$ $p=0.99$

Table 7-1. ANOVA test details for effect of orientation on tuning amplitude. Preferred orientation had significant effects on tuning amplitude only in layer 3 when trained up to 2.0°.

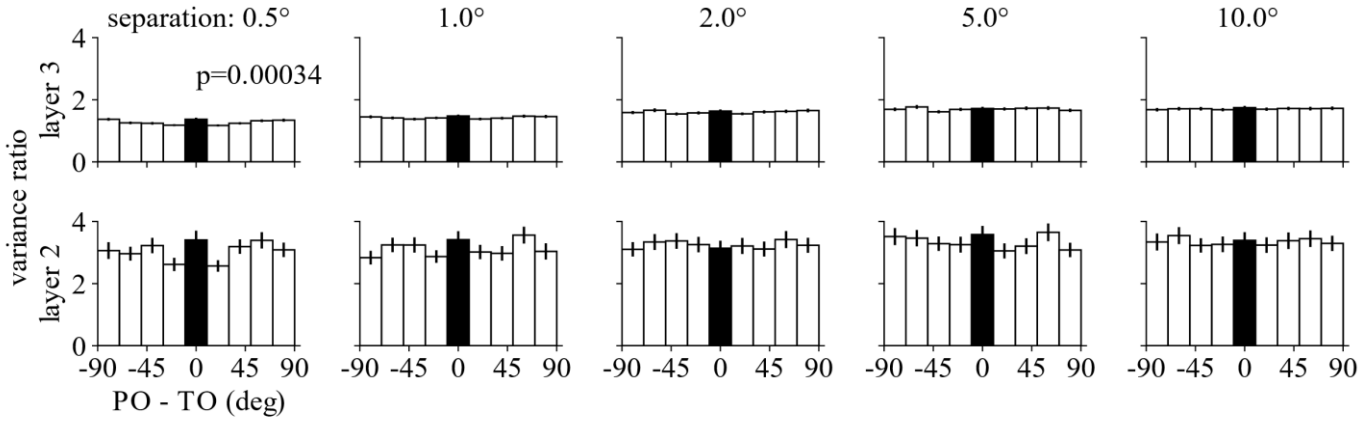

Figure 7-2. Variance ratio grouped by preferred orientation (PO) relative to trained orientation (TO) for the trained model unit population in the lower two layers. Two rows correspond to layers 3 and 2; columns are different training angle separations. Black bar indicates the bin that contains the trained orientation. PO had a significant effect on variance ratio only in layer 3 when trained at 0.5° ($p=0.0028$, see Table 7-1 for details).

Statistical test details for the effect of PO on variance ratio

separation		**0.5°**	**1.0°**	**2.0°**	**5.0°**	**10.0°**
layer	**3**	**$F(9,2777)=3.62$ $p=0.00034$**	$F(9,2829)=0.67$ $p=0.72$	$F(9,2835)=0.68$ $p=0.71$	$F(9,2835)=0.61$ $p=0.77$	$F(9,2833)=0.14$ $p=1.0$
	2	$F(9,1068)=1.85$ $p=0.064$	$F(9,1079)=1.00$ $p=0.44$	$F(9,1093)=0.20$ $p=0.99$	$F(9,1095)=0.77$ $p=0.69$	$F(9,1090)=0.16$ $p=1.0$

Table 7-2. ANOVA test details for the effect of preferred orientation (PO) on variance ratio. PO had a significant effect on variance ratio only in layer 3 when trained at 0.5°

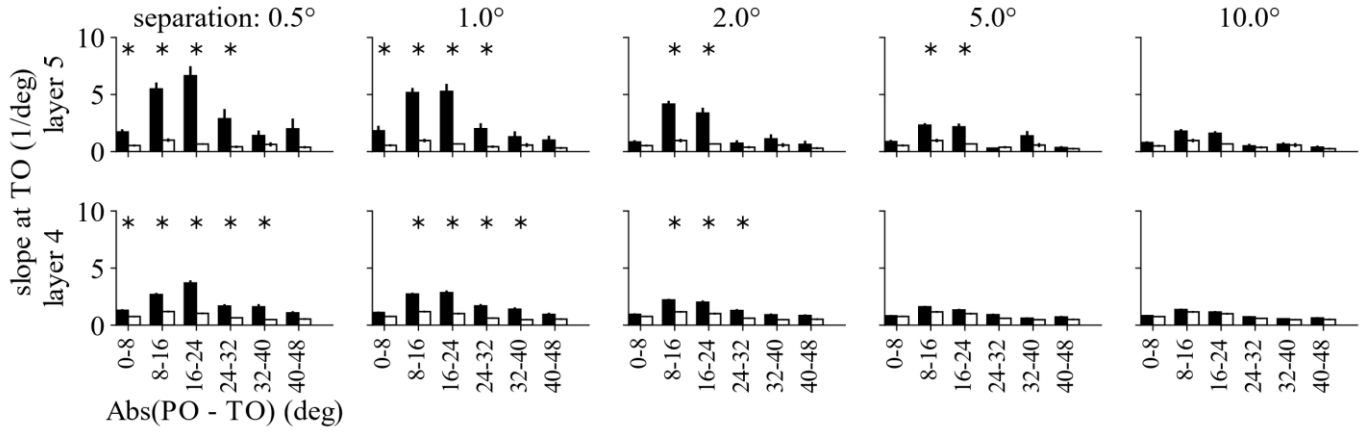

Figure 8-1. Slope at trained orientation (TO) grouped by preferred orientation (PO) for the trained (black) and naïve (white) model unit populations in the two higher layers. Two rows correspond to layers 5 and 4; columns are different training angle separations. Significant interactions of PO × training on the slope at TO were found for angle separations up to 5.0° in layer 5 and up to 2.0° in layer 4 ($p<0.001$, see Table 8-1 for details). (*) indicates significant increase in slope at TO (threshold $p=0.01$, Mann-Whitney U, Bonferroni-corrected for 6 bins).

Statistical test details for interaction of PO × training on the slope at TO

separation		**0.5°**	**1.0°**	**2.0°**	**5.0°**	**10.0°**
layer	**5**	**$F(5,196)=3.68$ $p=0.0033$**	**$F(5,202)=5.33$ $p=0.00013$**	**$F(5,203)=8.26$ $p=4\times10^{-7}$**	**$F(5,198)=5.61$ $p=7.4\times10^{-5}$**	$F(5,199)=2.79$ $p=0.018$
	4	**$F(5,1244)=12.69$ $p=4.8\times10^{-12}$**	**$F(5,1234)=9.48$ $p=7\times10^{-9}$**	**$F(5,1220)=5.78$ $p=2.8\times10^{-5}$**	$F(5,1215)=1.51$ $p=0.18$	$F(5,1224)=0.21$ $p=0.96$

Table 8-1. ANOVA test details for interaction of preferred orientation (PO) × training on the slope at trained orientation (TO). The individual main effects were included in the analyses. Significant interactions of PO × training on the slope at TO were found for angle separations up to 5.0° in layer 5 and up to 2.0° in layer 4.

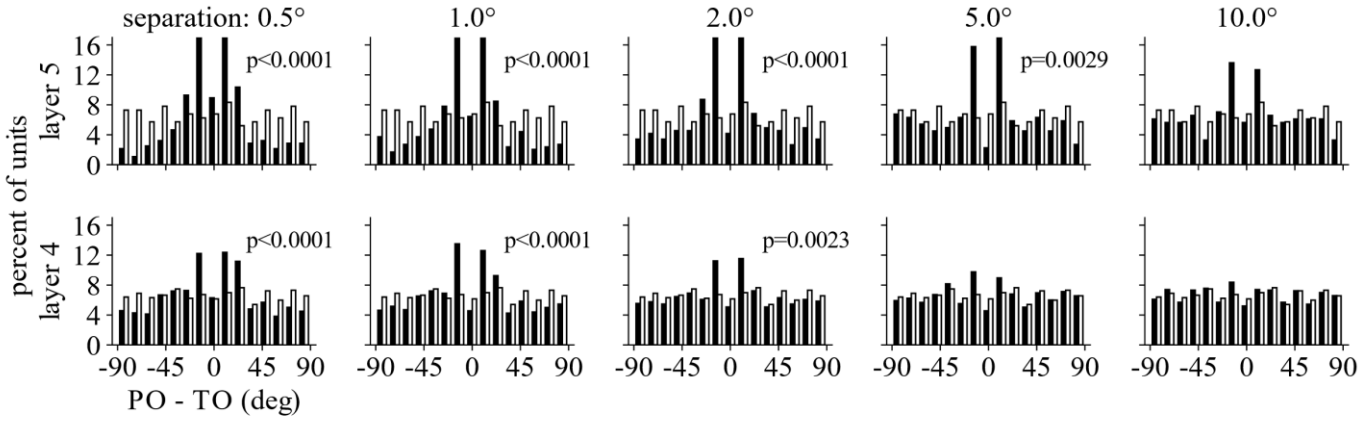

Figure 8-2. Distribution of preferred orientation (PO) relative to trained orientation (TO) for the trained (black) and naïve (white) model unit populations in the two higher layers. Two rows correspond to layers 5 and 4; columns are different training angle separations. Significant deviations in distribution from uniform were found after training for angle separations up to 5.0° in layer 5 and up to 2.0° in layer 4 ($p<0.003$, see Table 8-2).

Statistical test details for distribution of preferred orientation compared against a uniform distribution

separation		**0.5°**	**1.0°**	**2.0°**	**5.0°**	**10.0°**
layer	**5**	**d=0.21** **$p=1.3\times10^{-11}$** **n=280**	**d=0.21** **$p=5.2\times10^{-12}$** **n=295**	**d=0.15** **$p=1.7\times10^{-5}$** **n=264**	**d=0.12** **p=0.0029** **n=222**	d=0.083 p=0.10 n=213
	4	**d=0.095** **$p=6.4\times10^{-11}$** **n=1334**	**d=0.085** **$p=8\times10^{-9}$** **n=1318**	**d=0.051** **p=0.0023** **n=1301**	d=0.029 p=0.24 n=1250	d=0.017 p=0.88 n=1230

Table 8-2. K-S test details for distribution of preferred orientation (PO) compared against a uniform distribution. Significant deviations in distribution from uniform were found after training for angle separations up to 5.0° in layer 5 and up to 2.0° in layer 4.

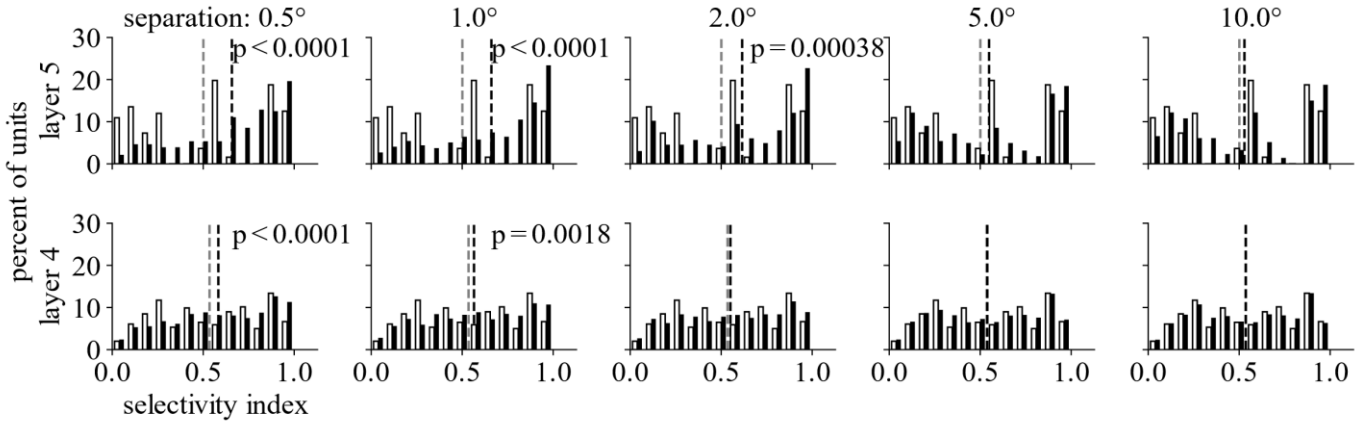

Figure 8-3. Distribution of selective index for the trained (black) and naïve (white) model unit populations in the two higher layers. Two rows correspond to layers 5 and 4; columns are different training angle separations. Black and gray dashed lines indicate the trained and naïve population means, respectively. Significant increases in selective index were found for angle separations up to 2.0° in layer 5 and up to 1.0° in layer 4 ($p<0.002$, See Table 8-3 for details).

Statistical test details for increase in selectivity index

separation		**0.5°**	**1.0°**	**2.0°**	**5.0°**	**10.0°**
layer	**5** ($MD_{naive}=0.55$, $n_{naive}=192$)	**MD=0.72** **$U=2.0\times10^4$** **$p=1.7\times10^{-6}$** **n=280**	**MD=0.75** **$U=2.1\times10^4$** **$p=1.1\times10^{-6}$** **n=295**	**MD=0.66** **$U=2.1\times10^4$** **p=0.00038** **n=264**	MD=0.58 $U=2\times10^4$ p=0.1.0 n=222	MD=0.56 $U=1.9\times10^4$ p=0.21 n=213
	4 ($MD_{naive}=0.54$, $n_{naive}=1203$)	**MD=0.60** **$U=7.2\times10^5$** **$p=2.0\times10^{-6}$** **n=1334**	**MD=0.57** **$U=7.4\times10^5$** **p=0.0018** **n=1318**	MD=0.56 $U=7.6\times10^5$ p=0.067 n=1301	MD=0.55 $U=7.4\times10^5$ p=0.34 n=1250	MD=0.54 $U=7.4\times10^5$ p=0.44 n=1230

Table 8-3. Mann-Whitney U test details for increase in selectivity index. Significant increases in selective index were found for angle separations up to 2.0° in layer 5 and up to 1.0° in layer 4.

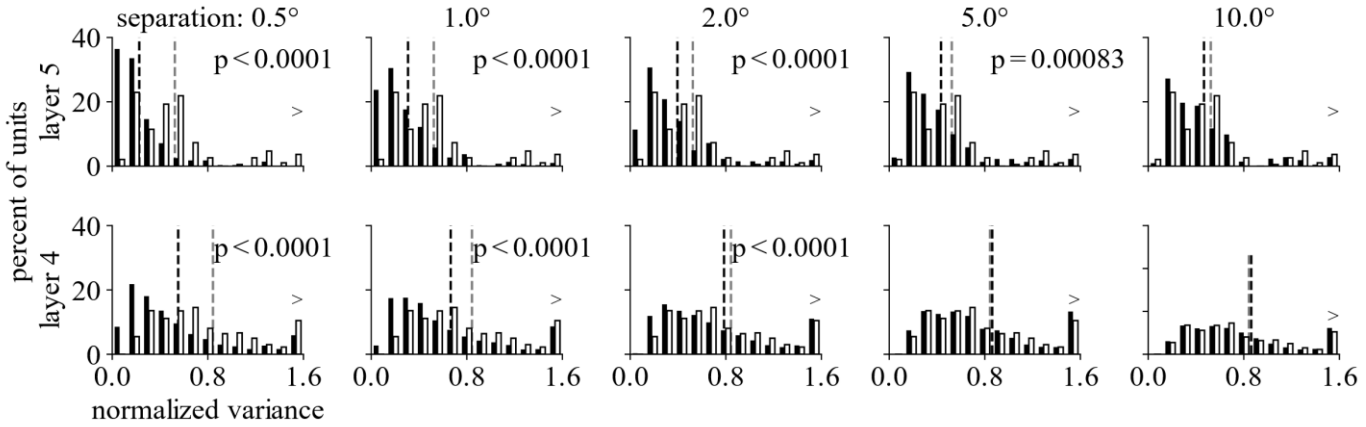

Figure 8-4. Distribution of normalized variance (or Fano factor) for the trained (black) and naïve (white) model unit populations in the two higher layers. Two rows correspond to layers 5 and 4; columns are different training angle separations. Black and gray dashed lines indicate the trained and naïve population means, respectively. Significant reductions in normalized variance were found for angle separations up to 5.0° in layer 5 and up to 2.0° in layer 4 ($p<0.001$, see Table 8-4 for details).

Statistical test details for decrease in normalized variance

separation		**0.5°**	**1.0°**	**2.0°**	**5.0°**	**10.0°**
layer	**5** (MD_{naive}=0.45, n_{naive}=192)	**MD=0.16** **U=1.02×10^4** **p=7.3×10^{-31}** **n=280**	**MD=0.22** **U=1.58×10^4** **p=6.5×10^{-17}** **n=295**	**MD=0.29** **U=1.84×10^4** **p=2.7×10^{-7}** **n=264**	**MD=0.34** **U=1.75×10^4** **p=0.00083** **n=222**	MD=0.38 U=1.84×10^4 p=0.041 n=213
	4 (MD_{naive}=0.66, n_{naive}=1203)	**MD=0.39** **U=4.87×10^5** **p=5.7×10^{-66}** **n=1334**	**MD=0.46** **U=5.81×10^5** **p=2.3×10^{-31}** **n=1318**	**MD=0.58** **U=6.93×10^5** **p=3.5×10^{-7}** **n=1301**	MD=0.65 U=7.37×10^5 p=0.20 n=1250	MD=0.66 U=7.35×10^5 p=0.39 n=1230

Table 8-4. Mann-Whitney U test details for decrease in normalized variance. Significant reductions in normalized variance were found for angle separations up to 5.0° in layer 5 and up to 2.0° in layer 4.

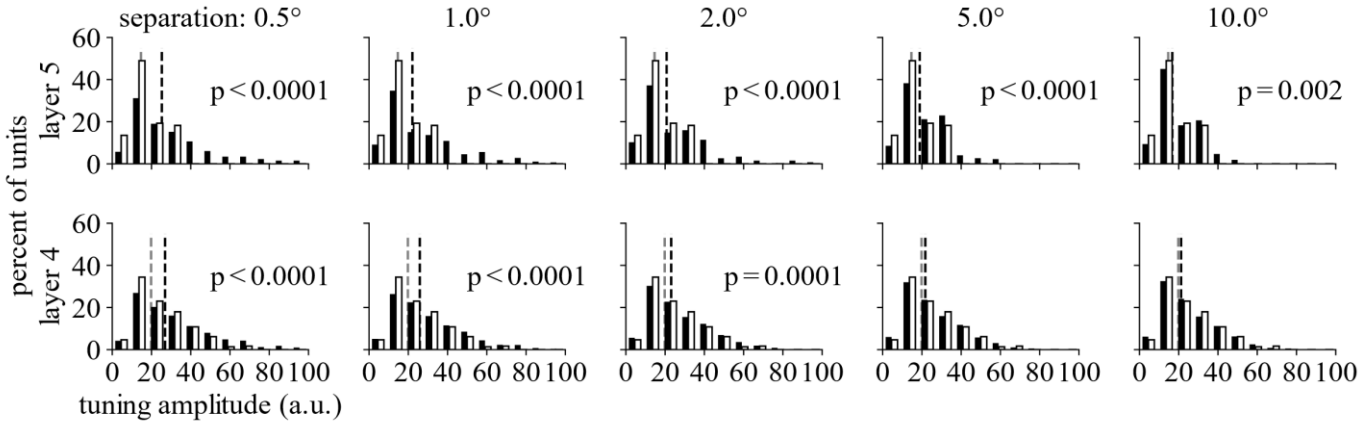

Figure 9-1. Distribution of tuning amplitude for the trained (black) and naïve (white) model units in the two higher layers. Two rows correspond to layers 5 and 4; columns are different training angle separations. Black and gray dashed lines indicate the trained and naïve population medians, respectively. Significant increases in tuning amplitude were found for angle separations under all precisions in layer 5 and up to 2.0° in layer 4 ($p<0.003$, see Table 9-1 for details).

Statistical test details for increase in response amplitude

separation		**0.5°**	**1.0°**	**2.0°**	**5.0°**	**10.0°**
layer	**5** (MD_{naive}=14.64, n_{naive}=192)	**MD=25.28** **U=1.54×10^4** **p=1.7×10^{-15}** **n=280**	**MD=22.09** **U=1.90×10^4** **p=4.2×10^{-10}** **n=295**	**MD=20.77** **U=1.88×10^4** **p=1.2×10^{-6}** **n=264**	**MD=18.88** **U=1.62×10^4** **p=1.5×10^{-5}** **n=222**	**MD=16.72** **U=1.71×10^4** **p=0.0020** **n=213**
	4 (MD_{naive}=19.79, n_{naive}=1203)	**MD=26.92** **U=6.41×10^5** **p=1.1×10^{-18}** **n=1334**	**MD=25.85** **U=6.64×10^5** **p=9.5×10^{-13}** **n=1318**	**MD=23.10** **U=7.15×10^5** **p=0.0001** **n=1301**	MD=21.73 U=7.25×10^5 p=0.062 n=1250	MD=21.22 U=7.24×10^5 p=0.17 n=1230

Table 9-1. Mann-Whitney U test details for increase in response amplitude. Significant increases in tuning amplitude were found for angle separations under all precisions in layer 5 and up to 2.0° in layer 4.

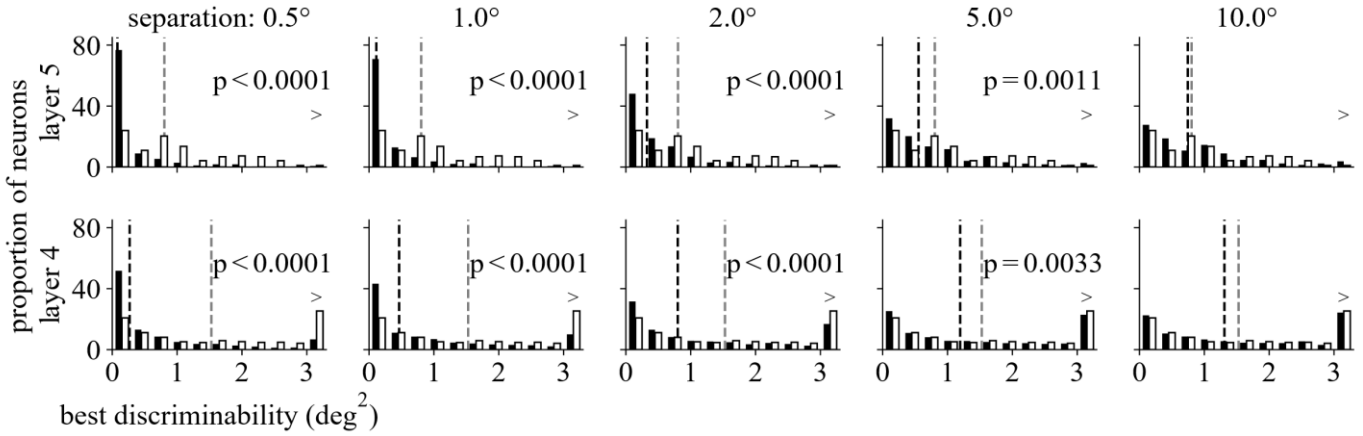

Figure 9-2. Distribution of best discriminability for the trained (black) and naïve (white) model units in the two higher layers. Two rows correspond to layers 5 and 4; columns are different training angle separations. Black and gray dashed lines indicate the trained and naïve population medians, respectively. Significant reductions in best discriminability were found for angle separations up to 5.0° in both layers ($p<0.005$, see Table 9-2 for details).

Statistical test details for decrease in best discriminability

separation		**0.5°**	**1.0°**	**2.0°**	**5.0°**	**10.0°**
layer	**5** (MD_{naive}=0.8, n_{naive}=192)	**MD=0.08** **U=8.08×10³** **p=1.9×10⁻³⁸** **n=280**	**MD=0.11** **U=9.48×10³** **p=1.1×10⁻³⁵** **n=295**	**MD=0.33** **U=1.46×10⁴** **p=4.8×10⁻¹⁵** **n=264**	**MD=0.55** **U=1.76×10⁴** **p=0.0011** **n=222**	MD=0.74 U=1.90×10⁴ p=0.10 n=213
	4 (MD_{naive}=1.53, n_{naive}=1203)	**MD=0.27** **U=4.19×10⁵** **p=2.3×10⁻⁹⁶** **n=1334**	**MD=0.46** **U=5.13×10⁵** **p=2.2×10⁻⁵³** **n=1318**	**MD=0.80** **U=6.34×10⁵** **p=1.1×10⁻¹⁶** **n=1301**	**MD=1.20** **U=7.04×10⁵** **p=0.0033** **n=1250**	MD=1.31 U=7.18×10⁵ p=0.10 n=1230

Table 9-2. Mann-Whitney U test details for increase in best discriminability. Significant reductions in best discriminability were found for angle separations up to 5.0° in both layers.

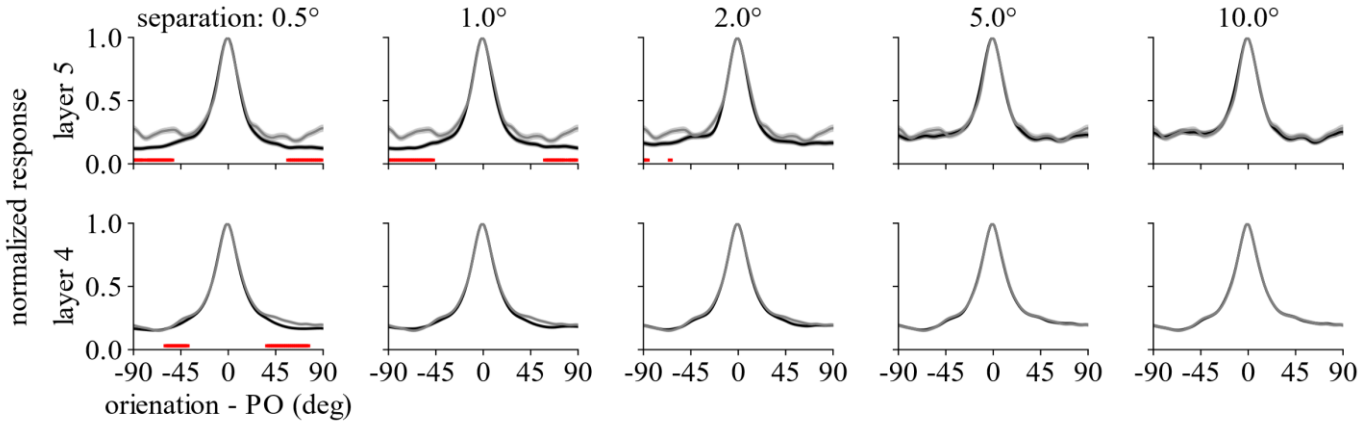

Figure 9-3. Average normalized tuning curves in for the trained (black) and naïve (gray) model units in the two higher layers. Two rows correspond to layers 5 and 4; columns are different training angle separations. Training reduced the activation away preferred orientations (PO) in layer 5 for small angle separations. By grouping the activations into two groups according to whether the test stimulus was oriented within 45° of the PO, two-way ANOVA revealed a significant interaction of orientation group × training on activation in layer 5 for up to 2.0° and in layer 4 for 0.5° angle separation ($p<0.01$, two-way ANOVA, see Table 9-3 for details), suggesting that the activation of non-POs was significantly more reduced than that of POs. Red lines indicate orientations where the trained activation was significantly less than naïve (threshold $p=0.01$, Mann-Whitney U, Bonferroni-corrected for 100 test orientations).

Statistical test details for interaction orientation group × training on activation

separation		**0.5°**	**1.0°**	**2.0°**	**5.0°**	**10.0°**
layer	**5**	**F(1,940)=13.02 p=0.00032**	**F(1,970)=15.86 $p=7.3\times10^{-5}$**	**F(1,908)=6.73 p=0.0096**	F(1,824)=0.36 p=0.55	F(1,806)=0.04 p=0.84
	4	**F(1,5070)=7.19 p=0.0074**	F(1,5038)=3.54 p=0.06	F(1,5004)=1.74 p=0.19	F(1,4902)=0.03 p=0.87	F(1,4862)=0.00 p=0.98

Table 9-3. ANOVA test details for interaction orientation group × training on activation. The individual main effects were included in the analysis. By grouping the activations into two groups according to whether the test stimulus was oriented within 45° of the preferred orientation, the analysis revealed a significant interaction of orientation group × training on activation in layer 5 for up to 2.0° and in layer 4 for 0.5° angle separation.

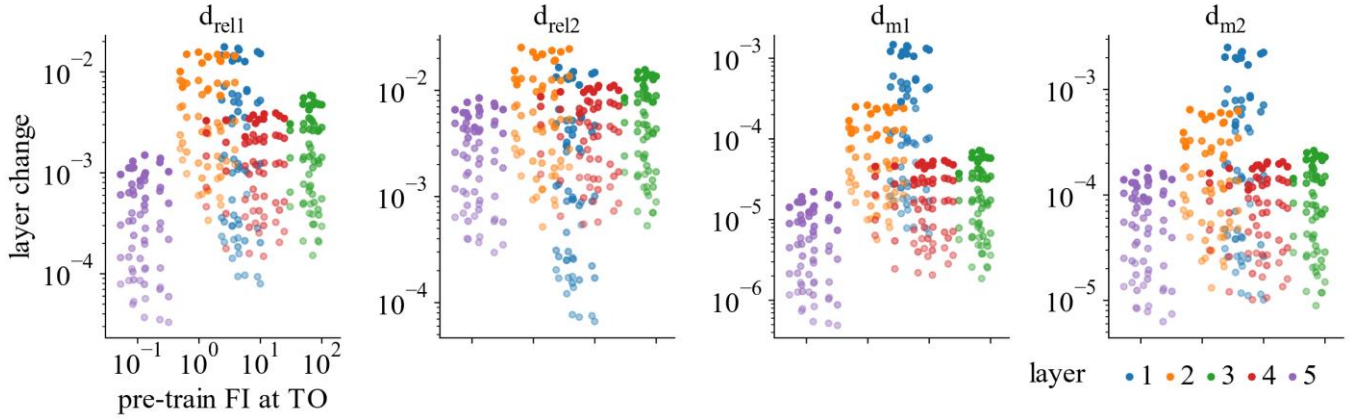

Figure 10-1. Weight change at each layer under the four different measures (columns) against pre-train linear fisher information (FI) at the trained orientation (TO). Darker color indicates smaller angle separation. The data were obtained from the simulations where unit tuning curves were taken (noise standard deviation 15, contrast 20% and wavelength 10 pixels). Units with responses less than 0.1 were excluded. The four measures are defined in Equations 4, 4a, 4b and 4c, respectively. Under all layer change measures, we conducted regression analyses using Equation 7, which revealed significant main effects of layer and angle separation ($p<0.0001$), but the effects involving FI were insignificant except for d_{m1} and d_{m2}. See Table 10-1 for statistical test details.

Statistical test details for various effects on layer weight change

Factors	$\boldsymbol{d_{rel1}}$	$\boldsymbol{d_{rel2}}$	$\boldsymbol{d_{m1}}$	$\boldsymbol{d_{m2}}$
pre-train Fisher information (FI)	F(1,282)=0.0069 p=0.93 R2=7.7×10^{-6}	F(1,282)=0.031 p=0.86 R2=1.9×10^{-5}	F(1,282)=0.00032 p=0.99 R2=5.6×10^{-7}	F(1,282)=0.00077 p=0.98 R2=1.3×10^{-6}
Layer (L)	**F(4,282)=48.27** **p=6.7×10^{-31}** **R2=0.21**	**F(4,282)=50.51** **p=5×10^{-32}** **R2=0.13**	**F(4,282)=40.58** **p=7.6×10^{-27}** **R2=0.28**	**F(4,282)=33.83** **p=4.7×10^{-23}** **R2=0.22**
angle separation (S)	**F(4,282)=103.3** **p=4.7×10^{-54}** **R2=0.46**	**F(4,282)=275.6** **p=4.3×10^{-96}** **R2=0.69**	**F(4,282)=28.38** **p=8×10^{-20}** **R2=0.2**	**F(4,282)=42.84** **p=4.6×10^{-28}** **R2=0.28**
L × FI	F(4,282)=0.02025 p=1.00 R2=9×10^{-5}	F(4,282)=0.1487 p=0.96 R2=0.00037	F(4,282)=0.001148 p=1.00 R2=8×10^{-6}	F(4,282)=0.005635 p=1.00 R2=3.7×10^{-5}
S × FI	F(4,282)=2.887 p=0.023 R2=0.013	F(4,282)=0.5435 p=0.70 R2=0.0014	**F(4,282)=3.83** **p=0.0048** **R2=0.027**	**F(4,282)=3.704** **p=0.0059** **R2=0.025**

Table 10-1. Statistical test details for various effects on layer weight change. The data were obtained from the simulations where unit tuning curves were taken (noise standard deviation 15, contrast 20% and wavelength 10 pixels). The four measures in the title row are defined in Equations 4, 4a, 4b and 4c, respectively. Bold indicates significant effects.

Statistical test details for various effects on unit weight change

Factors	d_{rel1}	d_{rel2}	d_{m1}	d_{m2}
pre-train Fisher information (FI)	**F(1,18984)=127.76** **$p=1.6\times10^{-29}$** **R2=0.0053**	**F(1,18984)=177.70** **$p=2.3\times10^{-40}$** **R2=0.0074**	**F(1,18984)=130.61** **$p=3.8\times10^{-30}$** **R2=0.0038**	**F(1,18984)=198.96** **$p=5.9\times10^{-45}$** **R2=0.006**
Layer (L)	**F(4,18984)=158.16** **$p=2.2\times10^{-133}$** **R2=0.026**	**F(4,18984)=32.60** **$p=4\times10^{-27}$** **R2=0.0054**	**F(4,18984)=819.25** **p=0** **R2=0.095**	**F(4,18984)=478.17** **p=0** **R2=0.058**
angle separation (S)	**F(4,18984)=750.58** **p≈0.0** **R2=0.12**	**F(4,18984)=973.87** **p≈0.0** **R2=0.16**	**F(4,18984)=1287.37** **p≈0.0** **R2=0.15**	**F(4,18984)=1625.12** **p≈0.0** **R2=0.2**
L × FI	**F(4,18984)=22.47** **$p=1.5\times10^{-18}$** **R2=0.0037**	**F(4,18984)=24.03** **$p=7.4\times10^{-20}$** **R2=0.004**	**F(4,18984)=191.88** **$p=1.5\times10^{-161}$** **R2=0.022**	**F(4,18984)=168.17** **$p=9.4\times10^{-142}$** **R2=0.02**
S × FI	**F(4,18984)=32.31** **$p=7\times10^{-27}$** **R2=0.0054**	**F(4,18984)=38.87** **$p=1.8\times10^{-32}$** **R2=0.0065**	**F(4,18984)=35.11** **$p=2.9\times10^{-29}$** **R2=0.0041**	**F(4,18984)=48.62** **$p=9.4\times10^{-41}$** **R2=0.0059**

Table 10-2. Statistical test details for various effects on unit weight change. The data were obtained from the simulations where unit tuning curves were taken (noise standard deviation 15, contrast 20% and wavelength 10 pixels). The four measures in the title row are defined in Equations 4, 4a, 4b and 4c, respectively. All effects are significant under all change measures, but angle separation accounted for the majority of variance.