



Double training downshifts the threshold vs. noise contrast (TvC) functions with perceptual learning and transfer

Xin-Yu Xie, Cong Yu*

ARTICLE INFO

Perceptual learning
Double training
External noise
Vernier acuity

ABSTRACT

Location specific perceptual learning can transfer to a new location if the new location is trained with a secondary task that by itself does not impact the performance of the primary learning task (double training). Learning may also transfer to other locations when double training is performed at the same location. Here we investigated the mechanisms underlying double-training enabled learning and transfer with an external noise paradigm. Specifically, we measured the Vernier thresholds at various external noise contrasts before and after double training. Double training mainly vertically downshifts the TvC functions at the training and transfer locations, which may be interpreted as improved sampling efficiency in a linear amplifier model or a combination of internal noise reduction and external noise exclusion in a perceptual template model at both locations. The change of the TvC functions appears to be a high-level process that can be remapped from a training location to a new location after double training.

1. Introduction

Visual perceptual learning is often specific to the retinal location and orientation/direction of the trained stimulus. Therefore, learning is often interpreted as a result of training induced neural plasticity, such as sharpened orientation/direction tuning of neurons, in retinotopic and feature selective early visual areas (Ball & Sekuler, 1982; Karni & Sagi, 1991; Schoups, Vogels, & Orban, 1995; Teich & Qian, 2003; Bejjanki, Beck, Lu, & Pouget, 2011). Alternatively, also bounded by learning specificity, reweighting theories propose that a decision stage reweights the inputs from stimulus-specific sensory neurons to improve the readout (Poggio, Fahle, & Edelman, 1992; Doshier & Lu, 1998; Yu, Klein, & Levi, 2004; Petrov, Doshier, & Lu, 2005; Law & Gold, 2009).

Our recent studies demonstrate that specificity is not an inherent property of perceptual learning. Visual perceptual learning can transfer to a new location if the new location is additionally trained with a secondary task that by itself has no impact on the performance of primary learning task (Xiao et al., 2008; Wang, Zhang, Klein, Levi, & Yu, 2012; Wang, Cong, & Yu, 2013). Sometimes when double training is performed at the same training location, learning can also transfer to new locations (Wang, Zhang, Klein, Levi, & Yu, 2014). Learning also transfers to a new orientation/direction when a secondary task is practiced at the new orientation/direction to eliminate feature specificity (Zhang, Zhang, Xiao, Klein, Levi, & Yu, 2010; Zhang & Yang,

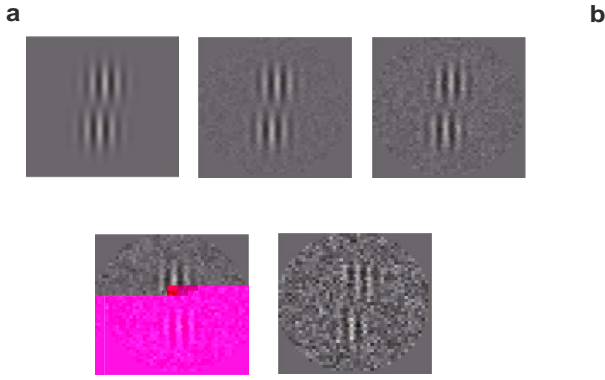
2014; Xiong, Xie, & Yu, 2016; Zhang & Yu, 2016). These results suggest that visual perceptual learning is at least in some situations a high-level process that occurs beyond the retinotopic and feature selective visual areas.

Orientation and direction learning also transfers with double training to physically distinct stimuli (e.g., orientations defined by luminance gratings vs. symmetry axes of random dot patterns; motion directions defined by first-order luminance vs. second-order contrast gratings) that are initially encoded by different neural substrates and discriminated at separate threshold ranges (Wang et al., 2016). These data indicate that what is learned is more likely the concept of a trained visual feature (e.g., an abstract concept of orientation or motion direction) that is independent of retinal location, feature dimension, physical property, and putative neural encoders. Moreover, either the top-down or the bottom-up influences produced by the stimulus at the new location or orientation in the secondary training task, when isolated by a revised continuous flash suppression method (Tsuchiya & Koch, 2005), can enable significant transfer of primary learning. This finding suggests that learning specificity may result from absent or weak functional connections between high-level learning and visual inputs at untrained conditions that are neither bottom-up stimulated nor top-down attended during training (Xiong, Zhang, & Yu, 2016).

Previous studies have applied external noise paradigms to study the mechanisms of perceptual learning (Burgess, Wagner, Jennings, &

* Corresponding author.

E-mail address: yucong@pku.edu.cn (C. Yu).



Barlow, 1981; Legge, Kersten, & Burgess, 1987; Pelli, 1991; Doshier & Lu, 1998, 1999). In these studies the contrast thresholds for performing a certain visual task are measured with the target stimuli presented in external noise of various contrasts. The contrast threshold plotted against the noise contrast in log-log axes is called a threshold vs. noise contrast (TvC) function. Training typically down-shifts the thresholds at all noise contrasts (i.e. the entire TvC function) vertically (Doshier & Lu, 1998; Gold, Bennett, & Sekuler, 1999). A linear amplifier model with two parameters (i.e., equivalent internal noise and sampling efficiency) would interpret this vertical downshift as improved sampling efficiency and unchanged internal noise (Burgess et al., 1981; Legge et al., 1987; Pelli, 1991). Alternatively, Doshier and Lu (1998, 1999) proposed a perceptual template model, which is more complex than the two-component model with additional considerations of internal multiplicative noise and nonlinearities. They attributed the TvC function down-shift to a combination of internal noise reduction and external noise exclusion.

We applied the external noise paradigm in the current study to investigate the mechanisms underlying double training. Like previous reports, our results showed that training led to a vertical downshift of the TvC functions at the training location. Moreover, the TvC functions at the transfer location were equally downshifted with learning transfer after double training, indicating that similar learning mechanisms underlie learning and transfer. These results are consistent with our theory that the same high-level rules govern learning and transfer in perceptual learning (Zhang et al., 2010).

2. Methods

2.1. Participants

The observers consisted of thirty-four undergraduate and graduate students (18–27 years old) with normal or corrected-to-normal vision. They were new to psychophysical experiments and naïve to the purposes of the study. Informed written consent, which was approved by the Peking University IRB, was obtained before data collection from each observer. This work was carried out in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki).

The stimuli were generated with Psychtoolbox-3 (Pelli, 1997) and presented on a 21-in CRT monitor (1024 pixel \times 768 pixel, 0.39 mm \times 0.39 mm pixel size, 120 Hz frame rate, and 33.4 cd/m² mean luminance). The screen luminance was linearized by an 8-bit look-up table. Viewing was monocular at a distance of 1 m, and a chin-and-head rest stabilized the head. Viewing was through a circular opening (diameter = 17°) of a black cardboard that covered the rest of the monitor screen. Experiments were run in a dimly lit room.

An Eyelink-1000 eye-tracker (SR Research, Kanata, Ontario, Canada) monitored eye movements in one-third of the observers in each experiment to double check the potential eye movement effects. Trials

were excluded from data analysis if eye positions deviated from the fixation point more than 2° immediately before and during the stimulus presentation. Our previous study indicated no significant changes of eye drifts after training with a peripheral task (Zhang et al., 2010), excluding the possibility that peripheral threshold improvements may result from eye movement pattern changes after training. Here we compared the results of observers with and without using the eye tracker. We pooled the double training data in Figs. 4b and 5b and contrasted the results obtained with vs. without eye tracking (N = 5 vs. 8). A repeated-measures ANOVA showed no significant effects of eye tracking at both training location ($F_{1,9} = 2.439$, $p = .153$) and transfer location ($F_{1,9} = 2.381$, $p = .157$), suggesting that learning and transfer results reported in this paper were not significantly compromised by eye movements.

2.2. Stimulus

The Vernier stimulus consisted of two identical Gabors (Gaussian-windowed sinusoidal gratings) presented on a mean luminance screen background. The Vernier was centered on one visual quadrant at 5° retinal eccentricity. The two Gabors had the same spatial frequency (3 cpd), standard deviation (0.67°), contrast (0.47), orientation (vertical) and phase (0°). The center-to-center distance of two Gabors was 1.33°. To form a specific Vernier offset, the position of each Gabor shifted half the Vernier offset away in opposite directions perpendicular to the Gabor orientation. The Vernier was imbedded in external noise in a circular window (radius = 2°) (Fig. 1a). Each noise element was 4 \times 4 pixel in size, and the luminance of each element was sampled from the look-up-table following a Gaussian distribution. The root mean square (rms) contrast of the external noise was 0%, 5%, 9%, 16%, or 29%. The Vernier and the noise stimuli were presented in alternating frames in actual experiments (6 frames each for a total duration of 100 ms).

The stimulus for orientation discrimination was a single Gabor presented at 5° retinal eccentricity either in the same quadrant or diagonal to the Vernier quadrant. The Gabor was identical to those forming the Vernier stimulus. The reference orientation of the Gabor was either 36° or 126°.

2.3. Procedure

The Vernier threshold was measured with a one-interval staircase procedure. In each trial, a small fixation cross preceded the Vernier by 500 ms and stayed throughout the trial. The Vernier was presented for 100 ms. Observers reported whether the lower Gabor was to the left or right of the upper Gabor by key press. Auditory feedback was given on incorrect responses.

The orientation discrimination threshold was measured with a two-interval forced-choice staircase procedure. In each trial, a small fixation

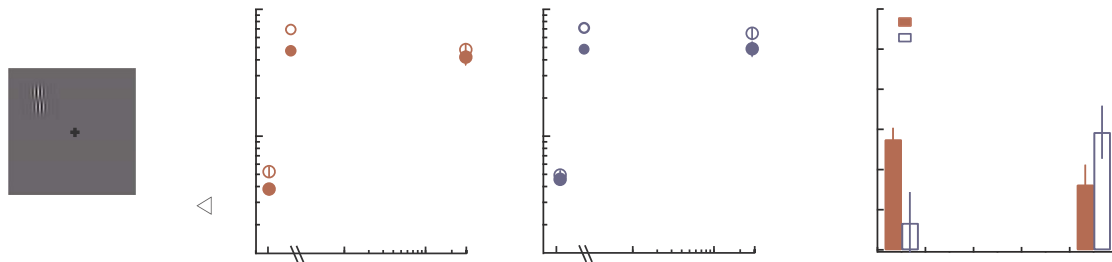


Fig. 2. Baseline Vernier training at zero noise. (a) Stimulus. Vernier was trained at the upper-left or lower-right quadrant, and learning transfer was tested at the diagonal quadrant. (b) Pre- and post-training Vernier thresholds at the training (left panel) and diagonal transfer (right panel) locations. c. The mean percent improvements at two measured noise contrasts at the training and transfer locations.

cross preceded the first interval by 500 ms and stayed throughout the trial. The Gabors at the reference orientation and the test orientation (reference + Δori) were shown in two 100-ms stimulus intervals, respectively, in a random order. The two stimulus intervals were separated by a 500-ms inter-stimulus interval. The observers judged which stimulus interval contained the more clockwise-oriented Gabor. Auditory feedback was given on incorrect responses.

Thresholds were estimated following a 3-down-1-up staircase rule that converged at a 79.4% correct response rate. Each staircase consisted of four preliminary reversals and six experimental reversals (approximately 50–60 trials). The step size of the staircase was 0.05 log units. The geometric mean of the experimental reversals was taken as the threshold for each staircase run.

In a pre- or post-training session in most experiments (Figs. 3–6), the Vernier thresholds at five noise contrasts were measured at two diagonal locations in a counterbalanced order, with each condition tested for 5 staircases, for a total of 50 staircases. These 50 staircases were completed in two daily sessions with the test sequence preset with a permuted table. The pre- and post- training sessions in Fig. 2 were shorter with fewer staircases, and were completed in a single daily session. The training sessions lasted six days, each consisting of 10 staircases of Vernier task at zero noise and/or 10 staircases of orientation discrimination task at zero noise. Each 20-staircase session lasted for 1.5–2 h.

2.4. f_i

We revised a linear-amplifier model (Burgess et al., 1981; Legge et al., 1987; Pelli, 1991) to fit the pre- and post-training TvC (Threshold vs. Noise Contrast) functions. The original format of the model is:

$$Th^2 = \frac{1}{k} (N_{ext}^2 + N_i^2)$$

in which Th stands for contrast threshold, N_{ext} stands for external noise, N_i stands for sampling efficiency, and k stands for equivalent internal noise that is additive. We noticed in our data that the Vernier thresholds at the highest noise were too high to be fitted by the model, which could suggest extra masking effect that increases with the noise contrast. Therefore, we introduced a new parameter γ to the model to represent this effect that mainly impacts the thresholds at high noise and thus the slope of the TvC function. The revised model is:

$$Th^2 = \frac{1}{k} (N_{ext}^{2\gamma} + N_i^2)$$

In this new model, reduced k by training would lead to lower thresholds at low noise (blue curve), and increased γ would lead to lower thresholds at high noise (green curve) because the noise contrast is less than 1. However, a larger γ , or a vertical downshift of the entire TvC function (red curve), has different interpretations. As described

earlier, it could either indicate improved sampling efficiency with unchanged equivalent internal noise (Burgess et al., 1981; Legge et al., 1987; Pelli, 1991), or a combination of internal noise reduction and external noise exclusion (Doshier & Lu, 1998, 1999, 2005).

Data fitting was performed with a nonlinear least square method (the Matlab lsqnonlin function) and weighted with the standard error of each data point. The model fitted the individual TvC functions quite well (the mean adjusted goodness of fit $R^2 = 0.93$). It should be noticed that reducing k and increasing γ together have the same effect as increasing k alone. Therefore, when fitting the TvC functions we either kept γ fixed or only allowed γ to vary. Details are provided with each specific experiment in the Results section.

It is noteworthy that the linear amplifier model (Burgess et al., 1981; Legge et al., 1987; Pelli, 1991) and the perceptual template model (Doshier & Lu, 1998, 1999, 2005) have equations that were derived for the functional form of the TvC curves for contrast thresholds. Our experiments instead measured Vernier offset thresholds. Nonetheless, we used the functional form for contrast TvC curves as a heuristic approximation, based on the observation that both measures are limited by discriminability d' , and that the Vernier TvC functions are well approximated by this form.

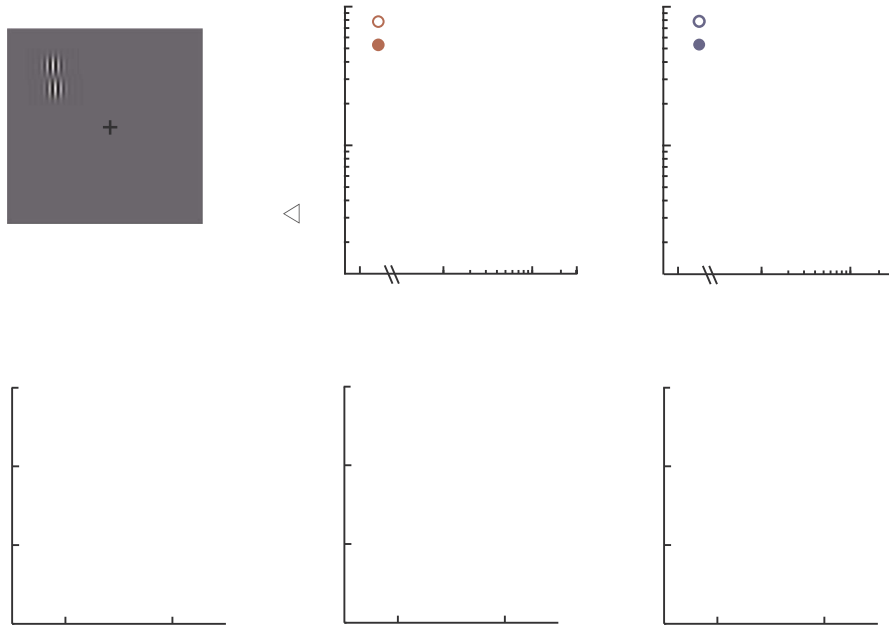
3. Results

3.1. Baseline Vernier training at zero noise

We first replicated location specificity in Vernier learning. Seven observers practiced Vernier discrimination at zero noise at one quadrant location for six days (Fig. 2a). In pre- and post-training sessions, Vernier thresholds were measured at zero and the highest noise contrast (0.29) at the training and diagonal transfer locations. The reason for not testing all noise levels was the concern that too many pretesting trials at the transfer location would lead to some degree of double training effects, as would be shown later in Fig. 3.

The learning effect was measured by calculating the percent improvement $[(1 - \text{post_threshold}/\text{pre_threshold}) * 100]$. Training improved Vernier thresholds at zero noise by $27.5 \pm 2.9\%$ ($t = 9.47$, $df = 6$, $p < .001$, paired two-tailed t -test here and in later analyses unless otherwise specified) and at the highest noise by $16.2 \pm 5.0\%$ ($t = 3.25$, $df = 6$, $p = .018$) at the training location (Fig. 2b, c). The learning transfer from zero noise to high noise was consistent with Doshier and Lu (2005). At the untrained diagonal location, Vernier performance did not change significantly at zero noise ($6.7 \pm 7.7\%$, $t = 0.89$, $df = 6$, $p = .41$), replicating location specificity. The Vernier performance at the highest noise was improved by $29.3 \pm 6.6\%$ ($t = 4.43$, $df = 6$, $p = .004$), indicating that learning transfer from zero noise to high noise was location unspecific (Fig. 2b, c).

In addition, we repeated the above experiment in another seven



observers with Vernier thresholds measured at all five levels of noise contrasts in pre- and post-training sessions (Fig. 3a). Like in Fig. 2, training improved Vernier thresholds at zero noise by $27.4 \pm 5.2\%$ ($t = 5.28$, $df = 6$, $p = .002$) and the highest noise by $26.9 \pm 7.2\%$ ($t = 3.74$, $df = 6$, $p = .010$) at the training location (Fig. 3b). At the untrained diagonal location, training did not change Vernier performance significantly at zero noise ($15.6 \pm 7.2\%$, $t = 2.16$, $df = 6$, $p = .074$) because of large individual differences, but it improved the performance at the highest noise ($22.0 \pm 5.8\%$, $t = 3.77$, $df = 6$,

$p = .009$) (Fig. 3b).

We first fitted the pre-training TvC functions to find the best values of the three model parameters (Fig. 3c). For post-training functions, because there was no training at high noise, we assumed that the parameter v , which indicated the effects of high noise, would not change. Thus we fixed v at the pre-training value and let σ and γ vary (Fig. 3c, smooth curves). The fitting results indicated increased σ at the training location ($t = 2.48$, $df = 6$, $p = .047$), which suggested improved sampling efficiency in a linear amplifier model or a combination

of internal noise reduction and external noise exclusion in a perceptual template model. However, there was no significant change of σ at the untrained location ($t = 1.75$, $df = 6$, $p = .130$) because of the large error bars. Data fitting also indicated no significant changes of σ at the training location ($t = 1.12$, $df = 6$, $p = .30$) and the transfer location ($t = 0.33$, $df = 6$, $p = .75$). These results together suggested that Vernier training with pretests at all noise contrasts failed to change the TvC functions significantly at the untrained transfer location. On the other hand, some observers did show more transfer effects at the untrained location, as suggested by higher Vernier improvement at zero noise and larger change of model parameter σ when compared to those in Fig. 2, even if these changes were not statistically significant due to large individual differences.

3.2. σ vs γ

Six new observers practiced the same Vernier task at zero noise, as well as an orientation discrimination task at zero noise at a diagonal quadrant location, in alternating blocks of trials within the same sessions (Fig. 4a). The orientation task served as the secondary location training in a double-training design. Vernier thresholds at five noise contrasts were measured at the training and diagonal transfer locations before and after training.

The double training improved Vernier thresholds at zero noise at the training location by $39.3 \pm 6.2\%$ ($t = 6.36$, $df = 5$, $p = .001$), as well as at the diagonal location by $28.4 \pm 7.4\%$ ($t = 3.82$, $df = 5$, $p = .012$) (Fig. 4b). The two improvements were not significantly different from each other ($t = 1.13$, $df = 10$, $p = .28$), replicating our previous results that Vernier learning becomes largely location-unspecific after double training (Xiao et al., 2008; Wang et al., 2012).

Again we fixed γ at its pre-training value because there was no training at high noise contrasts and let other parameters vary. We found increased σ at both training ($t = 3.19$, $df = 5$, $p = .024$) and transfer ($t = 2.60$, $df = 5$, $p = .048$) locations, along with no significant changes of γ ($t = 0.99$ and 0.97 , $df = 5$ and 5 , $p = .36$ and $.37$ at the training and diagonal locations, respectively) (Fig. 4c). These results indicate vertical downshifts of the TvC functions at both training and transfer locations, as a result of improved sampling efficiency (Burgess et al., 1981; Legge et al., 1987; Pelli, 1991), or a combination of internal noise reduction and external noise exclusion (Doshier & Lu, 1998, 1999, 2005).

3.3. σ vs ff

Vernier learning, when paired with orientation training at the same location, can also transfer to other retinal locations (Wang et al., 2014). Here we repeated this “piggybacking” effect with same-location Vernier and orientation training at zero noise in alternating blocks of trials (Fig. 5a), while pre- and post-training Vernier thresholds were measured at five noise contrasts at the training location and a diagonal location. Fig. 5b shows improvements of Vernier thresholds at all noise contrasts at both training and diagonal locations. At zero noise the Vernier thresholds were similarly ($t = 0.69$, $df = 12$, $p = .51$) reduced by $29.9 \pm 4.5\%$ ($t = 6.62$, $df = 6$, $p = .001$) at the training location and by $25.4 \pm 4.7\%$ ($t = 5.37$, $df = 6$, $p = .002$) at the diagonal location, consistent with our previous finding of complete learning transfer (Wang et al., 2014).

Following the same fitting procedure as in Figs. 3 and 4 (i.e., fixing γ and searching for best σ and ff), we found increased σ at both training ($t = 3.32$, $df = 6$, $p = .016$) and transfer ($t = 3.30$, $df = 6$, $p = .017$) locations, with no significant changes of ff ($t = 1.39$ and 1.10 , $df = 6$ and 6 , $p = .21$ and $.32$ at the training and diagonal locations, respectively) (Fig. 5c). Overall the experimental and fitting results are similar to those in Fig. 4.

3.4. σ vs γ

The control experiment tested the possibility that pretesting at all five noise contrasts and orientation training could improve Vernier performance in Figs. 4 and 5 without primary Vernier training. Seven new observers only practiced orientation discrimination at zero noise at one quadrant location (Fig. 6a). Before and after training Vernier thresholds at all noise contrasts were measured at the orientation training location and a diagonal location (Fig. 6b).

The results showed that pretesting at all five noise contrasts and orientation training at zero noise had no significant impact on Vernier thresholds at zero noise at the orientation-training location ($4.1 \pm 5.6\%$, $t = 0.72$, $df = 6$, $p = .50$) and a diagonal location ($-2.5 \pm 6.4\%$, $t = 0.40$, $df = 6$, $p = .70$) (Fig. 6b). However, Vernier performance was improved significantly at the highest noise, by $25.7 \pm 7.4\%$ ($t = 3.48$, $df = 6$, $p = .013$) at the orientation-training location and $23.6 \pm 6.1\%$ ($t = 3.86$, $df = 6$, $p = .008$) at the diagonal location (Fig. 6b), indicating that the performance improvements at high noise contrasts were task unspecific.

We first fitted the pre-training TvC functions with three free parameters. Because the post-training TvC functions revealed no changes of σ and γ , we initially looked for the best γ with σ and ff fixed at the pre-training values. After that we fixed γ and allowed σ and ff to vary. The fitting results showed significant changes of γ ($t = 3.33$ and 2.80 , $df = 6$ and 6 , $p = .016$ and $.031$ at the training and diagonal locations, respectively), but no significant changes of σ and ff at both locations ($ps > 0.05$) (Fig. 6c). Thus, pretesting at all five noise contrasts and the secondary orientation training were not responsible for the Vernier improvements at low noise and the parameter σ increase at the untrained transfer location in Figs. 4 and 5.

4. Discussion

Our study replicated previous results that training led to a vertical downshift of the TvC functions (Doshier & Lu, 1998; Gold et al., 1999). After double training, similar changes of the TvC functions are also shown at the transfer location. These results suggest that the same high-level mechanisms may be underlying the performance improvements at two locations. This is consistent with our previous ERP evidence that C1, P1, and N1 changes are similar at both the training location and a new location showing learning transfer (Zhang, Cong, Song, & Yu, 2013; Zhang, Li, Song, & Yu, 2015).

The vertical downshift of the TvC functions at both training and transfer locations after double training can be interpreted by two models as we mentioned earlier. The two-parameter linear amplifier model would interpret the downshift as a result of improved sampling efficiency. Sampling efficiency reflects the weights of spatial and temporal integration of signal and noise information. An ideal observer has a perfect sampling efficiency (i.e., $k = 1$) because the receptive field shape matches the stimulus shape perfectly, but real observers are less efficient ($k < 1$). Improved sampling efficiency would suggest refined spatial-temporal summation to better match the signals and improve performance at all noise levels. In this context, our current results would suggest that double training improves the match between the stimulus and the receptive field not only at the training location, but also at the transfer location. On the other hand, Doshier & Lu’s model would interpret the vertical downshift of the TvC functions as a combination of internal noise reduction and external noise exclusion. In this context, our results would suggest internal noise reduction and external noise exclusion at both training and transfer locations after double training.

Like those at the training location, the TvC functions at the transfer location also show the same downshift after double training, suggesting that the same high-level rules (Zhang et al., 2010) may govern the TvC function downshift at both locations. Recently we demonstrated that orientation and motion direction learning can transfer completely

different parameter values for internal noise and efficiency. Our model is more like a simplified Doshier and Lu model because it also contains a parameter to simulate the extra high noise masking effects, likely caused by multiple simultaneous speech streams.