Available online at www.sciencedirect.com
SCIENCE DIRECT®
Vision Research 46 (2006) 3223–3233
Vision Research
www.elsevier.com/locate/visres

Perceptual learning with spatial uncertainties

Thomas U. Otto [a,*], Michael H. Herzog [a,b], Manfred Fahle [b,c], Li Zhaoping [d]

[a] *Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland*
[b] *Institute for Human Neurobiology, University of Bremen, Germany*
[c] *The Henry Wellcome Applied Vision Research Centre, London, UK*
[d] *Psychology Department, University College London, UK*

Received 24 May 2005; received in revised form 25 January 2006

Abstract

In perceptual learning, stimuli are usually assumed to be presented to a constant retinal location during training. However, due to tremor, drift, and microsaccades of the eyes, the same stimulus covers different retinal positions on sequential trials. Because of these variations the mathematical decision problem changes from linear to non-linear (Zhaoping, Herzog, & Dayan, 2003). This non-linearity implies three predictions. First, varying the spatial position of a stimulus within a moderate range does not deteriorate perceptual learning. Second, improvement for one stimulus variant can yield negative transfer to other variants. Third, interleaved training with two stimulus variants yields no or strongly diminished learning. Using a bisection task, we found psychophysical evidence for the first and last prediction. However, no negative transfer was found as opposed to the second prediction.

Keywords: Ideal observer model; Bisection task; Recurrent networks; Positional coding; Stimulus uncertainty

1. Introduction

1.1. Perceptual learning

Perceptual learning is the ability to improve perception per se. Examples are the improvement of vernier discrimination (e.g., McKee & Westheimer, 1978; Poggio, Fahle, & Edelman, 1992), stereoscopic depth perception (e.g., Ramachandran & Braddick, 1973), grating waveform discrimination (e.g., Fiorentini & Berardi, 1980), orientation discrimination (e.g., Vogels & Orban, 1985), and motion direction discrimination (e.g., Vaina, Sundareswaran, & Harris, 1995; Watanabe, Nanez, & Sasaki, 2001). Perceptual learning gained increasing interest in recent years since it is concerned with the creation of the building blocks of perception. We perceive the world, at least partly, according to how we have learned to perceive it. However, the mechanisms underlying perceptual learning are still largely unknown (e.g., Fahle & Poggio, 2003).

Perceptual learning is specific for many stimulus parameters. For example, no transfer of learning occurs when the orientation, the spatial frequency, or the position of stimuli change (e.g., Ahissar & Hochstein, 1997; Crist, Kapadia, Westheimer, & Gilbert, 1997; Fahle, Edelman, & Poggio, 1995; Karni & Sagi, 1991; Schoups, Vogels, & Orban, 1995; Shiu & Pashler, 1992). After the change, observers have to re-train to improve performance. Most models of perceptual learning (for a review see Tsodyks & Gilbert, 2004) implicitly assume that stimuli are projected exactly to the same retinal position at each trial. However, this assumption is not met in the experimentation. Because of eye tremor, drifts, and microsaccades, eye positions change from trial to trial and, hence, stimuli are never presented exactly to just one retinal position. Nevertheless, our perceptual performance often achieves a spatial acuity much finer than the typical magnitudes in the positional changes of the stimulus. This is, for instance, the case in the bisection task studied in this contribution. As shown in Fig. 1,

* Corresponding author. Fax: +41 21 6931749.
E-mail address: tom.otto@epfl.ch (T.U. Otto).

0042-6989/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2006.03.021

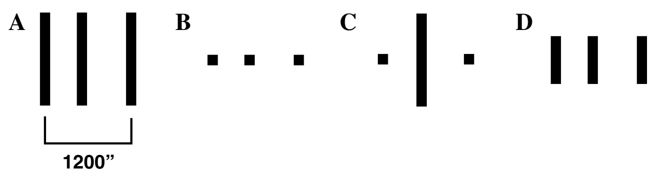

Fig. 1. Bisection stimuli. A spatial interval, delineated by two outer elements, is bisected in two unequal components by a central element. The task is to judge the resulting proportion, i.e., is the central element closer to the left or to the right outer delineator? The schematics correspond to bisection stimuli used in the experiments. (A) Line bisection (an outer line distance of 1200″ is shown). (B) Dot bisection. (C). Dot-line bisection. (D) Line bisection with lines having only half length.

observers are asked to judge whether a middle element is closer to one or the other of two outer elements. Thresholds in these tasks are in the order of a few tens of arc seconds, whereas the microsaccades are in the order of tens of arc minutes (Martinez-Conde, Macknik, & Hubel, 2004).

1.2. Linear versus non-linear computation

A recent ideal observer analysis showed that introducing positional uncertainties in the stimuli changes the underlying perceptual decision making problem from a linear to a non-linear problem which is approximately quadratic for small positional uncertainties (Zhaoping et al., 2003). Specifically, a decision about stimulus variants, in the form of a binary answer, may be made based on whether or not a certain quantity, determined in the ideal observer analysis, passes a threshold. As illustrated in Fig. 2A, a linear problem (or linear decision) means that this perceptual quantity can be computed as a *linear* weighted sum of responses of

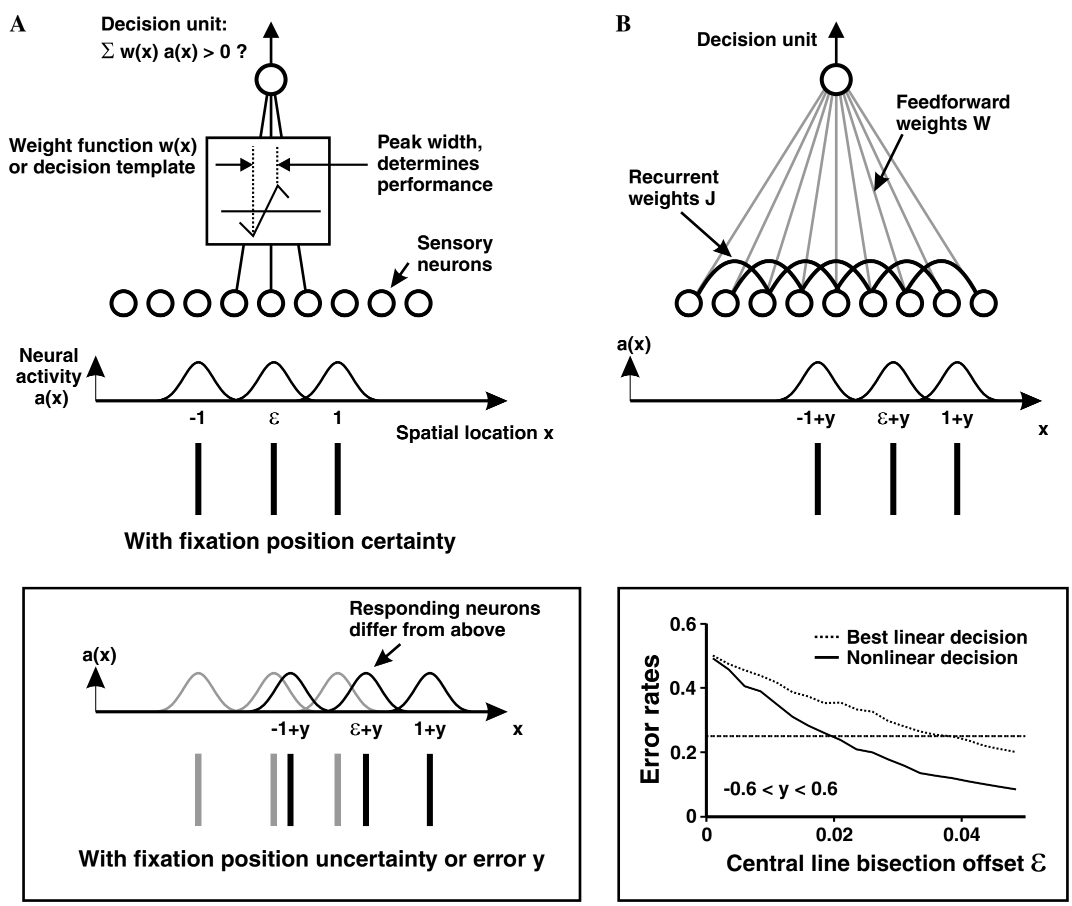

Fig. 2. (A) Linear decision on the bisection task. A line at location x evokes activities $a(x)$ in neurons tuned to that location. With fixed eye position (upper half), only the central line changes between trials ($x = \epsilon$), and the decision depends only on neurons responding to the central line, i.e., the weights $w(x) = 0$ for neurons responding to the outer lines (at normalized locations $x = \pm 1$). Performance threshold is determined by the spatial width of the weights $w(x)$. The decision is whether a weighted summation $\sum_x w(x)a(x)$ of the neural activations $a(x)$, with weights $w(x)$, passes a threshold. Eye position error y (bottom) shifts the stimulus array by y, and the decision will differ when y is larger than the width of $w(x)$. To compensate, this width needs to be larger, thus increasing the threshold. (B) Bisection task by non-linear decision. The decision network is augmented by recurrent connections J, and the feedforward weights $w(x)$ now examine inputs from neurons responding to all lines. Neurons have a sigmoid input-to-output function. The recurrent weights J are specific to the task or the outer element distance, linking neurons separated by roughly half of the outer element distance. They are near translation invariant, serve to amplify the task relevant signal ϵ, and reduce the effect of the task irrelevant noise y. Consequently, the error rate (bottom) is much smaller, compared to that of a best linear decision without the recurrent weights, reaching a threshold (at error rate 25%) of $\epsilon \sim 0.02$ even though the eye position uncertainty $-0.6 < y < 0.6$ is much larger.

early sensory neurons to the visual stimuli. For example, this quantity could be a linear weighted sum of responses from neurons tuned to orientation and/or spatial locations. A non-linear problem means that this quantity is a non-linear (e.g., quadratic) function of the same sensory responses. The purpose of the non-linear computation is to enable the perceptual decision to be insensitive to the positional variance of the stimuli. Consequently, the performance after training will be improved over a reasonable spatial range of stimulus positions.

Zhaoping et al. (2003) proposed that the non-linear computation could be achieved by recurrent connections between the non-linear sensory neurons before the neural responses are sent to the next, linear decision stage. In particular, the recurrent connections might be between V1 neurons, and could be modified during perceptual learning. Meanwhile, feedforward linear weights in the decision template of the subsequent linear decision stage depend on the resulting recurrent connections in the non-linear stage. They can be learned concurrently and, since mathematically they are much more easily and straight-forwardly learned, were suggested to correspond to the fast phase of perceptual learning (Zhaoping et al., 2003). To better understand the consequences of changing from a linear to a non-linear solution, we performed computer simulations with a recurrent network emulating this ideal observer model. In this network, the recurrent connections are specific to the task or the separation between the outer lines in the stimulus (Fig. 2B). They link neurons which are separated by distances about half of the outer element distance. In effect, they assess the distances between two stimulus lines separated by roughly half of the outer element distance, i.e., the left-to-center distance or the right-to-center distance, rather than the absolute locations of the lines. In doing so, they amplify the effect of the distance between elements and reduce the effect of the absolute stimulus locations.

Our non-linear recurrent model for perceptual learning is consistent with the observation that recurrent connections are known to be extensive between V1 neurons, extending over a distance up to a few millimeters (Gilbert & Wiesel, 1983; Rockland & Lund, 1983) and are modifiable by perceptual learning (Crist, Li, & Gilbert, 2001). In particular, the length of these V1 recurrent connections suggests that neurons can be linked with each other even though they are responding to two stimulus lines. Such connections are sufficient for our stimulus configurations. The degree of uncertainty in the spatial location of the visual stimulus is unlikely to be much more than one degree in visual angle for stimuli near the fovea. If the overall location of the stimulus changes from trial to trial by more than this amount, we assume that the brain could give a quick and rough estimate of the overall location with a similar degree of spatial precision or uncertainty, before the decision is made on the perceptual task at hand. Our analysis and the model lead to three predictions.

1.2.1. First prediction

If the bisection stimulus is presented at random positions, improvement of performance is comparable to when the bisection stimulus is presented constantly at one position. This positional invariance is reached by the recurrent neural connections in the network.

1.2.2. Second prediction

Training with a bisection stimulus of one given outer element distance can yield negative transfer to bisection stimuli with other outer element distances. Negative transfer means that performance in the pre-training conditions is superior to the post-training conditions. For positive transfer the opposite is true. This prediction arises from the fact that different recurrent connections would be used for different outer element distances. In particular, a connection between two neurons may be excitatory for one outer element distance and inhibitory for another outer element distance. Hence, if perceptual learning is achieved for the first outer element distance, the recurrent connections learned would be detrimental for the second outer element distance, causing negative transfer.

1.2.3. Third prediction

Improvement of performance can be diminished or even abolished if bisection stimuli with two different outer element distances are presented randomly interleaved during training. This prediction arises naturally from the second prediction, considering that perceptual learning would have to do the impossible: simultaneously achieve an excitatory connection between two neurons for one outer element distance and an inhibitory connection between the same neural pair for another outer element distance.

In the following, we will test these three predictions.

2. General methods

2.1. General set up

Stimuli appeared on an X–Y-display (HP 1332A, 1333A or Tektronix 608), either controlled by a Macintosh computer or PC via fast 16 bit D/A converters (1 MHz pixel rate). Lines were composed of dots drawn with a dot pitch of 250–350 μm at a dot rate of 1 MHz. The dot pitch was selected so that dots slightly overlapped, i.e., the dot size (or line width) was of the same magnitude as the dot pitch. Stimuli were refreshed at 200 Hz or else 100 Hz. Luminance of a dot grid (same dot pitch and refresh rate as above) measured with a Minolta LS-100 luminance meter was 80 cd/m^2. The room was dimly illuminated (0.5 lux) and background luminance on the screen was below 1 cd/m^2. Viewing distance was 2 m.

2.2. Observers

Data were obtained from two of the authors and paid graduate students from the University of Bremen, Germany. Before the experiment proper took place, the general purpose of the experiment was explained to every observer. Moreover, subjects were told that they could quit the experiment at any time they wished. After observers signed a consent form, acuity was determined by means of the Freiburg visual acuity test (Bach, 1996). To participate in the experiments subjects had to reach a value of 1.0 (corresponding to 20/20) at least for one eye.

2.3. Stimuli and task

We presented bisection stimuli comprising two outer markers, delineating a horizontal, diagonal, or vertical interval, and a central element positioned between the outer elements. In line bisection, line length was either 1200″ (arc sec) or 600″ (Figs. 1A and D). Dots of dot bisection stimuli were composed of four pixels for each dot (Fig. 1B). Bisection stimuli could also be a combination of dots and lines (Fig. 1C). Stimulus duration was 150 ms. No fixation spot was presented to prevent subjects from judging the position of the center element relative to this fixation spot stored in memory. Each trial started with four markers at the corners of the screen presented for 500 ms followed by a blank screen for 200 ms. The screen was blank for 500 ms between response and subsequent trial. A block of stimulus presentations consisted of 80 trials (except for Experiment 3).

Observers had to discriminate, in a binary forced choice task, whether the central element was closer to the left (or upper) or else to the right (or lower) outer element by pressing one of two buttons. Incorrect responses were followed by an auditory error signal produced by the computer. We determined thresholds of 75% correct responses with an adaptive staircase method and maximum likelihood estimation of the parameters of the psychometric function (PEST; Taylor & Creelman, 1967). In Experiment 3, we used also percentages of correct responses to determine bisection acuity.

Before and after the training of a particular task, we determined *baseline* performance for various bisection stimuli. For every subject, these pre- and post-training conditions were measured twice and the order of conditions was randomized to reduce the influence of learning or fatigue effects. After every condition had been measured once, the order of conditions was reversed for the second set of measurements in order to, at least partly, compensate for possible learning effects in these baselines.

2.4. Data analysis

Data points in graphs showing performance during training correspond to the mean threshold (or mean percentages of correct responses) across subjects for a given training block. Standard errors of the mean (s.e.m.) were computed across subjects. To determine a possible improvement of performance during training, we fitted regression lines to the data of each subject and computed a one sample t test ($\alpha = 0.05$) comparing the slopes of regression lines with the hypothesis of no change due to training, i.e., a slope of zero.

To determine individual pre-training performance, we averaged for each observer the two thresholds of a given baseline condition (for the trained condition in Experiments 1 and 2, these are the first two training blocks). Post-training performance was determined accordingly. Hence, individual baseline thresholds were determined on the basis of 160 trials per observer. To compare performance, we computed the ratio of post-training thresholds to pre-training thresholds for each observer. Data points in baseline graphs correspond to the mean ratios. Standard deviations (s.d.) were computed across subjects. To determine if learning has taken place, we computed one sample t tests ($\alpha = 0.05$) with the hypothesis of no change in performance, i.e., a ratio of one.

3. Results

3.1. Experiment 1: Spatial position uncertainty during training

The eye never stays still and therefore stimuli presented on one constant position at the screen stimulate retinal positions differing from trial to trial. Obviously, perceptual learning is possible under such conditions. However, the exact degree of the variation is unknown. The first prediction of the model states that learning is also possible if the stimulus position can moderately vary on the screen. To test this prediction, we presented a line bisection stimulus with its position randomly chosen in an area of twice the size of the stimulus itself.

3.1.1. Methods

Nine observers trained with line bisection stimuli with an outer line distance of 1200″ (Fig. 1A). The orientation of the lines was either vertical (four observers) or horizontal (five observers) but fixed for each subject. The center of the bisection stimulus appeared in a region from −600″ to 600″ either to the left or right of the center of the screen for vertical and above and beneath for horizontal lines. Consequently, the position of the whole stimulus array was varied (pseudo)-randomly from trial to trial within an area 2400″ wide, twice the size of the entire bisection stimulus. Hence, the leftmost line in one trial could be presented at the same position as the rightmost line in another trial. Subjects were trained in 14 blocks comprising 1120 trials.

Before and after training, we determined performance for the trained line bisection stimulus with its position fixed at the center, for a dot bisection stimulus, and for a line bisection stimulus oriented orthogonally to the trained stimulus. The stimulus position for dot and orthogonal stimuli was randomized from trial to trial as in the trained condition. Outer element distance was 1200″ in all conditions.

3.1.2. Results and discussion

Performance improves with a mean slope of linear regression lines of −1.62 during training (one sample t test, p value: 0.014; Fig. 3A). For seven observers, individual regression lines had negative slopes ranging from −0.71 to −4.11. Performance remained constant for one observer (slope: 0.01) and deteriorated slightly for another one (slope: 0.87).

For baseline conditions, mean ratios of post- to pre-training performance are plotted in Fig. 3B. The ratio for the trained, 'jittered' line bisection stimulus was significantly below one indicating that training had taken place (mean ratio: 0.66; one sample t test, p value: 0.005). We found a similar improvement of performance for the line bisection stimulus with fixed position (mean ratio: 0.73; one sample t test, p value: 0.012). The mean *threshold* difference between this stationary and the "random" line condition was 1.5″ in the pre- and −2.4″ in the post-training measurements. Hence, randomizing the position of the line bisection stimulus yields comparable performance to a stationary line bisection stimulus.

No significant change in performance occurs for the dot bisection stimulus (mean ratio: 1.03). On average, there was a slight improvement of performance for the orthogonal stimulus (mean ratio: 0.91), however, this was non-significant and mainly caused by one observer.

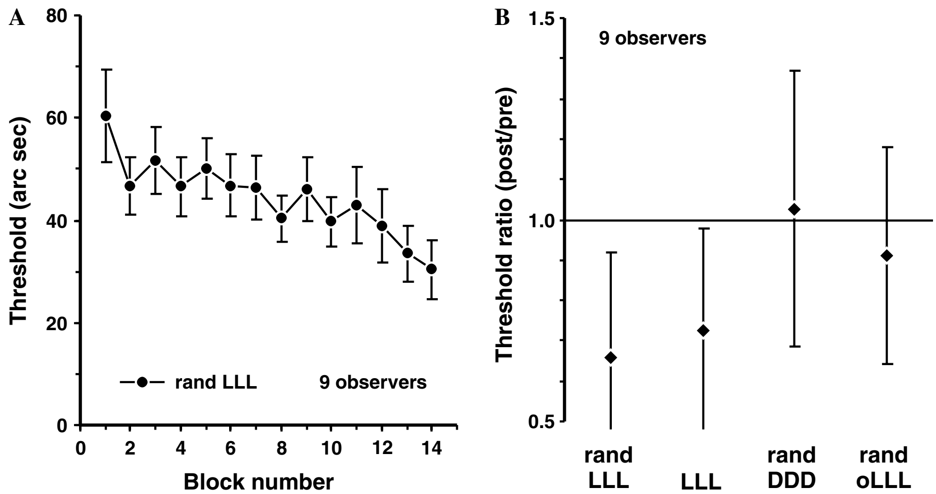

Fig. 3. (A) Bisection acuity as a function of training. Observers trained with line bisection stimuli whose position was shifted laterally by a random amount from trial to trial (rand LLL). The distance between the outer lines was 1200″. Performance improves significantly over blocks (means and s.e.m. of 9 observers). (B) Ratio of post- to pre-training thresholds. Performance significantly below the horizontal line indicates that learning has taken place. In addition to the trained condition (rand LLL), we determined performance for a line stimulus with fixed spatial position (LLL), a dot bisection stimulus with a randomized position (rand DDD), and a line stimulus with randomized position and its orientation orthogonal to the trained line stimulus (rand oLLL). Performance improves significantly for the trained 'jittered' line stimulus (rand LLL) as well as for its constant version (LLL). Transfer of learning occurs neither to the dot nor the orthogonal line stimuli. Means and s.d. of 9 observers. Note that we show s.e.m. for training data but s.d. for the baseline conditions. S.d. are used to provide an impression about statistical significance.

As predicted by the model, the human brain can improve performance with line bisection stimuli even though the positional variation of the bisection stimulus is about a factor of 50 larger than bisection acuity thresholds (see Fig. 2B, "error rates").

Moreover, this experiment clearly demonstrates that perceptual learning occurs under position variant conditions even if subjects would fixate perfectly.

3.2. Experiment 2: Transfer of learning

In the last experiment, the position of the bisection stimulus was varied. Strong transfer of learning occurred to a condition when the stimulus was displayed constantly at the same position in accordance with the first prediction of the model. In the second experiment, observers trained line bisection at a *constant* stimulus position with a line bisection stimulus of an outer line distance of 1200″. In the first part of the experiment, we determined the degree of transfer to a number of line stimuli with different outer line distances. In the second part, we reduced the number of baseline conditions to avoid possible training effects during extensive baseline testings. In both parts, we were particularly interested whether or not negative transfer occurs with one of these distances as the second prediction states.

3.2.1. Methods

Eight observers participated in the first part of the experiment. In the training session, we presented a line bisection stimulus with an outer line distance of 1200″ (Fig. 1A). The lines were oriented either vertically (six observers) or horizontally (two observers). The orientation was constant for each subject. Stimuli were presented always in the center of the screen. Subjects were trained in 14 blocks comprising 1120 trials.

The model predicts that negative transfer occurs for some outer line distances relative to the trained distance. However, the model did not confidently predict quantitatively the exact distances at which this negative transfer occurs. Therefore, we determined pre- and post-training performances for a variety of outer line distances to cover a wide range. We tested distances between 600″ and 1400″ in steps of 200″ for three subjects and between 600″ and 3600″ in steps of 600″ for another three subjects. Two observers were tested over the whole set of distances.

For all observers, we determined pre- and post-training performance for a dot, a dot-line-dot, and a line bisection stimulus with a line length of 600″, i.e., half the length of the trained stimulus (see Figs. 1B–D). The distance between the outer elements in these conditions was 1200″ as in the trained condition.

In the second part of the experiment, five additional observers joined. These observers were trained as in the first part. In order to avoid possible training effects throughout the large number of baseline conditions in the first part of the experiment, baseline performance was determined only with dot bisection with an outer dot distance of 1200″ and line bisection with an outer line distance of 1600″.

3.2.2. Results and discussion

In the first part of the experiment, line bisection improved significantly with a mean slope of linear regression lines of −0.62 during training (one sample t test, p value: 0.017; Fig. 4A). For seven observers, regression lines had negative slopes ranging from −0.38 to −1.66;

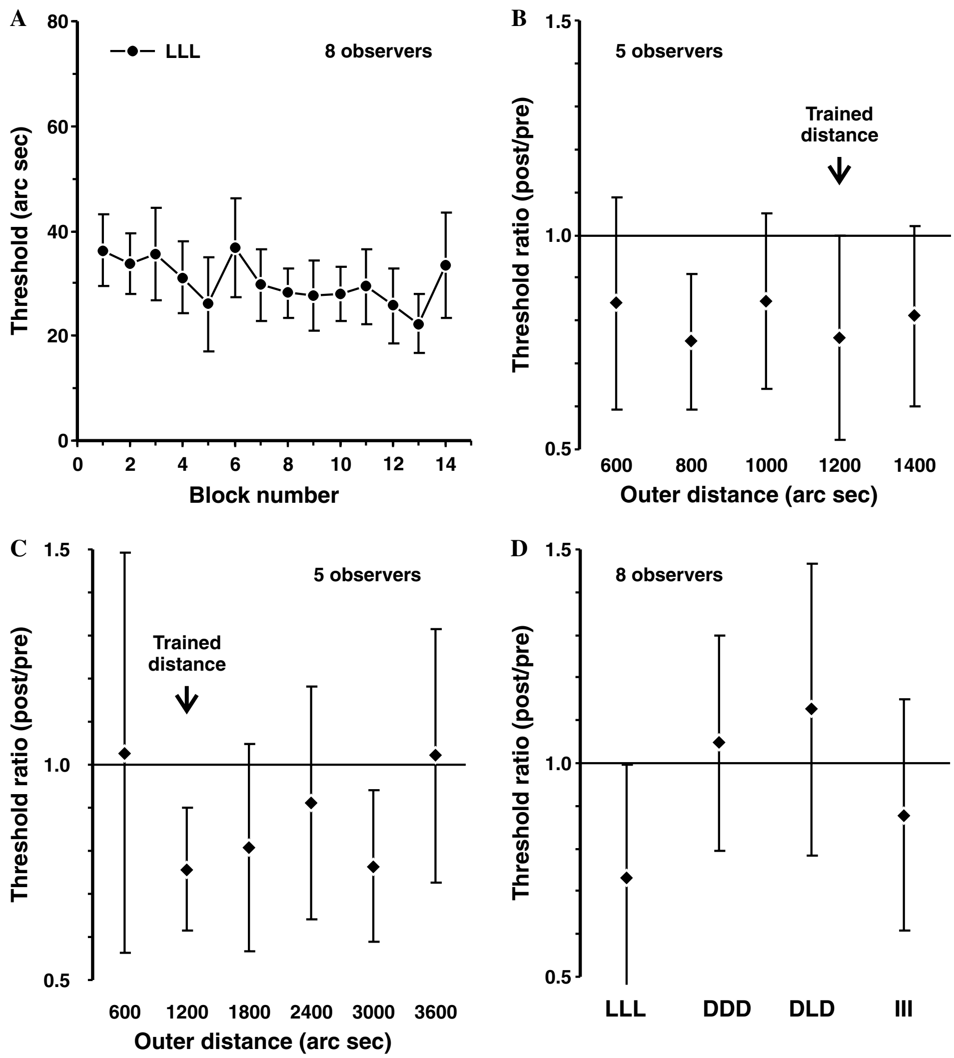

Fig. 4. (A) Bisection acuity as a function of training. Outer line distance was 1200″ (LLL, see Fig. 1A). Performance improves significantly (means and s.e.m. of 8 observers). (B and C), Ratio of post- to pre training thresholds as a function of outer line distance. We determined performance for outer line distances either in steps of 200″ between 600″ and 1400″ (B) or in steps of 600″ between 600″ and 3600″ (C). Bisection acuity seems to improve in some conditions while we find no evidence for negative transfer. Means and s.d. of 5 observers. (D) Ratios of post- to pre-training baseline performance for dot bisection (DDD), a dot-line-dot bisection (DLD), and a line bisection with 600″ long lines (lll, see Figs. 1B–D). Performance improves significantly for the trained stimulus (LLL) but not for dot, dot-line-dot, and short line stimuli (means and s.d. of 8 observers).

performance remained constant for one observer (slope: 0.01). Moreover, the ratio of post- to pre-training performance was significantly below one (Fig. 4D; LLL, mean ratio: 0.73; one sample t test, p value: 0.008).

Mean ratios of post- to pre-training performance for line bisection stimuli with various outer line distances are plotted in Fig. 4B and C. Performance seems to remain constant or to improve slightly. For outer line distances of 800″ and 3000″, the ratio of post- to pre-training performance was significantly below one (post hoc analysis: Fig. 4B, 800″; mean ratio: 0.75; one sample t-test, p value: 0.024; Fig. 4C, 3000″; mean ratio: 0.76; one sample t test, p value: 0.041). In summary, we do not find any indication for negative transfer.

Baseline threshold ratios for dot, dot-line-dot, and small line bisection stimuli are plotted in Fig. 4D. No change in performance occurs. These results show that the tentative improvement of performance cannot be attributed to learning of unspecific factors such as accommodation of the eyes, attentional allocation, or coping with the task in general.

Extensive testing of pre- and post-training conditions in the first part of this experiment may have yielded some improvement of performance and, thus, prevented observers from producing negative transfer. For this reason, five additional observers joined the second part of the experiment in which we tested only two baseline conditions (Fig. 5). Performance seems to improve with a mean slope of linear regression lines of −0.87 during training but the improvement fails to be significant because of one outlier (one sample t test, p value: 0.088). For four observers, regression lines had negative slopes ranging from −0.55 to −1.73 whereas performance deteriorated slightly for one observer (slope: 0.47). The mean threshold ratio was

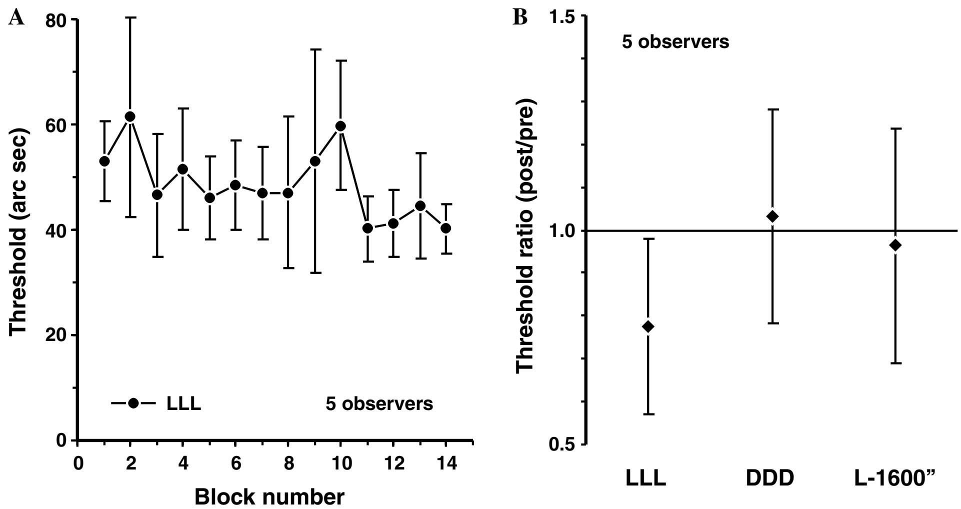

Fig. 5. (A) Bisection acuity as a function of training with a line bisection stimulus with an outer line distance of 1200″ (LLL, see Fig. 1A). Performance seems to improve (means and s.e.m. of 5 observers). (B) Ratios of post- to pre-training baseline performance. We determined bisection acuity only in two baseline conditions. Performance seems to improve for the trained stimulus (LLL) but neither for dot bisection stimuli with the same outer element distance (DDD) nor for the line stimuli with an outer line distance of 1600″ (L-1600″). Means and s.d. of 5 observers.

below one (mean ratio: 0.78), however, improvement failed to be significant due to the outlier as well (one sample *t* test, *p* value: 0.069).

Also in this experiment, there seems to be no evidence for negative transfer of performance since for line bisection with an outer line distance of 1600″ performance slightly improves but at least seems not to deteriorate (mean ratio: 0.96). Hence, we do not find evidence for the prediction of negative transfer neither in the first nor the second part of the experiment.

3.3. Experiment 3: Training with variable stimulus dimensions

In the last experiment, observers trained with one line bisection stimulus. Prediction three states that learning is strongly diminished when bisection stimuli with two different outer line distances are presented randomly interleaved. This prediction is tested here.

3.3.1. Methods

In the first part of the experiment, nine subjects were trained with vertical line bisection stimuli (Fig. 1A) with outer line distances of 1200″ and 1800″ (pseudo)-randomly interleaved between trials. Subjects were informed that either one of these two different line bisection stimuli was presented. Before and after training, we determined performance for each outer line distance separately with the adaptive method (PEST) in two blocks.

For five observers, thresholds for both bisection stimuli were determined by two independent staircase procedures during the training blocks. To achieve reliable thresholds, 60 line bisection stimuli of each outer line distance were presented per block. Observers were trained in 16 blocks, i.e., 1920 trials (plus 320 trials in baseline conditions).

For four additional observers, we determined the percentage of correct responses during the training session. This is to avoid confusion of observers with rapidly changing positions of the central line during two interleaved staircase procedures. Moreover, learning is often faster with constant stimuli since this method avoids presentation of supra-threshold offset values that are often useless for the learning since performance is already perfect for those stimuli (Ahissar & Hochstein, 1997; Herzog & Fahle, 1998). For each observer, offset sizes were chosen individually aiming to achieve a performance level around 65-70% for each of the two bisection stimuli at the beginning of the training session. Offset sizes ranged from 20″ to 40″ at 1200″ and from 30″ to 50″ at 1800″ outer line distance. Observers were trained in 24 blocks each consisting of 40 trials of each outer line distance, i.e., 1920 trials.

In the second part, we tested further pairs of outer line distances to better cover the parameter space. For five observers, we used 1200″ and 2400″ and for another five subjects 800″ and 1400″ as outer line distances. Again, we used the percentage of correct responses to determine performance during the training session. For outer line distances of 1200″ and 2400″, the central line offset sizes ranged from 15″ to 35″ and 31″ to 75″, respectively. For outer line distances of 800″ and 1400″, the central line offset sizes ranged from 10″ to 20″ and 13″ to 25″, respectively. In the case of outer line distances of 1200″ and 2400″, we used diagonal line bisection stimuli (the whole horizontal stimulus array (Fig. 1A) is rotated 45 deg counterclockwise) for three observers and vertical lines for two observers. In the case of 800″ and 1400″ outer line distances, we used horizontal lines. Subjects trained the interleaved tasks in 12 blocks consisting of 40 trials of each outer line distance (960 trials).

Fifteen observers participated in this experiment, and three joined in more than one training condition. For these observers orientation of bisection stimuli was changed. Moreover, there were several weeks between testing different conditions. Results of these observers seem not to differ from those observers participating in one condition only.

3.3.2. Results and discussion

In the first part of the experiment, performance seems not or only weakly to improve during training in the mean of nine observers (Fig. 6A and B). Please note that the amount of training was more than doubled in comparison with Experiments 1 and 2 (1920 versus 800 trials between baselines). In case of an outer line distance of 1200″, two observers showed an initial performance level above the predefined range from 65% to 70%.

Whereas there seems to be no obvious improvement during training, post-training is better than pre-training performance for the 1200″ condition (Fig. 6C, mean ratio: 0.83). However, improvement fails to be significant (one sample t test, p value: 0.094). For the outer line distance of 1800″, performance remains almost unchanged (mean ratio: 0.97).

Thresholds in the baseline conditions were determined for each outer line distance separately whereas presentation was interleaved during training. Still, baseline and training thresholds are in the same range. Therefore, an interleaved presentation seems to disturb learning whereas not performing the task itself.

When outer line distances of 1200″ and 2400″ are interleaved (Fig. 7A and B), performance seems to improve slightly during training (mean slopes: 0.49/0.37). However, baseline performance remains almost unchanged (1200″, mean ratio: 0.90; 2400″, mean ratio: 1.02). When outer line distances of 800″ and 1400″ are interleaved (Fig. 7C and D), slopes are around and often even below zero (mean slopes: −0.15/−0.09), and performance remains almost constant for the baseline conditions (800″, mean ratio: 0.90; 1400″, mean ratio: 0.81).

Interleaved training with two different outer line distances seems to make learning with either distance rather difficult compared to Experiments 1 and 2. It is surprising to find strong learning if the position of the entire stimulus is randomly chosen (Fig. 3) but no or only weak learning when training occurs concurrently with two different outer line distances. It seems that stronger improvement occurs for a stimulus with a constant outer line distance in spite of random overall position than for stimuli with varying outer line distances in spite of almost constant position of the middle element.

The model suggests that different, even conflicting, recurrent networks are needed if bisection stimuli with more than one outer line distance have to be trained. Since perceptual learning involves learning the optimal recurrent network for each condition, one might expect that interleaving the conditions during learning makes it impossible to learn either network and, thus, no or diminished improvement of performance should result in either condition. This is what we found.

4. Discussion

4.1. Learning under uncertainty in spatial positions

In Experiment 1, we presented the line bisection stimulus with a substantial spatial jitter during the training

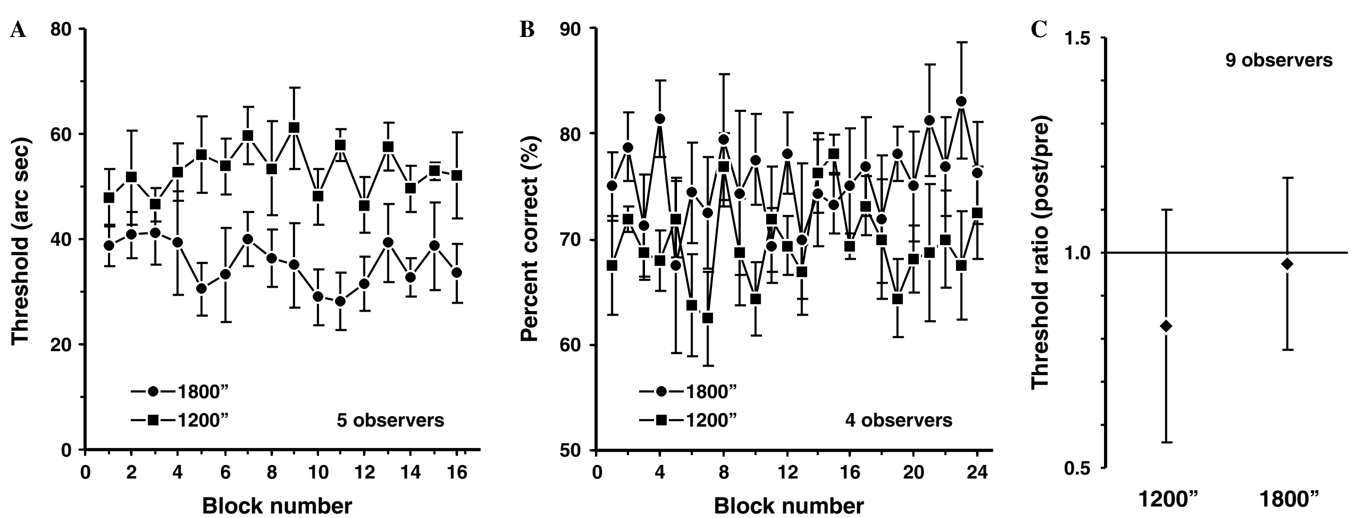

Fig. 6. (A and B) Bisection performance as a function of training. Line bisection stimuli with two different outer line distances (1200″ and 1800″) were presented randomly interleaved during training. For five observers (A), we determined thresholds and, for another four observers (B), percentages of correct responses. We presented 960 trials in each condition, i.e., together with baseline measurements more trials per condition than in Experiments 1 and 2. Improvement of performance is weak compared to Experiments 1 and 2, if present at all. (C) Ratios of post- to pre-training baseline performance. For all nine observers, we determined thresholds for both trained outer line distances separately in additional blocks. In both cases, mean ratios do not significantly differ from one. However, there might be some improvement for the 1200″ condition.

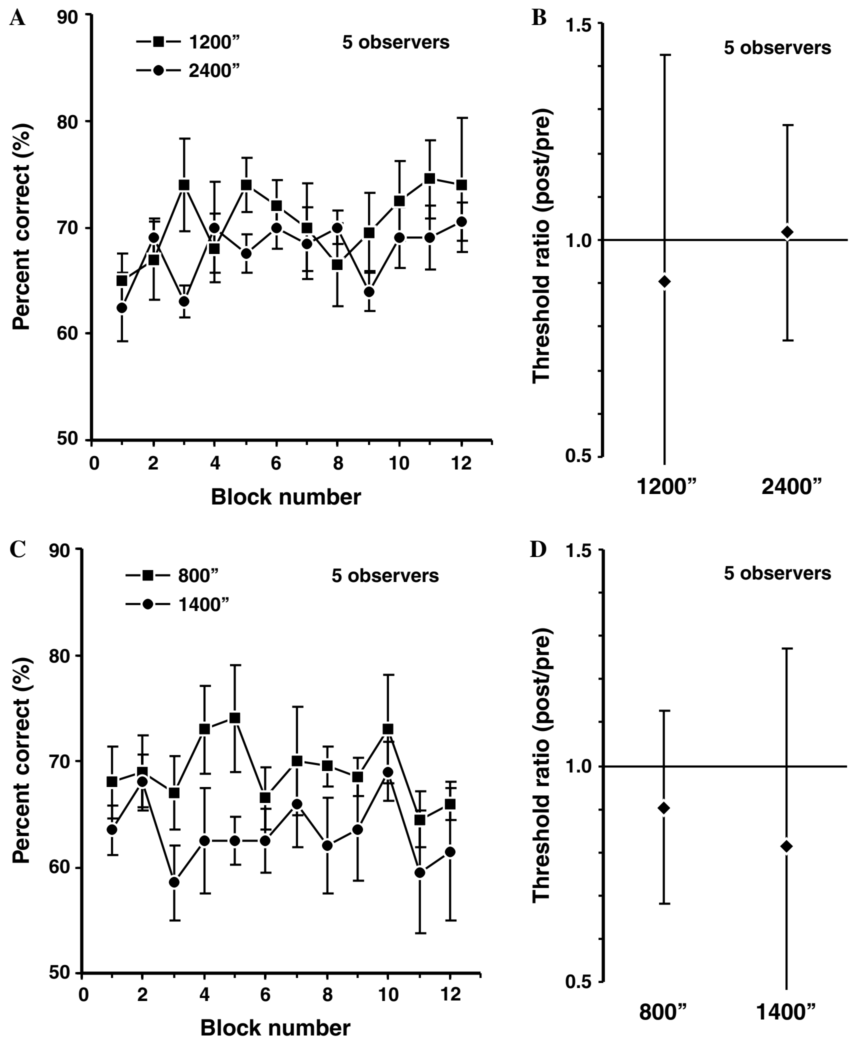

Fig. 7. (A) Bisection performance as a function of training. For five observers, line bisection stimuli with outer line distances of 1200″ and 2400″ were randomly interleaved during training (960 trials). We used the percentage of correct responses measure to determine performance. For both outer line distances, performance seems to improve slightly. (B) Ratios of post- to pre-training baseline performance. We determined thresholds for both trained outer line distances separately. There seems to be no improvement of performance for both outer line distances. (C and D) We repeated the experiment with outer line distances of 800″ and 1400″ for five new observers. Again, we found no significant training effects.

session. Strong learning (Fig. 3A) and positive transfer to a centrally presented line stimulus with a constant spatial position occurs (Fig. 3B). Moreover, varying the stimulus position seems not to make learning less effective since the ratio of post- to pre baseline performance is comparable regardless of whether subjects were trained in conditions with randomized or fixed stimulus position (Figs. 3B, 4D, and 5B). These results are in good accordance with the model prediction. Since our model does not specify the learning algorithm for the recurrent weights, it does not predict whether learning should be faster with or without varying the stimulus positions, other than that they should both be effective.

Line bisection learning does not transfer to dot bisection or to an orthogonal line stimulus. Hence, learning is specific for the stimulus orientation and type but not for its exact position during learning. Westheimer, Crist, Gorski, and Gilbert (2001) found a pronounced transfer from line to dot bisection stimuli. The difference in transfer between their and our study might be explained by the fact that we trained our observers foveally while Westheimer et al. (2001) trained their subjects at 3.5 deg peripherally.

In Crist et al. (1997), bisection was trained at one peripheral position and pre- and post-training conditions were determined at other peripheral positions for the same line bisection stimulus. Results show some position specificity as well as a strong component of transfer across positional changes up to 8 deg (these results are in contrast to findings with vernier acuity in which no transfer between peripheral positions was found when observers had previously been trained to attend to the periphery; Fahle et al., 1995).

It is surprising that bisection learning is very unspecific for the exact position of training but very specific for the exact stimulus type. Hence, a fine tuning of neurons dedicated to spatial separation analysis at one spatial location only is very unlikely the cause for perceptual learning. On the other hand, high level unspecific learning related to keeping the eyes fixed, improving accommodation, learning the motor responses, or adjusting overall decision criteria is unlikely either for the same reason. The specificity for the exact type of stimulus, moreover, rules out the possibility that improvement is based only on unspecific recruitment of neural resources as it is proposed to occur, for example, with training in the auditory system (Recanzone, Schreiner, & Merzenich, 1993). Our experiments indicate the strong need for incorporating a "form" analyzing component involved in the learning of spatial separation as proposed by Westheimer et al. (2001).

Our results are in good agrement with those of other learning experiments in which the position of stimuli between trials was permanently varied. For example, Sireteanu and Rettenbach (1995, 2000) showed that in visual search experiments serial processing becomes parallel while the position of target and distractors varied.

In our experiments, we focus on the positional variation from trial to trial rather than on a positional jitter occurring *during* one trial caused by tremor or microsaccades that may improve performance (e.g., Greschner, Bongard, Rujan, & Ammermuller, 2002; Hennig, Kerscher, Funke, & Wörgötter, 2002).

Our results indicate that perceptual learning with bisection stimuli is accomplished after at least some spatial invariance is reached, and our model (Zhaoping et al., 2003) suggests that modifications of lateral connections in V1 may be partly responsible (Crist et al., 2001; Furmanski, Schluppeck, & Engel, 2004; Ghose, Yang, & Maunsell, 2002; Schoups, Vogels, Qian, & Orban, 2001; Schwartz, Maquet, & Frith, 2003; Skrandies & Fahle, 1994; but see Mollon & Danilova, 1996).

4.2. Negative transfer between conditions of different dimensions

The model was not definitive concerning at which outer line distance negative transfer would occur for line bisection stimuli. Therefore, we tested a large range of outer line distances. Bisection learning for a line bisection stimulus with an outer line distance of 1200″ seems to transfer to outer line distances from 600″ up to 3000″ (Fig. 4B and C). Our data, averaged over the subjects, indicate that no negative transfer occurred. This holds regardless of whether pre-training is rather extensive or restricted to two conditions (Fig. 5B). Still, we cannot exclude that negative transfer may be found in some conditions not tested. Moreover, negative transfer might occur for different observers at different outer line distances and, therefore, may be obliterated in the averaged data. Also, negative transfer might correlate with the strength of learning. Indeed, a post-hoc analysis of data presented in Fig. 4C showed that among the four subjects who improved in the trained outer-distance condition, the strongest negative transfer was observed for the two subjects who improved most in the trained condition. However, more subjects are needed for a more sophisticated analysis.

Crist et al. (1997) found strong transfer of improvement when doubling the outer line distance from a trained to an untrained line distance in an extended nine week training paradigm with stimuli presented 5 deg peripherally. The authors determined pre- and post-learning performance only for this untrained outer line distance and, hence, did not address negative transfer.

4.3. No learning for variable stimulus dimensions

In Experiment 3, learning is strongly diminished when line bisection stimuli with two outer line distances are presented randomly interleaved according to the third model prediction.

These results are in good agreement with the findings of Adini, Sagi, and Tsodyks (2002), Adini, Wilkonsky, Haspel, Tsodyks, and Sagi (2004), Yu, Klein, and Levi (2004). In these studies, no learning occurs in a contrast discrimination task if the reference contrasts are unpredictably interleaved (contrast roving). Yu et al. (2004) explain their results as evidence for an improvement of templates whereas our computer simulations favor an explanation in terms of recurrent connections (Zhaoping et al., 2003).

There seems to be some positive but at least no negative transfer from one to another outer line distance (Fig. 4) while learning is strongly diminished when two outer line distances were trained simultaneously (Figs. 6 and 7). This might be explained by the different temporal ordering of stimuli in the experiments. Kuai, Zhang, Klein, Levi, and Yu (2005) showed that a fixed temporal stimulus patterning can enable perceptual learning for contrast and motion direction discrimination, which are "unlearnable" if different stimulus alternatives are randomly interleaved (see also Liu & Vaina, 1998).

4.4. Summary

Perceptual learning requires the ability of the human brain to cope with positional uncertainties. Indeed, this was experimentally found (Fig. 3). Mathematically, coping with spatial uncertainties changes the decision problem from linear to non-linear (Zhaoping et al., 2003). A non-linear model suggests that learning two outer line distances simultaneously would require the creation, on the same neural population, of two set of recurrent connections, one for each outer line distance. Since these two sets of connections are likely to conflict with each other, learning would be difficult or impossible as found in our data (Figs. 6 and 7). However, the prediction of negative transfer, implied by this model, could not be verified (Figs. 4 and 5). Hence, while non-linear recurrent networks may achieve positional invariance, the brain is able to prevent negative transfer by learning mechanisms not considered by our model.

Acknowledgments

Thomas Otto was supported by the Swiss National Science Foundation. Li Zhaoping was supported by the Gatsby Charitable Foundation. Manfred Fahle and Michael Herzog were supported by the German Research Council (SFB 517). We thank Peter Dayan for useful discussions arising from a previous collaborative project modelling the computations in perceptual learning under positional uncertainties. We thank Walter Senn for comments, Marc Repnow for excellent technical support, and Ingo Enskat for help in the experiments.

References

Adini, Y., Sagi, D., & Tsodyks, M. (2002). Context-enabled learning in the human visual system. *Nature, 415*, 790–793.

Adini, Y., Wilkonsky, A., Haspel, R., Tsodyks, M., & Sagi, D. (2004). Perceptual learning in contrast discrimination: the effect of contrast uncertainty. *Journal of Vision, 4*(12), 993–1005.

Ahissar, M., & Hochstein, S. (1997). Task difficulty and the specificity of perceptual learning. *Nature, 387*(6631), 401–406.

Bach, M. (1996). The "Freiburg Visual Acuity Test." Automatic measurement of visual acuity. *Optometry and Vision Science, 73*(1), 49–53.

Crist, R. E., Kapadia, M. K., Westheimer, G., & Gilbert, C. D. (1997). Perceptual learning of spatial localization: specificity for orientation, position, and context. *Journal of Neurophysiology, 78*(6), 2889–2894.

Crist, R. E., Li, W., & Gilbert, C. D. (2001). Learning to see: experience and attention in primary visual cortex. *Nature Neuroscience, 4*(5), 519–525.

Fahle, M., Edelman, S., & Poggio, T. (1995). Fast perceptual learning in visual hyperacuity. *Vision Research, 35*(21), 3003–3013.

Fahle, M., & Poggio, T. (Eds.). (2003). *Perceptual learning*. MIT Press.

Fiorentini, A., & Berardi, N. (1980). Perceptual learning specific for orientation and spatial frequency. *Nature, 287*(5777), 43–44.

Furmanski, C. S., Schluppeck, D., & Engel, S. A. (2004). Learning strengthens the response of primary visual cortex to simple patterns. *Current Biology, 14*(7), 573–578.

Gilbert, C. D., & Wiesel, T. N. (1983). Clustered intrinsic connections in cat visual cortex. *Journal of Neuroscience, 3*(5), 1116–1133.

Ghose, G. M., Yang, T. M., & Maunsell, J. H. R. (2002). Physiological correlates of perceptual learning in monkey V1 and V2. *Journal of Neurophysiology, 87*(4), 1867–1888.

Greschner, M., Bongard, M., Rujan, P., & Ammermuller, J. (2002). Retinal ganglion cell synchronization by fixational eye movements improves feature estimation. *Nature of Neuroscience, 5*(4), 341–347.

Hennig, M. H., Kerscher, N. J., Funke, K., & Wörgötter, F. (2002). Stochastic resonance in visual cortical neurons: does the eye-tremor actually improve visual acuity? *Neurocomputing, 44–46C*, 115–120.

Herzog, M. H., & Fahle, M. (1998). Modelling perceptual learning: difficulties and how they can be overcome. *Biological Cybernetics, 78*, 107–117.

Karni, A., & Sagi, D. (1991). Where practice makes perfect in texture discrimination: evidence for primary visual cortex plasticity. *Proceedings of the National Academy of the Science United States of America, 88*(11), 4966–4970.

Kuai, S. G., Zhang, J. Y., Klein, S. A., Levi, D. M., & Yu, C. (2005). The essential role of stimulus temporal patterning in enabling perceptual learning. *Nature Neuroscience, 8*(11), 1497–1499.

Liu, Z., & Vaina, L. M. (1998). Simultaneous learning of motion discrimination in two directions. *Cognitive Brain Research, 6*(4), 347–349.

Martinez-Conde, S., Macknik, S. L., & Hubel, D. H. (2004). The role of fixational eye movements in visual perception. *Nature Reviews Neuroscience, 5*(3), 229–240.

McKee, S. P., & Westheimer, G. (1978). Improvement in vernier acuity with practice. *Perception & Psychophysics, 24*(3), 258–262.

Mollon, J. D., & Danilova, M. V. (1996). Three remarks on perceptual learning. *Spatial vision, 10*(1), 51–58.

Poggio, T., Fahle, M., & Edelman, S. (1992). Fast Learning in visual Hyperacuity. *Science, 256*(5059), 1018–1021.

Ramachandran, V. S., & Braddick, O. (1973). Orientation-specific learning in stereopsis. *Perception, 2*(3), 371–376.

Recanzone, G. H., Schreiner, C. E., & Merzenich, M. M. (1993). Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *Journal of Neuroscience, 13*(1), 87–103.

Rockland, K. S., & Lund, J. S. (1983). Intrinsic laminar lattice connections in primate visual cortex. *Journal of Comparative Neurology, 216*, 303–318.

Schoups, A. A., Vogels, R., & Orban, G. A. (1995). Human perceptual learning in identifying the oblique orientation: retinotopy, orientation specificity and monocularity. *Journal of Physiology, 483.3*, 797–810.

Schoups, A., Vogels, R., Qian, N., & Orban, G. (2001). Practising orientation identification improves orientation coding in V1 neurons. *Nature, 412*(6846), 549–553.

Schwartz, S., Maquet, P., & Frith, C. (2003). Neural correlates of perceptual learning: a functional MRI study of visual texture discrimination. *Proceedings of the National Academy of Sciences United States of America, 99*(26), 17137–17142.

Shiu, L., & Pashler, H. (1992). Improvement in line orientation discrimination is retinally local but dependent on cognitive set. *Perception & Psychophysics, 52*(5), 582–588.

Sireteanu, R., & Rettenbach, R. (1995). Perceptual learning in visual search—fast, enduring, but nonspecific. *Vision Research, 35*(14), 2037–2043.

Sireteanu, R., & Rettenbach, R. (2000). Perceptual learning in visual search generalizes over tasks, locations, and eyes. *Vision Research, 40*(21), 2925–2949.

Skrandies, M., & Fahle, M. (1994). Neurophysiological correlates of perceptual learning in the human brain. *Brain Topography, 7*(2), 163–168.

Taylor, M. M., & Creelman, C. D. (1967). PEST: efficient estimates on probability functions. *The Journal of the Acoustical Society of America, 41*, 782–787.

Tsodyks, M., & Gilbert, C. (2004). Neural networks and perceptual learning. *Nature, 431*(7010), 775–781.

Vaina, L. M., Sundareswaran, V., & Harris, J. G. (1995). Learning to ignore: psychophysics and computational modeling of fast learning of direction in noisy motion stimuli. *Cognitive Brain Research, 2*(3), 155–163.

Vogels, R., & Orban, G. A. (1985). The effect of practice on the oblique effect in line orientation judgments. *Vision Research, 25*(11), 1679–1687.

Watanabe, T., Nanez, J. E., & Sasaki, Y. (2001). Perceptual learning without perception. *Nature, 413*, 844–848.

Westheimer, G., Crist, R. E., Gorski, L., & Gilbert, C. D. (2001). Configuration specificity in bisection acuity. *Vision Research, 41*(9), 1133–1138.

Yu, C., Klein, S. A., & Levi, D. M. (2004). Perceptual learning in contrast discrimination and the (minimal) role of context. *Journal of Vision, 4*(3), 169–182.

Zhaoping, L., Herzog, M. H., & Dayan, P. (2003). Quadratic Ideal Observation and recurrent Preprocessing in perceptual learning. *Network, 14*, 233–247.