

# Differences in perceptual learning transfer as a function of training task

C. Shawn Green

Department of Psychology, University of Wisconsin–Madison,  
Madison, WI, USA



Florian Kattner

Department of Psychology, University of Wisconsin–Madison,  
Madison, WI, USA



Max H. Siegel

Department of Brain and Cognitive Sciences,  
Massachusetts Institute of Technology,  
Cambridge, MA, USA



Daniel Kersten

Department of Psychology, University of Minnesota,  
Minneapolis, MN, USA



Paul R. Schrater

Department of Psychology, University of Minnesota,  
Minneapolis, MN, USA



A growing body of research—including results from behavioral psychology, human structural and functional imaging, single-cell recordings in nonhuman primates, and computational modeling—suggests that perceptual learning effects are best understood as a change in the ability of higher-level integration or association areas to read out sensory information in the service of particular decisions. Work in this vein has argued that, depending on the training experience, the “rules” for this read-out can either be applicable to new contexts (thus engendering learning generalization) or can apply only to the exact training context (thus resulting in learning specificity). Here we contrast learning tasks designed to promote either stimulus-specific or stimulus-general rules. Specifically, we compare learning transfer across visual orientation following training on three different tasks: an orientation categorization task (which permits an orientation-specific learning solution), an orientation estimation task (which requires an orientation-general learning solution), and an orientation categorization task in which the relevant category boundary shifts on every trial (which lies somewhere between the two tasks above). While the simple orientation-categorization training task resulted in orientation-specific learning, the estimation and moving categorization tasks resulted in significant orientation learning generalization. The general framework tested here—that task specificity or generality can be predicted via an examination of the optimal learning solution—may be useful in building future training paradigms with certain desired outcomes.

## Introduction

For nearly as long as there has been dedicated study of human learning, there has been interest in conditions wherein dedicated practice results in enhanced performance only on the trained task (i.e., *specificity*), and conditions wherein practice results in benefits not only on the trained task, but on some number of new untrained tasks as well (i.e., *transfer* or *generalization* [Thorndike & Woodworth, 1901]). While instances of both types of outcome have been documented throughout the field of psychology, for example in education (Barnett & Ceci, 2002) and cognitive training (Melby-Lervag & Hulme, 2013), the extent of learning specificity in the field of perceptual learning is particularly striking (Sagi, 2011). For instance, subjects trained to discriminate whether a line segment is tilted clockwise or counterclockwise around one reference angle may show substantial improvements in making this particular decision, but then show no enhancement in the ability to make the same discrimination around a new reference angle (Crist, Kapadia, Westheimer, & Gilbert, 1997). Similarly, participants trained to detect a texture patch presented in one part of the visual field may show no transfer of learning when the texture patch is moved to a new location (Ahissar & Hochstein, 1997). In all, specificity of perceptual learning has been

Citation: Green, C. S., Kattner, F., Siegel, M. H., Kersten, D., & Schrater, P. R. (2015). Differences in perceptual learning transfer as a function of training task. *Journal of Vision*, 15(10):5, 1–14. doi:10.1167/15.10.5.

noted under some training conditions for myriad low-level visual features and training characteristics including orientation and spatial frequency (Fiorentini & Berardi, 1980), texture (Karni & Sagi, 1991), retinal position (Fahle, Edelman, & Poggio, 1995), motion direction (Ball & Sekuler, 1982), motion speed (Saffell & Matthews, 2003), and even the trained eye (Fahle, 2004).

However, while there are many examples wherein perceptual training results in highly specific learning, a large and growing body of research strongly confirms that specificity is not a necessary aspect of perceptual learning. For example, there are now many cases wherein training on complex tasks—often corresponding to “real-life” activities such as aerobic activity (Hillman, Erickson, & Kramer, 2008), participation in sports (Mann, Williams, Ward, & Janelle, 2007), meditation (Davidson & McEwen, 2012), music training (Schellenberg, 2004), or playing action video games (Green & Bavelier, 2012)—result in significant learning transfer. There are also many instances where significant transfer of learning has been observed using more standard perceptual learning stimuli. For example, in seminal work by Ahissar and Hochstein (1997), training on an odd-element detection task resulted in either orientation-specific or orientation-general learning, depending on the nature of the training stimuli. Training with odd elements that differed greatly from the background elements (e.g., by 90°) or where odd elements could appear only in a restricted set of locations (e.g., two possible locations) resulted in significant generalization of learning. Conversely, training with odd elements that differed less strongly from the background elements (e.g., by 16°) or in a wider set of positions resulted in much more specific learning.

Significant transfer has also been observed in a double training procedure, wherein individuals are trained on one feature dimension (such as stimulus contrast) in one location, a second feature dimension (such as stimulus orientation) in a second location, and then tested on the initial feature dimension in the second location. Given this design, behavior typically reflects full transfer of the learning that occurred in the initial feature training to the new location (Wang, Zhang, Klein, Levi, & Yu, 2012, 2014; Xiao et al., 2008; Zhang, Cong, Song, & Yu, 2013; Zhang, Xiao, Klein, Levi, & Yu, 2010). Other work in this domain has outlined, as just a few examples, the role of training task difficulty (Garcia, Kuai, & Kourtzi, 2013; Liu, 1999), stimulus complexity (McGovern, Webb, & Peirce, 2012), training time (Jeter, Doshier, Liu, & Lu, 2010), and characteristics of the transfer tasks (Jeter, Doshier, Petrov, & Lu, 2009) in determining the specificity or generality of learning. In addition to the clear theoretical importance of these results, techniques

to induce learning transfer have obvious real-world relevance in rehabilitation paradigms for cortically based visual disorders such as amblyopia (Li, Ngo, Nguyen, & Levi, 2011; Polat, Ma-Naim, & Spierer, 2009; Zhang, Cong, Klein, Levi, & Yu, 2014), cortical blindness (Das, Tadin, & Huxlin, 2014), or age-related declines in vision (DeLoss, Watanabe, & Anderson, 2015), and in improving vision in individuals with normal vision whose jobs or activities involve significant visual demands (Deveau, Ozer, & Seitz, 2014; Schlickum, Hedman, Enochsson, Kjellin, & Fellander-Tsai, 2009).

In one highly influential framework, perceptual learning has been thought to reflect changes in receptive field properties of sensory areas, with the degree of behavioral learning specificity corresponding to the receptive field specificity in the neural locus of learning, which in turn is determined by the signal-to-noise ratio demanded by the training task (Ahissar & Hochstein, 2004; Ahissar, Nahum, Nelken, & Hochstein, 2009). However, an alternative framework suggests that perceptual learning instead reflects improvements in the ability of higher-level areas to read out task-relevant sensory information in the service of given decisions (Bavelier, Green, Pouget, & Schrater, 2012; Bejjanki, Beck, Lu, & Pouget, 2011; Doshier & Lu, 1998; Kahnt, Grueschow, Speck, & Haynes, 2011; Law & Gold, 2008; Lu & Doshier, 2009; Zhang, Zhang et al., 2010). For instance, the results seen in double training studies are consistent with a model wherein higher-level decision units learn a set of rules for weighting inputs from low-level areas. Appropriate training or experience then allows for functional connections to be formed that support the application of these learned rules to untrained contexts (e.g., new locations).

In our own work, we have contrasted learning solutions/rules that can be formulated as a mapping between stimulus attributes and responses (termed *policy learning* in the language of machine learning) and solutions that utilize predictive methods and look ahead. This distinction between learning simple mappings and learning more predictive methods in perception is similar to the division that exists between model-free and model-based solutions in the reinforcement learning literature (Dayan & Daw, 2008; Glascher, Daw, Dayan, & O’Doherty, 2010). Critically, the former type of rule will tend to be more task and stimulus specific than the latter, in that the mapping that is learned in the former case is necessarily dependent on the exact stimuli and responses (Bavelier et al., 2012; Fulvio, Green, & Schrater, 2014). In our framework, the extent to which one or the other type of rule is learned depends on the demands of the response policies impossible to learn (e.g., estimation tasks that do not allow one to learn a single discriminatory

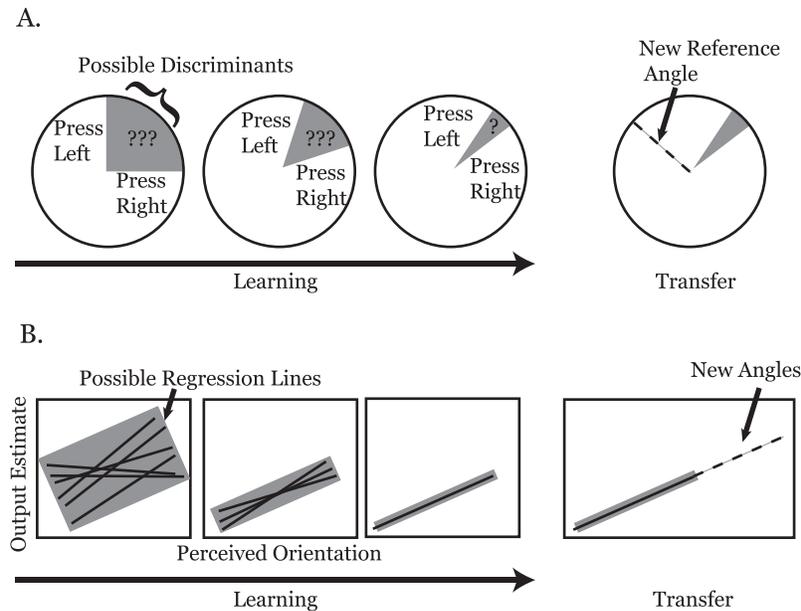


Figure 1. Learning solutions and their effect on transfer. (A) In learning a simple discriminative mapping, there is initial uncertainty about where the boundary lies (gray region), the extent of which is reduced over time. However, this discriminative mapping is of no value at an orthogonal orientation. (B) In learning a continuous relationship between perceived orientation and output estimate, there is initial uncertainty regarding the slope of the relationship (i.e., many possible lines that are consistent with the data, here represented as many individual black lines within the gray region indicative of overall uncertainty). As data are observed over time, the degree of uncertainty is reduced (e.g., the space of possible lines is narrowed to include only the true relationship). Finally, because the relationship is continuous with orientation, the learning is applicable to the orthogonal orientation, represented as an extrapolation to a new angle in the final panel.

feature) or in which learning response policies would require an inordinate number of trials (e.g., where there is an extremely large or indeterminate space of categories), subjects should utilize predictive methods (Fulvio et al., 2014). Conversely, if the task demands make learning policies reasonably simple, subjects should eschew the computational costs of prediction in favor of exploiting the simple mappings. We have recently tested this framework in a trajectory extrapolation learning task (Fulvio et al., 2014). Throughout this training, participants were presented either with a small number of distinct trajectories ( $N = 4$ ) to extrapolate or a large number of trajectories ( $N = 20$ ) to extrapolate. A small number of distinct trajectories favors the learning of distinct stimulus–response mappings in ideal learning agents and, consistent with this, significant behavioral learning specificity was observed in the human learners. Conversely, a large number of distinct mappings makes learning stimulus–response mappings difficult in ideal learning agents and thus, a more predictive approach is favored. This outcome was also mirrored in the behavioral results.

In the work above, we biased the learning solution by altering stimulus characteristics, but it should also be possible to see similar results when holding the stimulus characteristics constant across training groups, and manipulating the task in such a way to

favor a mapping versus a predictive solution. For instance, consider an orientation discrimination task wherein the subject is, on each trial, presented with a Gabor patch with the orientation drawn from a uniform distribution ( $30^\circ$ – $60^\circ$ ). The subject is instructed to press left if the Gabor is oriented counterclockwise from a reference angle ( $45^\circ$ ) and otherwise press right. The optimal solution for this task is to learn a mapping that divides the orientation space into two decision regions (clockwise and counterclockwise). While this will result in efficient performance on the trained task, the improvements will be highly orientation specific, as the learned mapping is of little use if the reference orientation is changed (see Figure 1A).

Conversely, in an angle estimation task in which the stimuli are identical as above, the participant must estimate the angle of the presented stimulus. In contrast to categorizing the stimuli as clockwise or counterclockwise, what needs to be learned in this case is a regression line mapping stimulus attributes relevant to orientation to an estimate of the angle (i.e., a function that translates information available to the visual system into a motor output). Significant orientation transfer would be predicted given this learning solution, as extrapolation to previously unseen orientations is a natural consequence of the constant relationship

between input and output across the continuous circular space of orientations (see Figure 1B).

Here we contrast the degree of orientation learning specificity that arises from training on three different tasks: (a) an orientation categorization task, (b) an orientation estimation task, and (c) a hybrid task, wherein the participant responses are categorical (clockwise or counterclockwise), but the critical boundary is changed on every trial (i.e., the reference angle changes on every trial). We hypothesized that: (a) consistent with previous work, significant orientation learning specificity would be observed in the categorization task; (b) significant, if not complete orientation learning transfer would be observed in the estimation task; and (c) that an intermediate degree of orientation learning transfer would be observed in the moving discriminant categorization task. We further hypothesized that little transfer from task to task would be observed. Indeed, the simple mapping that we hypothesize will be learned from a categorization task obviously provides no benefit in an estimation task. Similarly, in the estimation task, even if a participant has learned to estimate the angle of a line, this, in and of itself, does not provide information relevant to category membership (i.e., one may be able to match a presented angle, but not the discriminant line that determines whether it belongs to category A or category B).

To test our hypotheses, individuals underwent 3,800 trials of training across 4 days on one of the three tasks above. In order to ensure that any differences in transfer could not be attributed to differences in visual experience, all subjects viewed the exact same stimuli across training. Finally, after training was completed, both orientation and task transfer were assessed.

Consistent with our hypotheses, and as has been repeatedly observed (Sagi, 2011), no significant orientation transfer was observed in the participants trained on the categorization task. In contrast, significant orientation transfer was observed in the estimation and moving categorization trained groups.

## Methods

### Participants

Twenty healthy young adults ( $M_{\text{age}} = 19.7$ ) with normal or corrected-to-normal vision were recruited for the study. Observers were randomly assigned to either the categorization training group ( $n = 8$ ; five female, three male), the estimation training group ( $n = 7$ ; four female, three male), or the moving-categorization training group ( $n = 5$ ; three female, two male).

### Apparatus and display

The stimuli were generated on a Dell XPS computer running Windows XP using MATLAB (MathWorks, Natick, MA) and the Psychophysics Toolbox (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007; Pelli, 1997). The stimuli were displayed on a 20-in. Dell LCD monitor at a resolution of  $1680 \times 1050$  pixels by a NVIDIA GeForce 8800 GTX video card. Subjects were seated 59 cm from the screen.

### Stimuli and task

The stimuli were identical for all three groups (see Figure 2). The stimulus display consisted of a small centrally presented “T” and a peripherally presented full contrast Gabor (presented  $10^\circ$  below the T stimulus directly along the vertical axis). The central T was presented either right-side up or upside down, while the orientation of the Gabor was drawn from a uniform distribution (i.e.,  $30^\circ$ – $60^\circ$  during training, or  $120^\circ$ – $150^\circ$  at some test stages; see below).

Each trial started with the presentation of a central fixation cross for 750 ms. Following the brief presentation (130 ms) of the stimulus display, a white noise mask was presented in the location of the Gabor for 300 ms. In all versions of the task, following the presentation of the mask, participants were first required to indicate the orientation of the central T (by hitting either the “w” or “s” key for right-side up and upside down, respectively). Then, in the categorization task, observers were asked to indicate whether the orientation of the Gabor was clockwise or counterclockwise from a constant reference angle (e.g.,  $45^\circ$  during training; an oriented line was shown during the choice period as a reference). In the estimation task, observers used the mouse to orient a line so as to match the orientation of the Gabor. Finally, the moving categorization task was identical to the categorization task, but the orientation of the reference line (i.e., the discriminatory orientation) was drawn from a uniform distribution between  $30^\circ$  and  $60^\circ$  on each trial during training, again with an oriented line being shown during the choice period as a reference. In all three tasks, participants had unlimited time to respond; consequently, average response times and the total duration of the experiment were longer in the estimation group than in the categorization and moving categorization group. There was a 1500 ms intertrial interval starting immediately after the response. Feedback was presented during the intertrial interval during training but not during pretest or posttest in all three tasks (complete feedback showing both whether the answers were correct or incorrect,

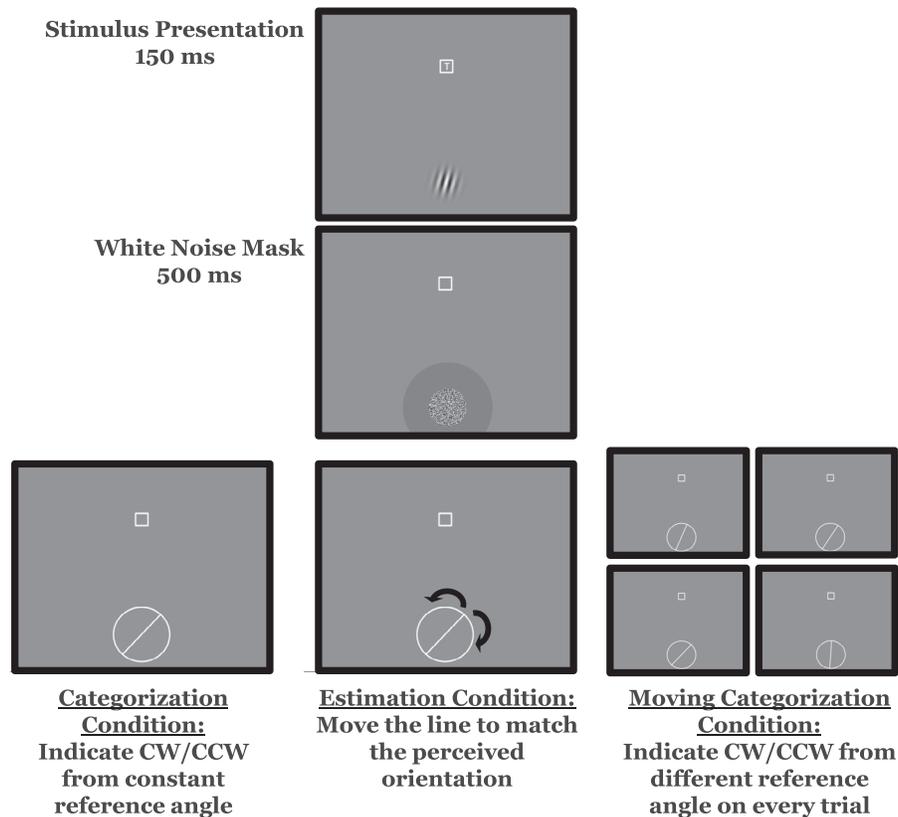


Figure 2. Typical trial. In all three conditions participants see the same stimuli. A small “T” presented right-side up or upside down at fixation and a peripheral Gabor presented  $10^\circ$  directly below fixation. This is followed by a white noise mask at the Gabor location. Participants in all three conditions are next asked to indicate the orientation of the central T (not pictured). Then, in the categorization condition (bottom left), participants indicate whether the presented Gabor was tilted clockwise or counterclockwise from the constant reference angle. In the estimation condition (bottom center), the participants use the mouse to change the orientation of the central line so as to match the orientation of the presented Gabor. In the moving categorization condition (bottom right), the participants indicate whether the presented Gabor was tilted clockwise or counterclockwise from the presented reference angle, which changes on every trial.

as indicated by a green or red outline around the central T square and around the peripheral discrimination circle, and with the actual stimuli that were presented overlaid on top).

## Procedure

The entire experiment consisted of a pretest session, 19 blocks (200 trials each) of training, and a posttest session, which, together, took place across five separate days (over a span of no more than nine total days). Day 1 of the experiment included a short block of trials run at a reduced speed for the participants to familiarize themselves with the dual-task procedure, a pretest (see below), and the first block of training. Days 2 through 4 consisted of six blocks of training on the respective training condition. Day 5 consisted only of the posttest (see below).

## Pre- and posttest assessments

The pretest session consisted of two 200-trial blocks, one block with the same task as during training but with the orientations of the Gabor rotated by  $90^\circ$  (e.g., in the categorization training group, the Gabor stimuli were drawn from a uniform distribution from  $120^\circ$ – $150^\circ$  and the reference angle was  $135^\circ$ ), and one block with a different task at the same orientation that they would be trained on (e.g., the categorization training group performed the estimation task with stimuli centered on  $45^\circ$ ). At least one day after the end of the last training block, all observers underwent three posttest blocks. The posttest consisted of the same two tasks as were performed during pretest, plus an additional block of the same task as was performed during training. The selected pre- and posttest conditions allowed comparison of the extent of specificity of perceptual learning (concerning the stimuli and the task participants were trained on) in either the estimation

Stage	Block	Training group		
		Categorization	Estimation	Moving categorization
Pretest	1	ET 30°–60°	ET 120°–150°	CT 30°–60°
	2	CT 120°–150°	CT 30°–60°	MT 120°–150°
Training	1–19	CT 30°–60°	ET 30°–60°	MT 30°–60°
Posttest	1	ET 30°–60°	CT 30°–60°	CT 30°–60°
	2	CT 30°–60°	ET 30°–60°	MT 30°–60°
	3	CT 120°–150°	ET 120°–150°	MT 120°–150°

Table 1. Training procedure. Participants performed either the categorization task (CT), the estimation task (ET), or the moving categorization task (MT) with different ranges of orientations of the stimulus Gabors.

task or the moving categorization task with the standard categorization task. Comparing the estimation task with the moving categorization tasks was considered less crucial. The detailed procedures in the three training groups are shown in Table 1.

## Assessment of learning

To assess learning on the trained task, for each participant, 79% accuracy thresholds were calculated in two ways, which differed between the two categorization tasks and the estimation task. For the categorization tasks, in the first method (block model), thresholds were calculated by fitting psychometric curves to each block of response data using the `glm()` function in R (McCullagh & Nelder, 1989; Yssaad-Fesselier & Knoblauch, 2006). Specifically, for each block, we randomly sampled the 200 stimuli and associated participants' responses with replacement 500 times. We fit a psychometric curve to each of these 500 samples, and compared the mean threshold during the first block to the mean threshold from the final block. A power function was then fit to the mean thresholds across blocks.

We also utilized a second fitting method designed to ameliorate two issues with the block-by-block fitting described above. The first is that the large number of blocks (19) entails a substantial number of free parameters (38 to estimate thresholds within the blocks, plus an additional set of three parameters to fit to the block estimates). This, in turn, leads to the concern that any estimates will be overfit. Furthermore, the block-by-block fitting implicitly instantiates an assumption that is almost certainly incorrect, namely that performance is static within a block and can only change between blocks. Thus, in a second method (time model), we fitted the proportion of clockwise responses,  $P(CW)$ , to a dynamic logistic function of orientations ( $x$ ) with the slope parameter  $\beta_1$  evolving linearly over time ( $t$ ), and a constant bias term  $\beta_0$  (see Equation 1 and Figure 3A). Estimated 79% thresholds (Figure 3B) were compared between the first and the last trial of the entire training.

$$P(CW) = \frac{1}{1 + e^{(\beta_0 - \beta_1 \cdot x)}} \quad (1)$$

with  $\beta_1 = \chi + \alpha \cdot t$ .

For the estimation task, we also fit a block model and a time model. In the block model, the absolute angular errors of each 200-trial block of the training were fitted by a simple power function and the mean threshold during the first block was compared to the mean threshold from the final block. In the time model, we fit a dynamically normal distribution to the absolute angular errors (with  $M_{error} = 0$  and  $SD_{error}$  evolving with time; see Equation 2) and compared the mean estimate of first training trial performance with the mean estimate of final training trial performance.

$$SD_{error} = \alpha \cdot e^{(\beta \cdot t)} + \alpha \quad (2)$$

## Assessment of transfer

Using the block fitting methods (above), performance thresholds were calculated for pretest and posttest blocks. To assess orientation transfer, posttest performance on the trained task and trained orientation was compared against posttest performance on the trained task and untrained orientation. To assess task transfer, posttest performance on the untrained task (trained orientation) was compared both to pretest performance (within an individual) and to the posttest performance of the group trained on the task (e.g., comparing posttest estimation task performance in the categorization trained group with posttest estimation task performance in the estimation trained group).

## Results

### Learning

First, we assess the extent to which perceptual learning was found over the course of training in all three groups, using both the block and time-varying

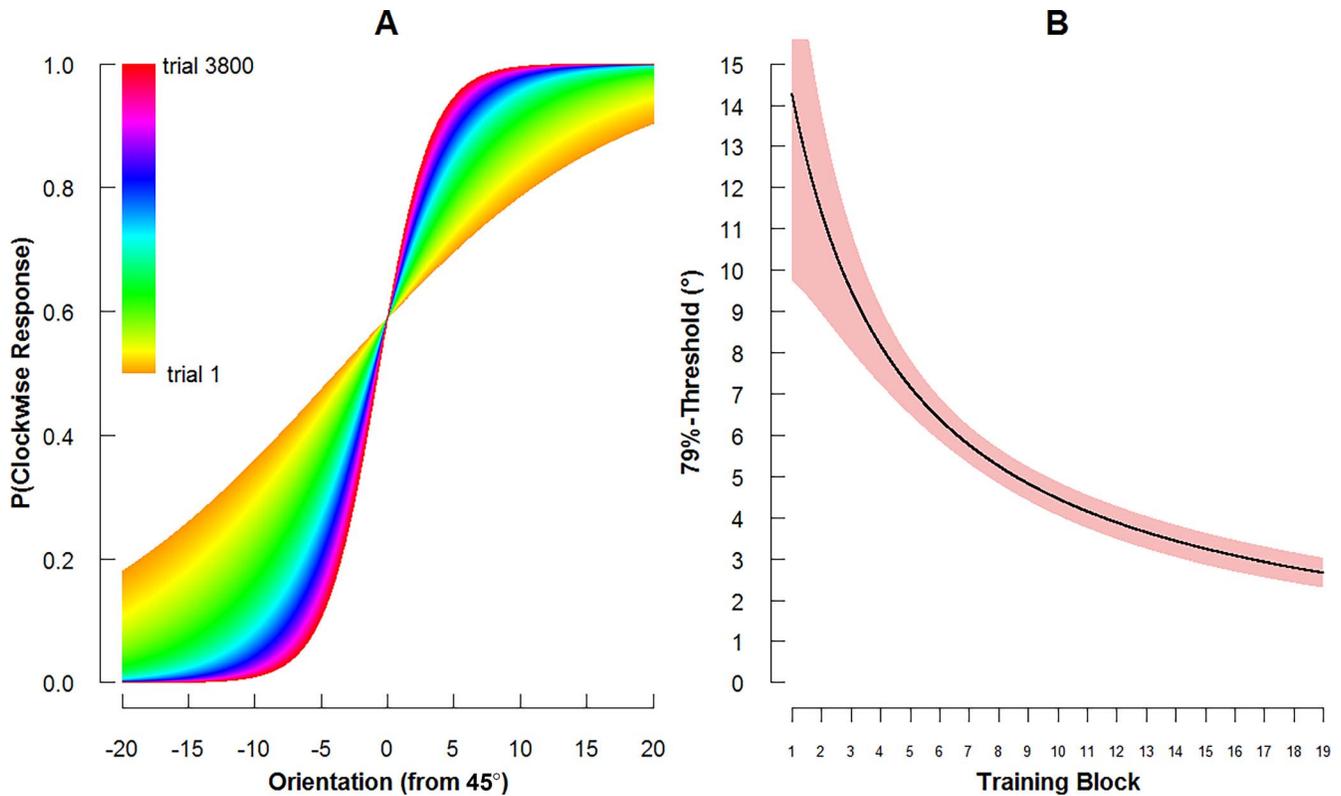


Figure 3. Illustrative example of the time-evolving psychometric function (A) and the resulting 79% threshold estimates (B) fitted to the data of subject 1 (categorization training group).

models discussed above. In the block model, this involved comparing estimated performance in the first block of training with estimated performance in the final block of training. In the time-varying model, this involved comparing estimated performance on the first trial of training with estimated performance on the final trial of training. No meaningful thresholds could be calculated for one participant in the categorization-trained group, and the following analyses are based on the data from the remaining participants only.

### Block model

The categorization group demonstrated a significant decrease in the average discrimination threshold from  $11.6^\circ \pm 26.4^\circ$  ( $MD \pm IQR$ ) in block 1 to  $MD = 4.0 \pm 1.7^\circ$  in block 19 (see Figure 4),  $V = 28$ ;  $Z = -2.42$ ;  $p = 0.007$  (Wilcoxon signed rank test, one-tailed). Similarly, the estimation group demonstrated a decrease in the average absolute error from  $8.1^\circ \pm 4.9^\circ$  in block 1 to  $5.1^\circ \pm 0.5^\circ$  in block 19 (see Figures 4 and S2),  $V = 28$ ;  $Z = -2.42$ ;  $p = 0.007$ . Finally, the moving categorization group also showed significant improvements in median thresholds across training with  $26.6^\circ \pm 33.5^\circ$  in block 1 and  $9.8^\circ \pm 3.8^\circ$  in block 19 (see Figures 4 and S3),  $V = 15$ ;  $Z = 1.86$ ;  $p = 0.03$ .

### Time model

The estimated 79% thresholds resulting from the time models for the orientation categorizations are illustrated as red lines in the supplementary Figures S1 and S3. Bootstrapped confidence intervals indicate that the thresholds decreased reliably over time. The best fits of the absolute errors in the estimation task obtained with the dynamic normal distribution model (with time-evolving  $SD$ ) are illustrated in Figure S2, which similarly indicate reliable improvements in estimation accuracy. The categorization group demonstrated a significant decrease in the average discrimination threshold from  $8.1^\circ$  ( $IQR = 5.48^\circ$ ) on trial 1 to  $3.8^\circ$  ( $IQR = 1.78^\circ$ ) on trial 3800 (block 19) (Figure S1),  $V = 28$ ;  $Z = -2.42$ ;  $p = 0.008$  (Wilcoxon signed rank test). Similarly, the estimation group demonstrated a decrease in the average absolute error from  $13.0^\circ$  ( $IQR = 17.8^\circ$ ) on trial 1 to  $6.9^\circ$  ( $IQR = 1.3^\circ$ ) on trial 3,800 (see Figure S2),  $V = 27$ ;  $Z = -2.15$ ;  $p = 0.016$ . Finally, the moving categorization group showed improvements (though nonsignificant) in average thresholds across training with  $32.5^\circ$  ( $IQR = 116.9^\circ$ ) on trial 1 and  $7.1^\circ$  ( $IQR = 1.6^\circ$ ) on trial 3,800 (see Figure S3),  $V = 10$ ;  $Z = -1.53$ ;  $p = 0.063$ .

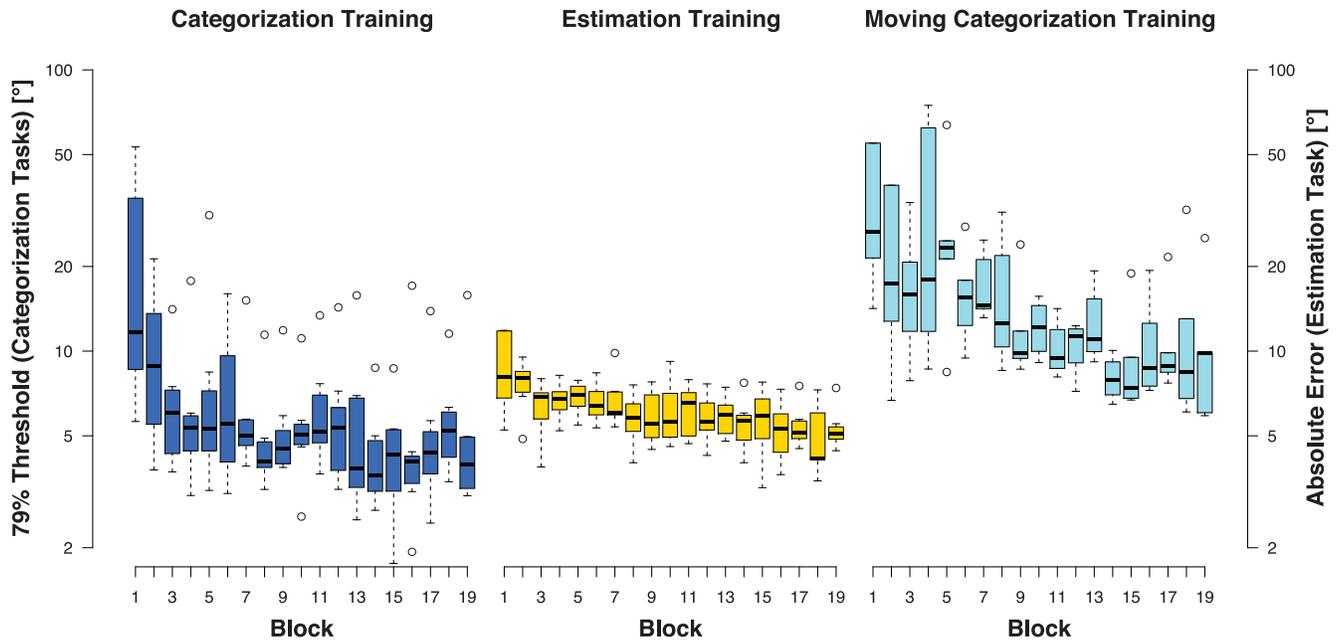


Figure 4. Overall learning curves (by block) across the three training tasks (black lines represent the medians, boxes indicate interquartile ranges, and whiskers refer to 1.5 interquartile ranges below and above the first and third quartile, respectively). Significant improvements were observed in all three conditions (left, categorization group; middle, estimation group; right, moving categorization group).

## Comparison of learning models

When evaluating learning, it is typical in the literature to first estimate thresholds separately by block, either by fitting a psychometric curve to data by block or by using a staircase procedure to estimate thresholds by block, and then to fit a time varying function (e.g., an exponential or power function) to those independently derived block thresholds. Here we have developed an alternative approach, wherein behavioral data is fit with a time-varying psychometric function, where the slope, or  $\beta_1$  component of the logistic function, changes linearly with time. The difference in free parameters in these two models is substantial. The block model is comprised of 19 separate psychometric functions with two free parameters each plus a power function with two free parameters for a total of 40 free parameters. The time-varying model is comprised of additive and multiplicative components of the time varying  $\beta_1$  parameter plus a constant  $\beta_0$  parameter for a total of three free parameters. Thus, the extent to which our time-varying model performs equivalently to the more standard approach is of significant interest.

The Akaike information criteria for the two different models are listed in Table 2. The time model, with three free parameters, reached better fits (relative to model complexity) than the block model in all participants that were trained on the categorization or moving categorization task. For the estimation training,

however, the difference between the block model and the time model was less obvious. In some observers, fitting the absolute angular errors block-by-block was superior to fitting a time-evolving normal-distribution model across the entire course of training.

## Orientation transfer

Next, we assess the extent to which the observed perceptual learning that occurred in each condition transferred to a new orthogonal orientation. Our prediction was that we would see significant orientation specificity in the categorization task (i.e., that performance on the trained orientation at posttest would be significantly better than performance on the untrained orientation), but less evidence for specificity in the estimation and moving categorization tasks (i.e., that there would be no significant differences between the trained and untrained orientations at posttest).

The 79% thresholds and absolute angular errors in the posttests, respectively, are depicted in Figure 5. Consistent with our predictions, in the categorization-training group, performance in the posttest block with the untrained orientation (median threshold = 8.6°;  $IQR = 1.0^\circ$ ) was significantly worse than in the posttest block with the trained orientation ( $MD = 3.4^\circ$ ;  $IQR = 0.7^\circ$ ),  $V = 0$ ;  $Z = -2.15$ ;  $p = 0.016$  (Wilcoxon signed rank test, two-tailed), indicating that categorization training produced perceptual learning that is at least partially

Subject	Training group					
	Categorization		Moving categorization		Estimation	
	$AIC_{Block}$	$AIC_{Time}$	$AIC_{Block}$	$AIC_{Time}$	$AIC_{Block}$	$AIC_{Time}$
1	132.97	104.34	157.90	124.78	116.47	118.30
2			178.38	153.22	120.03	145.04
3	140.74	103.88	172.55	138.83	78.59	91.54
4	146.61	136.40	163.00	129.82	258.04	228.03
5	91.80	61.48	179.06	156.83	104.41	105.36
6	184.00	148.93			102.85	124.07
7	124.41	98.70			102.25	95.49
8	110.98	80.88				

Table 2. Akaike information criteria (AIC) for the block model and the time model in the three training groups.

specific to the trained stimulus orientations. In addition, the difference between pretest ( $MD = 8.2^\circ$ ;  $IQR = 10.4^\circ$ ) and posttest thresholds (at  $135^\circ$ ) was not significant,  $V = 20$ ;  $Z = -0.89$ ;  $p = 0.188$  (one-tailed), suggesting that the 3,800-trial categorization training with a  $45^\circ$  boundary did, in fact, not lead to any detectable benefit to  $135^\circ$  orientation discrimination performance.

In contrast, in the estimation training group, there was no significant difference in performance between the trained (median angular error =  $5.7^\circ \pm 1.3^\circ$ ) and the orthogonal orientations ( $135^\circ$ ) at posttest (median angular error =  $5.4^\circ \pm 0.6^\circ$ ),  $V = 21$ ;  $Z = -0.53$ ;  $p = 0.297$ . Furthermore, posttest estimation errors at  $135^\circ$  were significantly reduced as compared to pretest errors ( $MD = 13.2^\circ$ ;  $IQR = 4.2^\circ$ ),  $V = 28$ ;  $Z = -2.42$ ;  $p = 0.008$  (one-tailed). Together, these results are also consistent with our hypothesis that estimation training would engender a significant degree of transfer.

Finally, in the moving categorization training group, posttest performance with orthogonal orientations was numerically worse ( $MD = 8.6^\circ$ ;  $IQR = 0.9^\circ$ ) than with the trained orientations ( $7.2^\circ \pm 1.0^\circ$ ); however, this effect did not reach significance,  $V = 1$ ;  $Z = -1.15$ ;  $p = 0.125$ . Moreover, the threshold of moving orientation categorizations (with a  $135^\circ$  boundary) significantly decreased from pretest ( $MD = 36.6^\circ$ ;  $IQR = 21.5^\circ$ ) to posttest,  $V = 15$ ;  $Z = 1.53$ ;  $p = 0.031$  (one-tailed), indicating that the  $45^\circ$  training did lead to improvements at a different boundary.

To further quantify the degree of specificity of learning across the three tasks, a specificity score was calculated by subtracting posttest performance (i.e., either 79% threshold or absolute error) with the trained orientation from posttest performance with the orthogonal orientations, standardized by the improvement from the initial training block to the posttest (Equation 3).

$$\text{Specificity} = \frac{\text{Post}_{120-150^\circ} - \text{Post}_{30-60^\circ}}{\text{Initial}_{30-60^\circ} - \text{Post}_{30-60^\circ}} \quad (3)$$

In the categorization training group, the average specificity score ( $MD = 0.56$ ;  $IQR = 0.41$ ) differed significantly from zero,  $V = 28$ ;  $Z = 2.15$ ;  $p = 0.016$ . In contrast, the average specificity score in the estimation training group ( $MD = -0.06$ ;  $IQR = 0.04$ ) did not differ significantly from zero,  $V = 7$ ;  $Z = -0.53$ ;  $p = 0.297$ . In the moving categorization training group, the degree of specificity lay in between the other two groups ( $MD = 0.10$ ;  $IQR = 0.06$ ), and it was not significantly different from zero,  $V = 14$ ;  $Z = -1.15$ ;  $p = 0.125$ , implying that this type of training produced some amount of transfer. Specificity scores differed significantly between the categorization training and estimation training groups,  $W = 44$ ;  $p = 0.011$ , as well as between the categorization training and moving categorization training groups,  $W = 35$ ;  $p < 0.003$ . The degree of specificity (or its absence) differed also between the estimation training and the moving categorization groups,  $W = 5$ ;  $p < 0.05$ .

## Task transfer

Finally, we examined the extent to which training on one task resulted in benefits on a different task (around the trained orientations; see Figure 6). This is of interest because many models of perceptual learning have, at their core, the idea that perceptual learning is supported by a change in the quality of sensory representations. If this were true, one would expect to see improvements on different tasks that utilize the same sensory stimuli. Conversely, our hypothesis, that perceptual learning is supported by a change in the ability to translate sensory stimuli into relevant actions or decisions, would predict limited to no task transfer. To test these alternatives, we compared performance on the untrained task at pretest with performance on the untrained task at posttest.

As predicted, the categorization training did not lead to a reduction of the absolute angular errors from pre- to posttest in the estimation task,  $V = 16$ ;  $Z = -0.24$ ;  $p = 0.406$  ( $MD_{pre} = 9.4^\circ$ ;  $IQR_{pre} = 4.1^\circ$ ;  $MD_{post} = 9.1^\circ$ ;

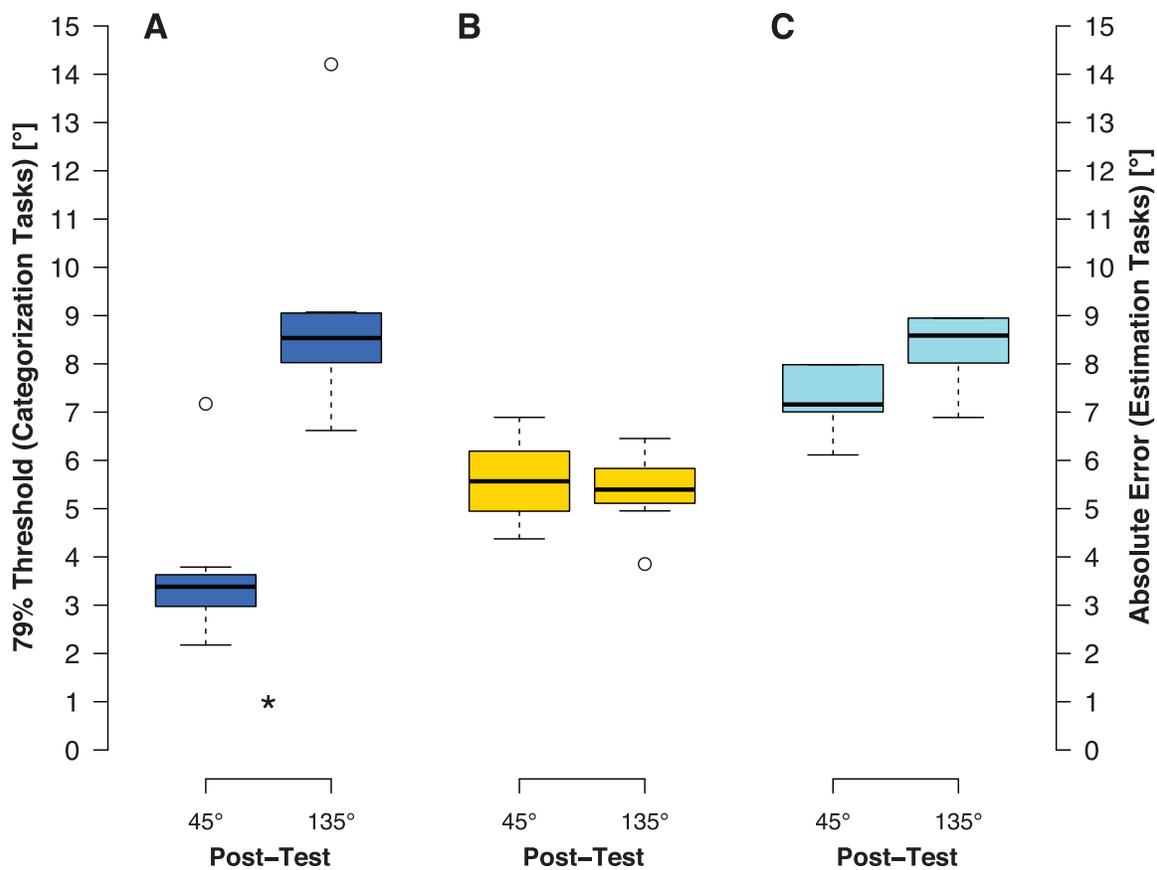


Figure 5. Orientation transfer. Posttest performance with trained and orthogonal orientations, trained task only. (A) Significant orientation specificity was observed in the categorization trained group, with posttest performance at the trained orientation (45°) significantly exceeding posttest performance at the untrained orientation (135°). (B) No significant orientation specificity was observed in the estimation trained group, with posttest performance on both the trained and untrained orientations being similar. (C) While participants in the moving categorization trained group were numerically slightly worse at the untrained orientation (by approximately 1°), this did not reach statistical significance.

$IQR_{post} = 2.7^\circ$ ), consistent with no task transfer. It was also the case that subjects in the categorization training group were significantly worse in the posttest estimation task (45°) than subjects in the estimation training group ( $MD = 5.6^\circ$ ;  $IQR = 1.2^\circ$ ),  $W = 49$ ;  $Z = -3.25$ ;  $p < 0.001$ . Likewise, in the estimation-training group, there were no significant improvements from pre- to posttest in orientation categorization thresholds,  $V = 15$ ;  $Z = -0.78$ ;  $p = 0.226$  ( $MD_{pre} = 23.0^\circ$ ;  $IQR_{pre} = 29.2^\circ$ ;  $MD_{post} = 10.0^\circ$ ;  $IQR_{post} = 41.9$ ) and posttest categorization thresholds (45°) differed significantly between the categorization-training ( $MD = 3.4^\circ$ ;  $IQR = 0.7^\circ$ ) and the estimation training groups,  $W = 1$ ;  $Z = -3.04$ ;  $p = 0.001$ . Finally, the moving categorization training did produce significant transfer to a nonmoving categorization task,  $V = 15$ ;  $Z = -1.86$ ;  $p = 0.031$  ( $MD_{pre} = 30.3^\circ$ ;  $IQR_{pre} = 24.5^\circ$ ;  $MD_{post} = 5.1^\circ$ ;  $IQR_{post} = 0.9^\circ$ ). The difference in posttest categorization threshold between the moving categorization training and the categorization training ( $MD = 3.4^\circ$ ;  $IQR = 0.7^\circ$ ) groups, however, was not significant,  $W = 23$ ;  $Z = -0.57$ ;  $p = 0.28$ .

## Discussion

As has been repeatedly observed throughout the perceptual learning literature, observers trained on a simple orientation categorization task showed evidence of significant orientation specificity when tested on the orthogonal orientation. Conversely, observers trained on either an orientation estimation task or a categorization task with a moving reference angle showed significant transfer to the orthogonal orientation. Critically, because the exact same stimuli were utilized for all three training groups (with the only difference being related to the way the sensory information had to be translated into a decision), the results cannot be explained by low-level stimulus attributes. Instead, this pattern of results is consistent with the emerging view that perceptual learning is mediated by rules that govern how higher-level integration areas read out sensory information in order to make particular decisions with the degree of learning transfer depending

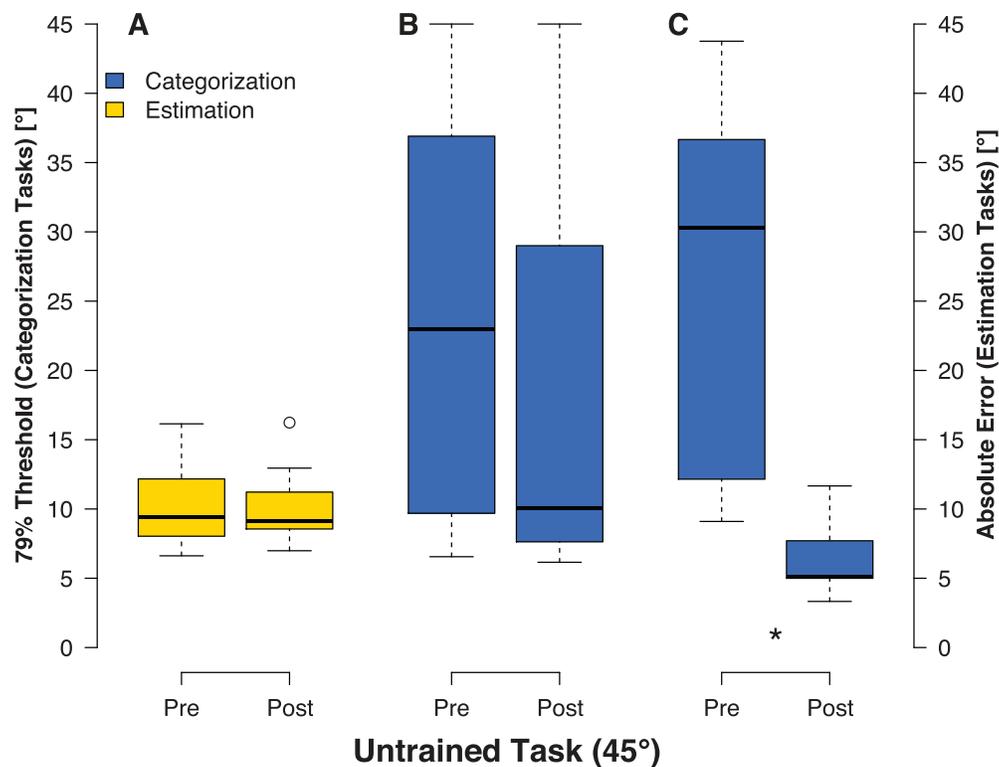


Figure 6. Task transfer. (A) No transfer from categorization training to the estimation task was observed. Participants in the categorization training group did not improve on the estimation task from pretest to posttest. (B) While numerically the estimation group did improve on the categorization task from pretest to posttest, this did not reach statistical significance. (C) Significant improvements on the categorization task from pretest to posttest were observed in the moving categorization task group. This was the expected effect, given that the categorization task is a subset of the moving categorization task.

on the extent to which those rules are applicable to new stimuli, tasks, or contexts (Bavelier et al., 2012; Bejjanki et al., 2011; Doshier & Lu, 1998; Kahnt et al., 2011; Law & Gold, 2008; Lu & Doshier, 2009; Zhang, Zhang et al., 2010).

Specifically, we have laid out a framework by which the type of rule that will be learned—and thus, the extent of learning generalization that will be observed—can be predicted a priori. In our model, the extent to which one or the other type of rule is learned depends on the optimal solution to the trained task. The optimal solution to a categorization task is a discriminant that separates the sensory space into two decision regions. However, such a discriminant has limited to no applicability to new categorization problems in which the boundaries are different. Thus, tasks with discriminants as optimal solutions should produce stimulus-specific learning. The optimal solution to an estimation task corresponds to a regression line that continuously maps stimulus orientation to a motor response. Because the relationship between stimulus orientation and motor responses holds over the range of orientations, such a regression line can be extrapolated to previously unseen orientations. Thus,

tasks with continuous regression-type mappings should produce significant transfer along that dimension.

An optimal-solution framework may also be able to account for the observation that perceptual learning on a target task (e.g., orientation discrimination) can be enhanced by an interleaved training on two tasks (e.g., blocks of spatial frequency and orientation discriminations), as compared to training on only the target task. Specifically, the interleaved training requires observers to not only learn a single orientation discriminant, but to map both orientation and spatial-frequency information to certain motor responses (Szpiro, Wright, & Carrasco, 2014). It may also account for the finding that specificity increases with longer training durations, in that longer training durations provide the additional needed samples that bias an optimal agent toward a mapping solution (Fulvio et al., 2014; Jeter et al., 2010). We believe that the more general framework (i.e., that the degree of transfer observed as a result of a learning experience can be predicted a priori given an understanding of the optimal solution to both the training and transfer tasks) can be extended to other important results in the field, for example, work showing that specificity is

dependent on the precision of the training and transfer tasks (Jeter et al., 2009; Hung & Seitz, 2014). However, it will be for future research to test this explicitly.

Interestingly, there was no task transfer from categorization training to the estimation task or estimation training to the categorization task. This is further evidence that the mechanism or mechanisms underlying improved performance via perceptual training include elements other than changes in sensory representations, as enhanced sensory representations should result in better performance across any tasks that utilize the same stimuli. This is of particular relevance in the case of the estimation training, which led to significant within-task, across-orientation transfer, but no across-task, within-orientation transfer. The only evidence for cross-task transfer was seen in the moving categorization trained group, which showed significant enhancements on the static categorization trained task following training. However, because the static categorization task is itself a subset of the moving categorization task, it is unlikely that this can truly be considered task transfer.

Outside of the empirical results, we also present a novel method of analyzing behavioral perceptual learning data. Specifically, we compared the fits derived from the standard approach to data analysis (fitting thresholds to blocks of trials and then fitting a time-varying function to the thresholds) with the fits given by a time-varying psychometric function. We found that our much simpler time-varying model (three free parameters compared to 40 free parameters in the block model) provided a better fit to the learning data when accounting for the difference in free parameters. In addition to providing a better fit to the data, the time-varying model has the additional benefit that it provides a continuous estimate of performance, thus implicitly recognizing that Trial 1 performance should be poorer than Trial 200 performance despite both being in the same block. Indeed, one failing of the block approach is that it implicitly assumes that performance is static within blocks, which is clearly false.

In all, our results add to a growing body of behavioral research suggestive of a higher-level neural locus for perceptual learning. This framework is also consistent with recent neurophysiological data—both in single-cell recordings in nonhuman primates, as well as in structural and functional neuroimaging in humans—wherein perceptual learning effects have been associated primarily with changes in higher-level association or integration areas, rather than lower-level sensory areas (Kahnt et al., 2011; Law & Gold, 2008; Li, Mayhew, & Kourtzi, 2009; Sathian, Deshpande, & Stilla, 2013).

*Keywords:* perceptual learning, categorization, estimation, orientation discrimination, transfer

## Acknowledgments

This work was supported by Office of Naval Research grants N00014-07-1-0937 and N00014-12-1-0883.

Commercial relationships: none.

Corresponding author: C. Shawn Green.

Email: csgreen2@wisc.edu.

Address: Department of Psychology, University of Wisconsin–Madison, Madison, WI, USA.

## References

- Ahissar, M., & Hochstein, S. (1997). Task difficulty and the specificity of perceptual learning. *Nature*, *387*, 401–406.
- Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences*, *8*(10), 457–464.
- Ahissar, M., Nahum, M., Nelken, I., & Hochstein, S. (2009). Reverse hierarchies and sensory learning. *Philosophical Transactions of the Royal Society B*, *364*, 285–299.
- Ball, K. K., & Sekuler, R. (1982). A specific and enduring improvement in visual motion discrimination. *Science*, *218*, 697–698.
- Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin*, *128*, 612–637.
- Bavelier, D., Green, C. S., Pouget, A., & Schrater, P. (2012). Brain plasticity through the life span: Learning to learn and action video games. *Annual Review of Neuroscience*, *35*, 391–416.
- Bejjanki, V. R., Beck, J. M., Lu, Z. L., & Pouget, A. (2011). Perceptual learning as improved probabilistic inference in early sensory areas. *Nature Neuroscience*, *14*(5), 642–648.
- Brainard, D. G. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436.
- 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*, 2889–2894.
- Das, A., Tadin, D., & Huxlin, K. R. (2014). Beyond blindsight: Properties of visual relearning in cortically blind fields. *The Journal of Neuroscience*, *34*, 11652–11664.
- Davidson, R. J., & McEwen, B. S. (2012). Social

- influences on neuroplasticity: Stress and interventions to promote well-being. *Nature Neuroscience*, *15*, 689–695.
- Dayan, P., & Daw, N. D. (2008). Decision theory, reinforcement learning, and the brain. *Cognitive, Affective, and Behavioral Neuroscience*, *8*(4), 429–453.
- DeLoss, D. J., Watanabe, T., & Anderson, G. J. (2015). Improving vision among older adults: Behavioral training to improve sight. *Psychological Science*, *26*, 456–466.
- Deveau, J., Ozer, D. J., & Seitz, A. R. (2014). Improved vision and on-field performance in baseball through perceptual learning. *Current Biology*, *24*, R146–R147.
- Dosher, B. A., & Lu, Z. L. (1998). Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting. *Proceedings of the National Academy of Sciences*, *95*, 13988–13993.
- Fahle, M. (2004). Perceptual learning: A case for early selection. *Journal of Vision*, *4*(10):4, 879–890, doi: 10.1167/4.10.4. [PubMed] [Article]
- Fahle, M., Edelman, S., & Poggio, T. (1995). Fast perceptual learning in hyperacuity. *Vision Research*, *35*, 3003–3013.
- Fiorentini, A., & Berardi, N. (1980). Perceptual learning specific for orientation and spatial frequency. *Nature*, *287*, 43–44.
- Fulvio, J. M., Green, C. S., & Schrater, P. R. (2014). Task-specific response strategy selection on the basis of recent training experience. *PLoS Computational Biology*, *10*(1), e1003425.
- Garcia, A., Kuai, S. G., & Kourtzi, Z. (2013). Differences in the time course of learning for hard compared to easy training. *Frontiers in Psychology*, *4*(110), 1–8.
- Glascher, J., Daw, N., Dayan, P., & O’Doherty, J. P. (2010). States versus rewards: Dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*, *66*, 585–595.
- Green, C. S., & Bavelier, D. (2012). Learning, attentional control, and action video games. *Current Biology*, *22*, R197–R206.
- Hillman, C. H., Erickson, K. I., & Kramer, A. F. (2008). Be smart, exercise your heart: Exercise effects on brain and cognition. *Nature Reviews Neuroscience*, *9*, 58–65.
- Hung, S. C., & Seitz, S. R. (2014). Prolonged training at threshold promotes robust retinotopic specificity in perceptual learning. *The Journal of Neuroscience*, *34*(25), 8423–8431.
- Jeter, P. E., Dosher, B. A., Liu, S. H., & Lu, Z. L. (2010). Specificity of perceptual learning increases with increased training. *Vision Research*, *50*, 1928–1940.
- Jeter, P. E., Dosher, B. A., Petrov, A., & Lu, Z. L. (2009). Task precision at transfer determines specificity of perceptual learning. *Journal of Vision*, *9*(3):1, 1–13, doi:10.1167/9.3.1. [PubMed] [Article]
- Kahnt, T., Grueschow, M., Speck, O., & Haynes, J. D. (2011). Perceptual learning and decision-making in human medial frontal cortex. *Neuron*, *70*(3), 549–559.
- Karni, A., & Sagi, D. (1991). Where practice makes perfect in texture discrimination: Evidence for primary visual cortex plasticity. *Proceedings of the National Academy of Sciences, USA*, *88*(11), 4966–4970.
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What’s new in Psychtoolbox-3? *Perception*, *36*(14), ECVP Abstract Supplement.
- Law, C. T., & Gold, J. I. (2008). Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. *Nature Neuroscience*, *11*, 505–513.
- Li, S., Mayhew, S. D., & Kourtzi, Z. (2009). Learning shapes the representation of behavioral choice in the human brain. *Neuron*, *62*, 441–452.
- Li, R. W., Ngo, C., Nguyen, J., & Levi, D. M. (2011). Video-game play induces plasticity in the visual system of adults with amblyopia. *PLoS Biology*, *9*(8), e1001135.
- Liu, Z. (1999). Perceptual learning in motion discrimination that generalizes across motion directions. *Proceedings of the National Academy of Sciences, USA*, *96*, 14085–14087.
- Lu, Z. L., & Dosher, B. A. (2009). Mechanisms of perceptual learning. *Learning and Perception*, *1*(1), 19–36.
- Mann, D. T., Williams, A. M., Ward, P., & Janelle, C. M. (2007). Perceptual-cognitive expertise in sport: A meta-analysis. *Journal of Sport and Exercise Psychology*, *29*(4), 457–478.
- McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models*. London: Chapman & Hall.
- McGovern, D. P., Webb, B. S., & Peirce, J. W. (2012). Transfer of perceptual learning between different visual tasks. *Journal of Vision*, *12*(11):4, 1–11, doi: 10.1167/12.11.4. [PubMed] [Article]
- Melby-Lervag, M., & Hulme, C. (2013). Is working

- memory training effective? A meta-analytic review. *Developmental Psychology*, *49*(2), 270–291.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.
- Polat, U., Ma-Naim, T., & Spierer, A. (2009). Treatment of children with amblyopia by perceptual learning. *Vision Research*, *49*, 2599–2603.
- Saffell, T., & Matthews, N. (2003). Task-specific perceptual learning on speed and direction discrimination. *Vision Research*, *43*, 1365–1374.
- Sagi, D. (2011). Perceptual learning in *Vision Research*. *Vision Research*, *51*, 1552–1566.
- Sathian, K., Deshpande, G., & Stilla, R. (2013). Neural changes with tactile learning reflect decision-level reweighting of perceptual readout. *The Journal of Neuroscience*, *33*, 5387–5398.
- Schellenberg, E. G. (2004). Music lessons enhance IQ. *Psychological Science*, *15*(8), 511–514.
- Schlickum, M. K., Hedman, L., Enochsson, L., Kjellin, A., & Fellander-Tsai, L. (2009). Systematic video game training in surgical novices improves performance in virtual reality endoscopic surgical simulators: a prospective randomized study. *World Journal of Surgery*, *33*, 2360–2367.
- Szpiro, S. F., Wright, B. A., & Carrasco, M. (2014). Learning one task by interleaving practice with another task. *Vision Research*, *101*, 118–124.
- Thorndike, E. L., & Woodworth, R. S. (1901). The influence of improvement in one mental function upon the efficiency of other functions. *Psychological Review*, *8*, 247–261.
- Wang, R., Zhang, J. Y., Klein, S. A., Levi, D. M., & Yu, C. (2012). Task relevancy and demand modulate double-training enabled transfer of perceptual learning. *Vision Research*, *61*, 33–38.
- Wang, R., Zhang, J. Y., Klein, S. A., Levi, D. M., & Yu, C. (2014). Vernier perceptual learning transfers to completely untrained retinal locations after double training: A “piggybacking” effect. *Journal of Vision*, *14*(13):12, 1–10, doi:10.1167/14.13.12. [PubMed] [Article]
- Xiao, L., Zhang, J., Wang, R., Klein, S. A., Levi, D. M., & Yu, C. (2008). Complete transfer of perceptual learning across retinal locations enabled by double training. *Current Biology*, *18*, 1922–1926.
- Yssaad-Fesselier, R., & Knoblauch, K. (2006). Modeling psychometric functions in R. *Behavior Research Methods*, *38*(1), 28–41.
- Zhang, J. Y., Cong, L. J., Klein, S. A., Levi, D. M., & Yu, C. (2014). Perceptual learning improves adult amblyopic vision through rule-based cognitive compensation. *Investigative Ophthalmology & Visual Science*, *55*, 2020–2030. [PubMed] [Article]
- Zhang, G. L., Cong, L. J., Song, Y., & Yu, C. (2013). ERP P1-N1 changes associated with Vernier perceptual learning and its location specificity and transfer. *Journal of Vision*, *13*(4):19, 1–13, doi:10.1167/13.4.19. [PubMed] [Article]
- Zhang, T., Xiao, L. Q., Klein, S. A., Levi, D. M., & Yu, C. (2010). Decoupling location specificity from perceptual learning of orientation discrimination. *Vision Research*, *50*, 368–374.
- Zhang, J. Y., Zhang, G. L., Xiao, L. Q., Klein, S. A., Levi, D. M., & Yu, C. (2010). Rule-based learning explains visual perceptual learning and its specificity and transfer. *The Journal of Neuroscience*, *30*(37), 12323–12328.